



HAL
open science

The evolving doublecortin (DCX) superfamily.

Orly Reiner, Frédéric M. Coquelle, Bastian Peter, Talia Levy, Anna Kaplan,
Tamar Sapir, Irit Orr, Naama Barkai, Gregor Eichele, Sven Bergmann

► **To cite this version:**

Orly Reiner, Frédéric M. Coquelle, Bastian Peter, Talia Levy, Anna Kaplan, et al.. The evolving doublecortin (DCX) superfamily.. BMC Genomics, 2006, 7, pp.188. 10.1186/1471-2164-7-188 . hal-00096552

HAL Id: hal-00096552

<https://hal.science/hal-00096552>

Submitted on 13 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Research article

Open Access

The evolving doublecortin (DCX) superfamily

Orly Reiner*¹, Frédéric M Coquelle^{1,5}, Bastian Peter², Talia Levy¹, Anna Kaplan¹, Tamar Sapir¹, Irit Orr³, Naama Barkai¹, Gregor Eichele⁴ and Sven Bergmann^{1,2}

Address: ¹Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel, ²Department of Medical Genetics, University of Lausanne, Switzerland, ³Department of Biological Services, Weizmann Institute of Science, Rehovot, Israel, ⁴Max-Planck Institute, Hannover, Germany and ⁵CNRS – UMR 6026, Université de Rennes 1, Equipe SDM, Campus de Beaulieu – Bât. 13, 35042 Rennes cedex, France

Email: Orly Reiner* - orly.reiner@weizmann.ac.il; Frédéric M Coquelle - frederic.coquelle@univ-rennes1.fr; Bastian Peter - Bastian.Peter@unil.ch; Talia Levy - TaliaL@weizmann.ac.il; Anna Kaplan - Anna.Kaplan@weizmann.ac.il; Tamar Sapir - tamar.sapir@weizmann.ac.il; Irit Orr - Irit.orr@weizmann.ac.il; Naama Barkai - naama.barkai@weizmann.ac.il; Gregor Eichele - gregor.eichele@mpihan.mpg.de; Sven Bergmann - Sven.Bergmann@unil.ch

* Corresponding author

Published: 26 July 2006

Received: 26 April 2006

BMC Genomics 2006, 7:188 doi:10.1186/1471-2164-7-188

Accepted: 26 July 2006

This article is available from: <http://www.biomedcentral.com/1471-2164/7/188>

© 2006 Reiner 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: Doublecortin (DCX) domains serve as protein-interaction platforms. Mutations in members of this protein superfamily are linked to several genetic diseases. Mutations in the human DCX gene result in abnormal neuronal migration, epilepsy, and mental retardation; mutations in *RPI* are associated with a form of inherited blindness, and *DCDC2* has been associated with dyslectic reading disabilities.

Results: The DCX-repeat gene family is composed of eleven paralogs in human and in mouse. Its evolution was followed across vertebrates, invertebrates, and was traced to unicellular organisms, thus enabling following evolutionary additions and losses of genes or domains. The N-terminal and C-terminal DCX domains have undergone sub-specialization and divergence. Developmental *in situ* hybridization data for nine genes was generated. In addition, a novel co-expression analysis for most human and mouse DCX superfamily-genes was performed using high-throughput expression data extracted from Unigene. We performed an in-depth study of a complete gene superfamily using several complimentary methods.

Conclusion: This study reveals the existence and conservation of multiple members of the DCX superfamily in different species. Sequence analysis combined with expression analysis is likely to be a useful tool to predict correlations between human disease and mouse models. The sub-specialization of some members due to restricted expression patterns and sequence divergence may explain the successful addition of genes to this family throughout evolution.

Background

Mutations in the X-linked gene *doublecortin* (*DCX*) result in lissencephaly or subcortical band heterotopia (SBH) [1,2] (see review in [3]). Lissencephaly is a severe brain

malformation characterized by absent (agyria) or decreased (pachygyria) convolutions accompanied by thickening of the cortex [4]. SBH is a related disorder in which there are bilateral bands of gray matter interposed

in the white matter between the cortex and the lateral ventricles [5]. In general, lissencephaly and SBH are neuronal migration disorders. SBH is very common among females with mutations in *DCX* [1,2]. *DCX* [6-9] and its closely related gene product DCLK (doublecortin-like kinase) [10,11] are classical microtubule (MT) associated proteins (MAPs), and contain two evolutionary conserved tubulin binding repeat sequences. We have initially defined the evolutionary conservation based on a limited number of proteins [12], following the expansion in available sequences, but the recent explosion of available sequence data, strongly warrants revisiting of this issue of the evolution of the DCX protein family. Interestingly, most missense mutations that are found in lissencephaly patients cluster in the defined DCX (doublecortin) domain [12,13]. These tandem repeat protein motifs are 11 kDa and are located in the N-terminal part of the 30 kDa DCX protein. The structure of the DCX domain markedly differs from that of classical MAPs adopting a globular structure with a ubiquitin-like fold [14]. Cryo-electron microscopy studies revealed that DCX binds to MTs at a novel site, located in between the protofilaments [15]. This binding site can explain why MTs are stabilized by DCX. Two additional genes of this superfamily are implicated in human diseases. RP1 is the protein product of retinitis pigmentosa -1 (*RP1*) [16,17]. Mutations in this gene result in progressive blindness, not only in humans, but also in a mouse model [18]. *DCDC2* has been implicated in dyslexia [19,20]. Reduction of *Dcdc2* in the developing cortex using *in utero* electroporation inhibited neuronal migration [19].

The notion that DCX domain proteins may function in a redundant way during cortical development was raised when *Dcx* null mice exhibited a very mild cortical phenotype [21], whereas acute knockdown of *Dcx* using RNAi resulted in a doublecortex phenotype [22]. Recent experimental evidence indicated that DCLK was redundant with DCX. *Dclk* *-/-* mice also exhibited a rather mild cortical phenotype, but *Dcx/Dclk* double mutant mice displayed severe cortical defects [23,24]. Recently, an additional member of this protein family, doublecortin kinase 2 (or DCLK2), has been described, and found to possess MT binding activities [25]. Nevertheless, *Dcx* and *Dclk* are but two genes among eleven DCX domain proteins in the mouse genome [26]. Thus, the quest for how DCX domain-containing proteins exert their multiple functions during cortical development is still open [27]. Our previous study has demonstrated common and unique activities in regard to protein interactions, MAP activity, and the subcellular localization of nine of the mouse DCX domain superfamily members [26]. Two unifying functional aspects were detected; all proteins affected MT polymerization, and all interacted with the JNK scaffold proteins, JIP1 and JIP2. Thus, the DCX superfamily of pro-

teins is likely to mediate signal transduction pathways. Proteins with a tandem DCX domain stabilized MTs in transfected cells. Some of the transfected proteins exhibited localization to actin-rich subcellular structures, and in addition most of the DCX domain proteins exhibited a nuclear localization. A distinct set of proteins interacted with the scaffold protein neurabin 2, which binds to PP1 [28] as well as to actin [29]. Positive interactions were observed with DCX, DCLK, and DCLK2, closely related family members, but not with other DCX-domain containing gene products.

In the current study, we studied the evolution of the DCX superfamily from unicellular organisms to humans. It was possible to demonstrate emergence and disappearance of specific genes or individual domains. The N-terminal and C-terminal DCX domains have undergone sub-specialization and divergence. In addition, the expression pattern of gene members was studied using database mining and *in situ* hybridization tools. The sub-specialization of some members due to restricted expression patterns and sequence divergence may explain the successful addition of genes to this family throughout evolution.

Results

Identification of proteins with a DCX domain

Human and mouse proteomes were searched for sequences similar to that of the human DCX domain yielding a total of 22 proteins containing one or two DCX repeats (Table 1, the complete sequences used in the present study are found in supplementary Fig. 1).

Serine/threonine protein kinase domains were found in three human/mouse proteins (DCLK, DCLK2, and DCLK3), and a ricin domain predicted to bind carbohydrates was found in a human/mouse protein referred to as FLJ46154 [30]. The structure of the human FLJ46154 and DCDC2B proteins differed from other proteins with tandem repeats; they contained a repeat more similar to DCX C-terminal repeat, which appeared in the N-terminal part of this protein, and a second repeat more similar to DCX N-terminal repeat. In the mouse orthologs of these two proteins, only one DCX domain was present. All the mouse genes reside in chromosomal regions (Fig. 1b), which are syntenic to the human orthologs (supplementary Fig. 2). This includes also the location of DCDC1 and BAC26042, however they are not true orthologs since the sequence similarity is very low (52%, among 46 out of 86 amino acids) only in the DCX domain, and the phylogenetic and evolutionary analysis, described below, indicate that they are different. BAC26042 is also unique in its close physical proximity with FLJ46154, the distance between these two genes being only 2 kb, suggesting they may share common regulatory elements.

Table 1: Sequence relationships between human and mouse proteins containing one or two DCX domains.

Symbol of the <i>Homo sapiens</i> proteins	Accession	Chromosome	# of DCX repeats	Symbol of the <i>Mus musculus</i> proteins	Accession	Chromosome	# of DCX	% similarity of orthologs
DCX	O43602	X	2	DCX	NP_034155	X	2	358/360 99%
DCLK	NP_004725	13	2	DCLK	NP_064362	3	2	681/702 97%
DCLK2	AAH32726	4	2	DCLK2	NP_081815	3	2	679/691 98%
RP1	AAD44198	8	2	RP1	NP_035413	1	2	544/675 80%
RP1L1	NP_849188	8	2	RP1L1	Q8CGM2	14	2	812/1440 56%
DCDC2A	AAF23612	6	2	DCDC2A	AAH45136	13	2	290/307 94%
DCDC2B	XP_497666	1	IC+IN	DCDC2B	XP_357395	4	INC	207/303 68%
DCDC2C	XP_371476	2	INC	DCDC2C	XP_917846	12	INC	107/145 73%
DCLK3	BAB21856	3	IC	DCLK3	NP_766516	9	IC	411/633 64%
FLJ46154	NP_940864	11	IC+IN	FLJ46154	XP_489892	2	IC	494/649 76%
DCDC1	P59894	11	IC	None				
None	-	-	-	BAC26042	BAC26042	2	IN	

This study is focused on the DCX domains, and does not cover the full-length proteins. Phylogenetic analysis was conducted for the individual DCX domains, separating the N- and C-terminal parts (Fig. 2). Several interesting features emerged from the human and mouse DCX domain phylogenetic analysis. The majority of human genes had a mouse ortholog. Two genes do not obey this rule as they do not have unambiguous orthologs (human DCDC1 and mouse BAC26042). Furthermore, in most instances the N-terminally located DCX domains were more similar to other N-terminal domains than to the C-terminal domains of the same protein. The two exceptions were already mentioned; human DCDC2B and FLJ46154. Sequence analysis combining BLAT [31] and phylogenetic analysis identified the orthologous relationships listed in Table 1.

Next, we extended the sequence analysis by including several additional non-mammalian genomes. Initially, the analysis encompassed proteins found in the conserved domain database CDD [32]. Subsequently, these searches were broadened with extensive BLAST, TBLASTN, and BLAT searches. Using BLAT search [31] sequences from opossum, rat, and rhesus monkey were added. *Ciona* sequences were added using TBLATN analysis against the genomic data, and only those sequences corresponding to ESTs were included. Hence, the present phylogenetic analysis included DCX-motif-containing proteins from human, chimpanzee, mouse, cow, dog, chicken, fish, worms, insects, frogs, fungi, and sea squirts (multiple alignments are provided in supplementary Fig. 3). The analysis of the tandem DCX domain proteins (67 proteins) resulted in an unrooted tree with bootstrap values shown in Fig. 3.

Four groups of proteins are easily categorized within the tandem DCX domain tree, which contains 67 proteins.

From top to bottom, the group of RP1 and RP1L1 include orthologs from the frog *Xenopus laevis*, fish (zebrafish *Danio rerio*, and pufferfish *Tetraodon nigroviridis*), chicken, cow, dog, mouse, rat, chimpanzee, and human. The second group includes proteins similar to DCDC2A (previously known as DCDC2, name approved by the HUGO Gene Nomenclature Committee) from mammals, including opossum (a marsupial), as well as chicken, fish, frog, and simpler organisms such as the ascidian *Halocynthia roretzi*, and the sea squirt *Ciona intestinalis*. The third group of proteins is devoid of mammalian proteins, but contains proteins from the social amoebae *Dictyostelium discoideum*, and one protein from *Ciona intestinalis*. Similar proteins were identified in the fruit fly, *Drosophila melanogaster*, the malaria mosquito, *Anopheles gambiae*, and the honey bee, *Apis mellifera*. Furthermore, two similar proteins from the worms *Caenorhabditis elegans* (ZYG-8), and *Caenorhabditis briggsae* are detected in this group. The fourth group of proteins includes those most similar to DCX, DCLK, and DCLK2. This group included mammalian, chicken, fish proteins, and one protein from and *Ciona intestinalis*. This analysis of proteins with two domains, was followed by an analysis for the N- and C-terminal domain proteins (supplementary Figs. 4-5). One hundred and seven proteins were analyzed in the N-group, and one hundred and one proteins in the C-group, suggesting that there are slightly more proteins similar to the N-terminal part of DCX. The general subdivision into the four groups was preserved. Inspection of the proteins composing the N-terminal phylogenetic tree detected that additional proteins were added mainly to the third group containing the *Dictyostelium discoideum* protein (including 8 members). Proteins from flies and worms were also added to this group. The fruit fly genome contains five DCX proteins, four of which are single repeats. Furthermore, several mammalian proteins were added to this group as well. This group was increased to contain 26

FIGURE 1

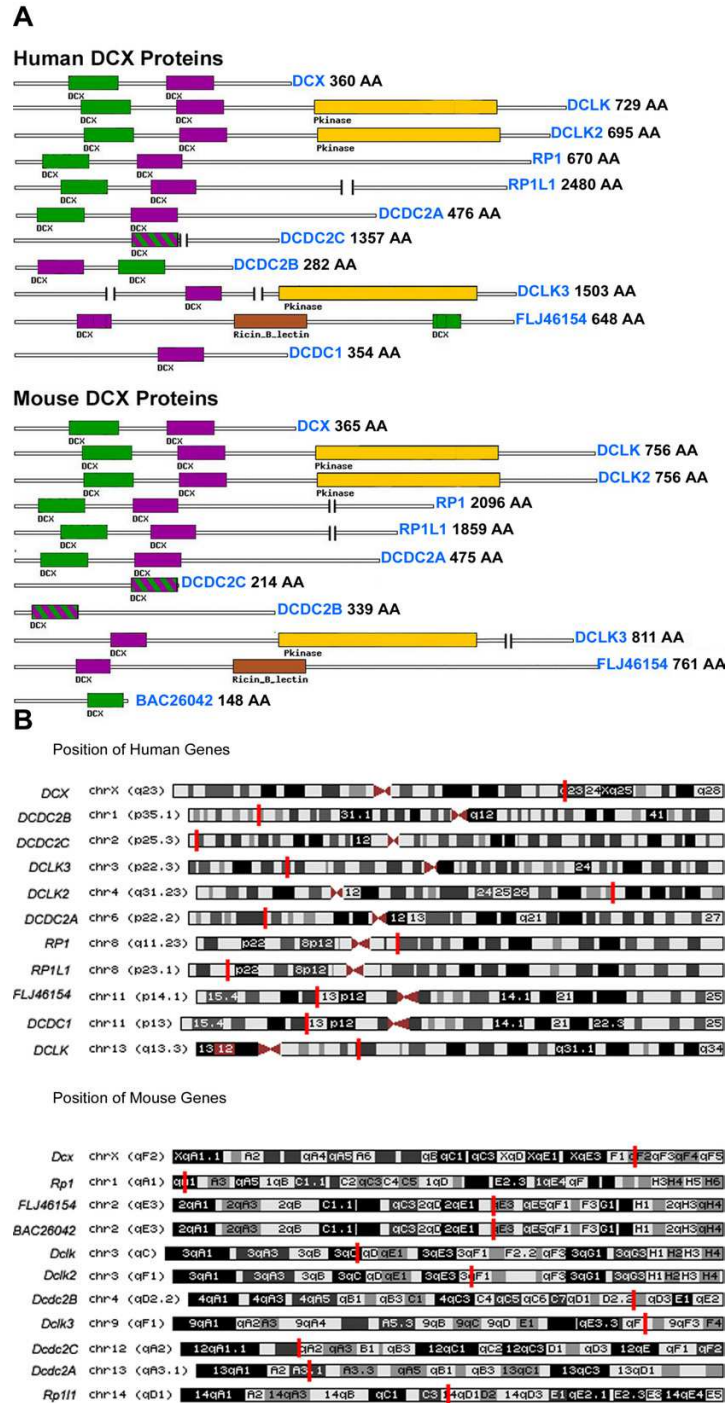


Figure 1

A. Schematic representation of human and mouse proteins containing DCX domains. DCX domains more similar to the N-terminal repeat of DCX were labeled in green, whereas those more similar to the C-terminal repeat were labeled in purple. Protein kinase domains are marked in yellow. The ricin domain is marked in brown. Human proteins are positioned on the top, the mouse proteins are shown below. B. Schematic presentation of the chromosomal location of the human and mouse DCX domain genes using UCSC site [51].

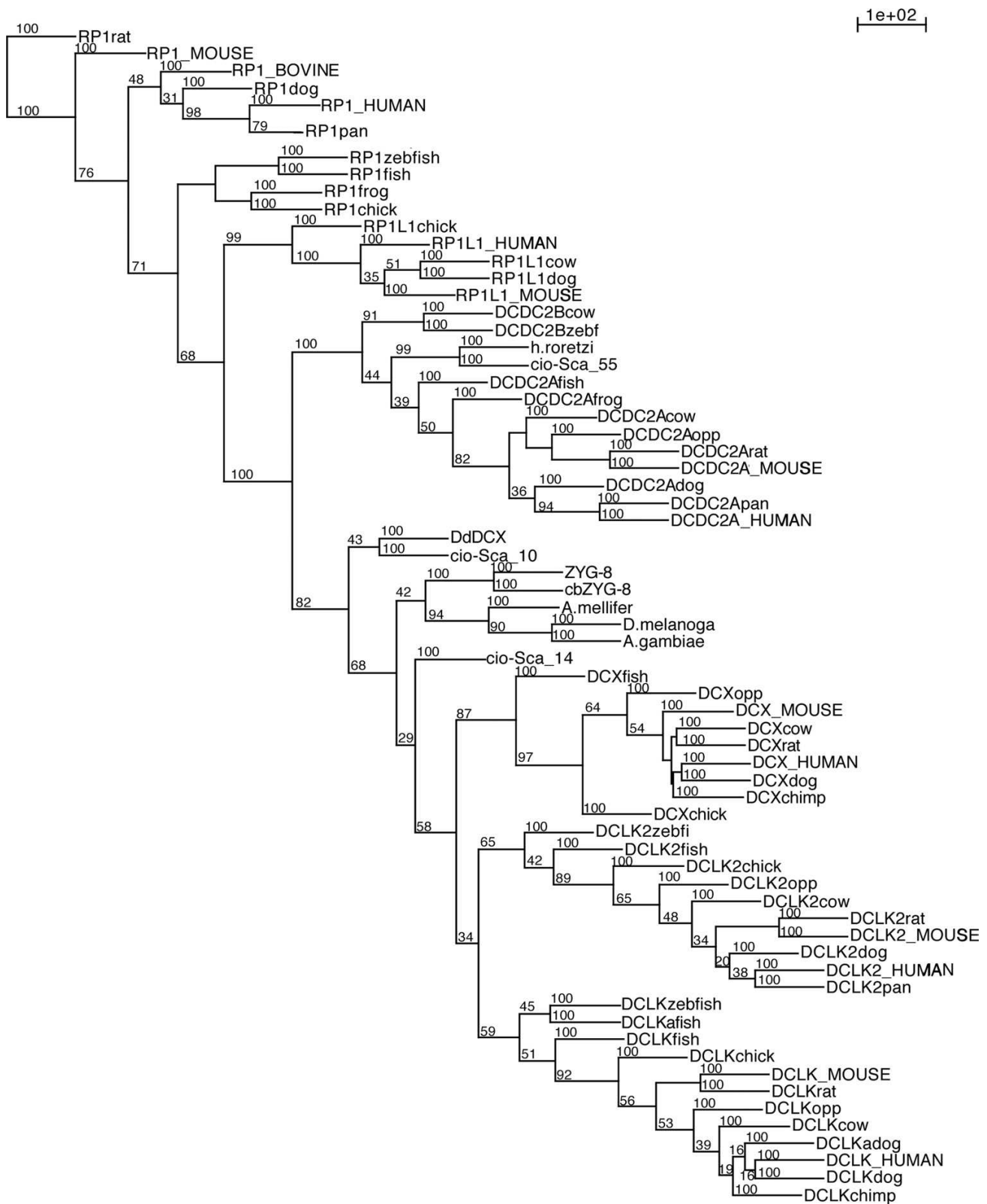


Figure 3
ML tree of the tandem DCX domain proteins from different species. Bootstrap values are indicated.

Table 2: Summary of the number of exons in DCX domains using BLAT [31]. The species analyzed are human (hum), chimpanzee (pan), rhesus monkey (rhe), dog, cow, mouse (mou), rat, opossum (opp), chicken (chi).

Gene	hum	pan	rhe	Dog	cow	mou	Rat	Opp	chi
<i>Dcx</i>	3e	3e	3e	3e	3e	3e	3e	3e	3e
<i>Dclk</i>	3e	3e	3e	3e	2e	3e	3e	3e	4e
<i>Dclk2</i>	4e	4e	4e	3e	4e	3e	3e	3e	4e
<i>Dclk3</i>	1e	1e	1e	1e	1e	1e	1e	1e	na
<i>Rpl</i>	2e	2e	2e	3e	3e	3e	2e	2e	1e
<i>Rpl11</i>	2e	1e	2e	1e	2e	2e	Na	1e	1e
<i>Dcdc2a</i>	5e	5e	5e	4e	5e	5e	5e	4e	5e
<i>Dcdc2b</i>	3e	3e	3e	3e	3e	2e	2e	1e	1e
<i>Dcdc2c</i>	1e	1e	1e	1e	2e	1e	1e	na	na
<i>Flj4615</i>	3e	3e	3e	3e	3e	2e	2e	na	na
<i>Dcdc1</i>	1e	1e	1e	1e	1e	na	Na	na	na
<i>Bac26042</i>	na	na	1e	Na	na	2e	1e	na	na

members in the N-group and 19 members in the C-group. This group included a protein from the unicellular organism *Plasmodium falciparum*, the malaria parasite.

Inspection of the proteins composing the C-terminal phylogenetic tree detected a group containing all the DCLK3 proteins. It should be noted that this group as a whole is quite distinct from DCX, DCLK, and DCLK2. Proteins in this group contain a single DCX domain from mammals (human, chimpanzee, cow, rat, and opossum), but also from fruit flies, honeybees, and malaria mosquitoes. An exception is the ciona protein demarking this group (Sca_10), which has a tandem repeat. One of the groups contain both DCDC2A and DCDC2B proteins, and yet an additional group contains several more DCDC2B proteins, suggesting probably less evolutionary conserved sequences in the C-terminal domains of this subset of proteins.

During the analysis of the DCX domain proteins, the presence of tandem or single DCX-domains was noted in corresponding orthologs. The simplest way to explain these differences may be through loss of intergenic sequences. The analysis of exon-intron boundaries included all the mammalian species and chicken since it is a non-mammalian vertebrate, close enough to mammal to make comparison possible (Table 2). In general, the location of the intron-exon boundaries is highly conserved. In some cases the presence of an additional exon, does not change the length of amino acids that are part of the DCX domains. Such is the case with DCDC2C; most species contain one exon, whereas the cow ortholog the corresponding amino acid sequence is divided into two exons. However, in most cases, the lack of an exon implies a reduction in the amino acid information. For example, FLJ46154 contains in most species three exons, whereas in mouse and in the

corresponding sequence in rat only two. Consequently, in mouse and rat only a single DCX domain was identified in the region corresponding to the human FLJ46154 DCX domains. This analysis also allows identifying key time points in the evolution of the DCX-domain proteins. The common vertebrate ancestor of mammals and birds is now believed to reach back 310 million years, marsupials split from the main (placental) group about 180 million years ago, and humans and rodents split off from their evolutionary family tree about 87 million years ago. The above analysis revealed that it is likely that BAC26042 was lost during evolution (in mouse two exons exist, while rat and rhesus monkey harbour only one exon). This analysis has been complicated due to a predicted sequence in rat (XM_230359) that is a fused sequence containing both FLJ46154 and BAC26042. However, we have experimental evidence that do not support the existence of this fused sequence. Antibodies we generated against the mouse FLJ46154 protein recognize a protein of the predicted size for FLJ46154 in mouse brain extract (supplementary figure 6). Thus, we have conducted our analysis based on the human data, which is derived from mRNA and EST data, and the mouse data that is based on EST data, supported by our experimental data. DCLK3 was generated after the mammals and birds split. BAC26042, FLJ46154, and DCDC2C were generated after the marsupials split from the main placental group. DCDC1 was generated after the humans and rodent split. According to this analysis the most conserved genes in this superfamily are DCX, DCLK, and DCDC2A.

Following analysis of the two groups including N- and C-terminal domains, analysis for all the DCX proteins was conducted (data not shown). As previously observed for the human and mouse proteins (Fig. 2), the N- and C-terminal domains were more similar to each other than to the corresponding repeat within the same protein. This result suggested that the DCX-domain duplications were ancient, and probably these two repeats have differed in their functions. Subspecialization of the N-terminal and C-terminal DCX motifs can be visualized at the level of logo sequences. Previously, four conserved blocks (A-D) within the DCX motif were identified [12], these conserved blocks are shown in the bottom of Fig. 4. When the N-terminal region was analyzed separately from the C-terminal region, it was obvious that the A and portions of B- and C- subdomains specify the N-terminus, while a portion of the C- subdomain specifies the C-terminus (Fig. 4). This result was obtained using the Lawrence Gibbs sampler motif-finding algorithm. Similar results were obtained with the Smith's MOTIF motif-finding algorithm (data not shown). This analysis indicates that although the tandem domains share a short sequence of similar amino acids, the N-terminal domain has a unique very conserved block of amino acids.

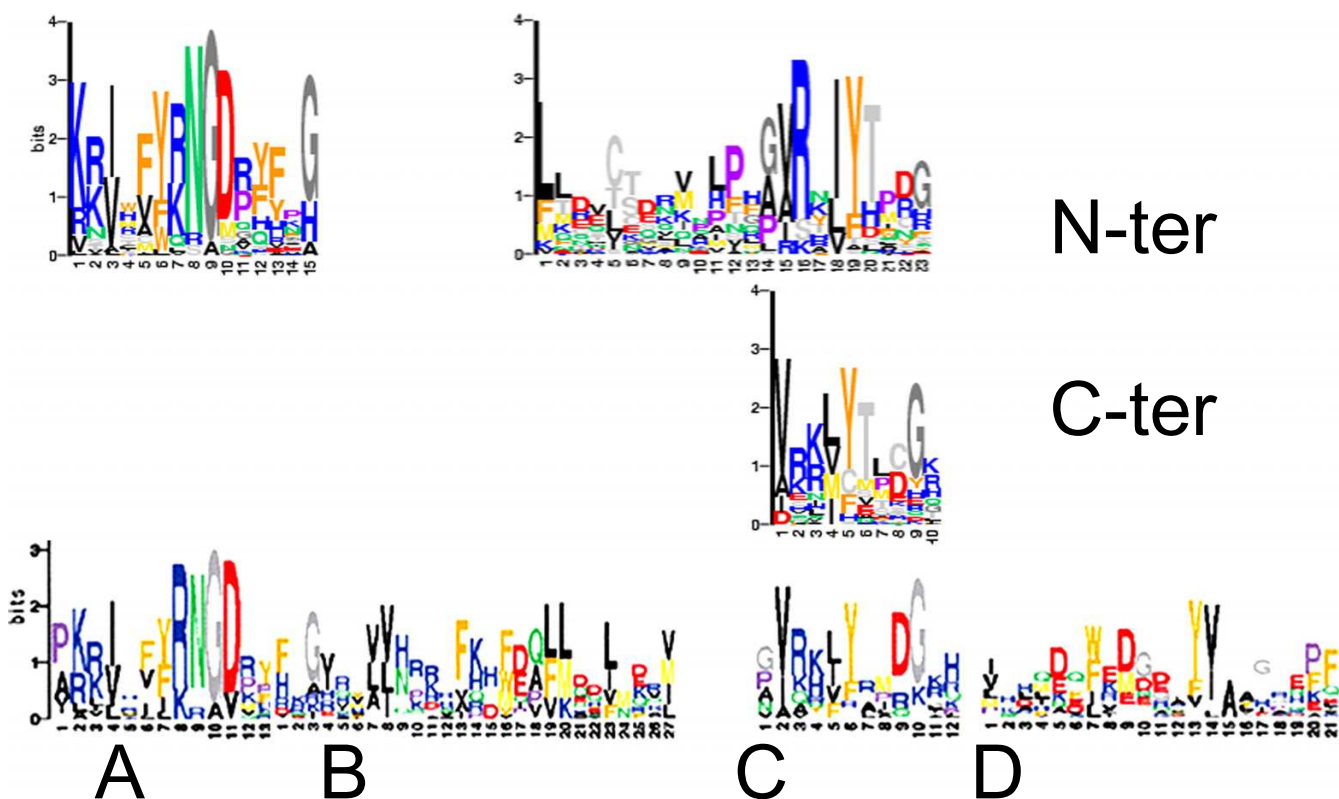


Figure 4
 Sequence logos of the N-terminal and C-terminal DCX motifs. Multiple alignments of the motifs from the DCX motifs are shown as sequence logos. The height of each amino acid represents bits of information and is proportional to its conservation at that position (y-axis), after the sequences have been weighted and frequencies adjusted by the expected amino acid frequency. Below the logos is the numbering of amino acids within the internal A-D subdomains. This SeqLogo represents the Lawrence Gibbs sampler motif-finding algorithm.

Expression analysis by in situ hybridization

Taken into consideration the similarities among the different DCX-domain paralogs, and their common functions in relation to signal transduction and microtubule regulation [26], it is important to establish when and where these genes are expressed. This will help in delineating their potential function. For example, the distinction whether a specific gene is expressed in proliferating, migrating, or differentiating cells is critical when trying to figure out gene function. Additionally, coexpression in a particular tissue may indicate that paralogs could cooperate or be redundant.

Our analysis was carried out by *in situ* hybridization at E14.5, a stage at which many differentiated cell types characteristic of an adult organism have formed yet at same time such mid-gestation embryonic tissues still contain progenitor cells. This analysis was performed with the goal to generate an expression profile "snapshot". With the exception of the ubiquitously expressed *Dcdc2B* (Fig. 5D), expression patterns of genes encoding DCX-repeat-

containing proteins are to a greater or lesser extent regional. *Dcx*, *Dclk* and *Dclk2* are expressed in the central and peripheral nervous system including the brain, spinal cord, cranial and dorsal root ganglia and in the parasympathetic ganglia (Fig. 5A-C). A high power view (Fig. 5E-H) shows that in the developing neocortex *Dcx* and *Dclk* transcripts are much more abundant in the preplate, but individual cells expressing the *Dcx* and *Dclk* genes can be detected in the ventricular zone. Both *Dclk2* and *Dcdc2B* are expressed in the developing neocortex, largely uniform and at low levels, but more pronounced in the ventricular zone than *Dcx* and *Dclk*. Outside the nervous system, prominent sites of *Dcx* and *Dclk* expression are the skeletal muscles, tongue muscles and individual cells of the olfactory epithelium (Fig. 5A,B). The latter tissue also expresses *Dclk2* (Fig. 5C).

BAC26042, *FLJ46154* and *Dcdc2A* exhibit highly regional expression patterns, which in the brain appear to be similar for *BAC26042* and *FLJ46154* (Fig. 5I-K). Fig. 5I and 5J show sagittal sections through the forebrain with

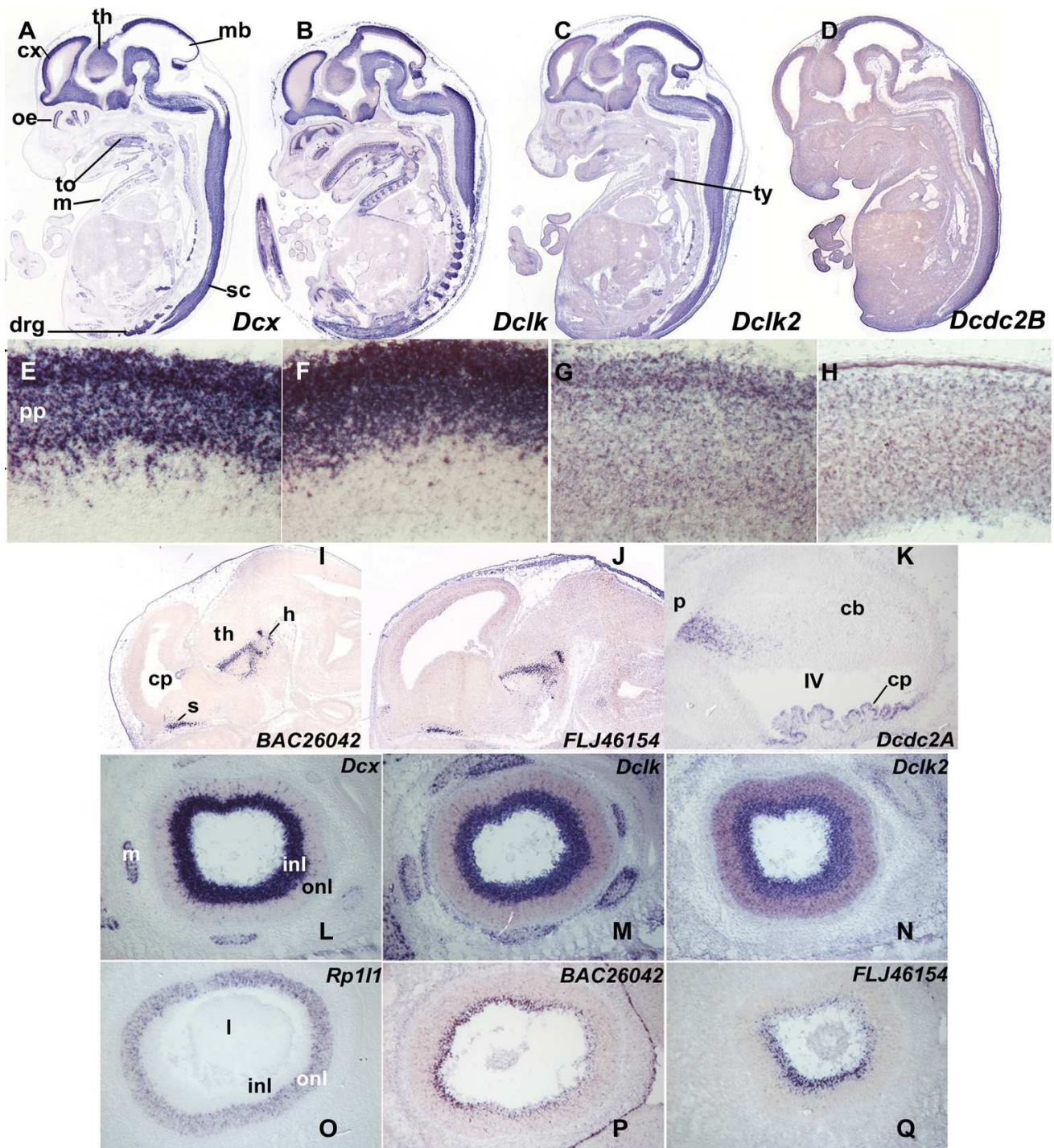


Figure 5
In situ hybridization patterns of genes containing the DCX protein domain. Blue stain denotes expression of the gene whose name is indicated in each panel. For details see Text. (A-D) Sagittal sections of whole E14.5 embryos. (E-H) Expression in frontal cortex. Note a few cells expressing in the ventricular zone. (I-K) Striking localized expression of *BAC26042* (I), *FLJ46154* (J), and *Dcdc2A* (K) in the E14.5 brain. In (I) "t" becomes "th" for thalamus. (L-Q) Expression patterns in the E14.5 eye. Abbreviations: cerebellum (cb), choroids plexus (cp), cortex (cx), dorsal root ganglia (drg), hypothalamus (h), inner neuroblastic layer (inl), midbrain (mb), muscles (m), olfactory epithelium (oe), outer neuroblastic layer (onl), pons (p), preplate (pp), septum (s), spinal cord (sc), thalamus (th), thymus (ty). tongue (t), ventricular zone (VZ).

FIGURE 6

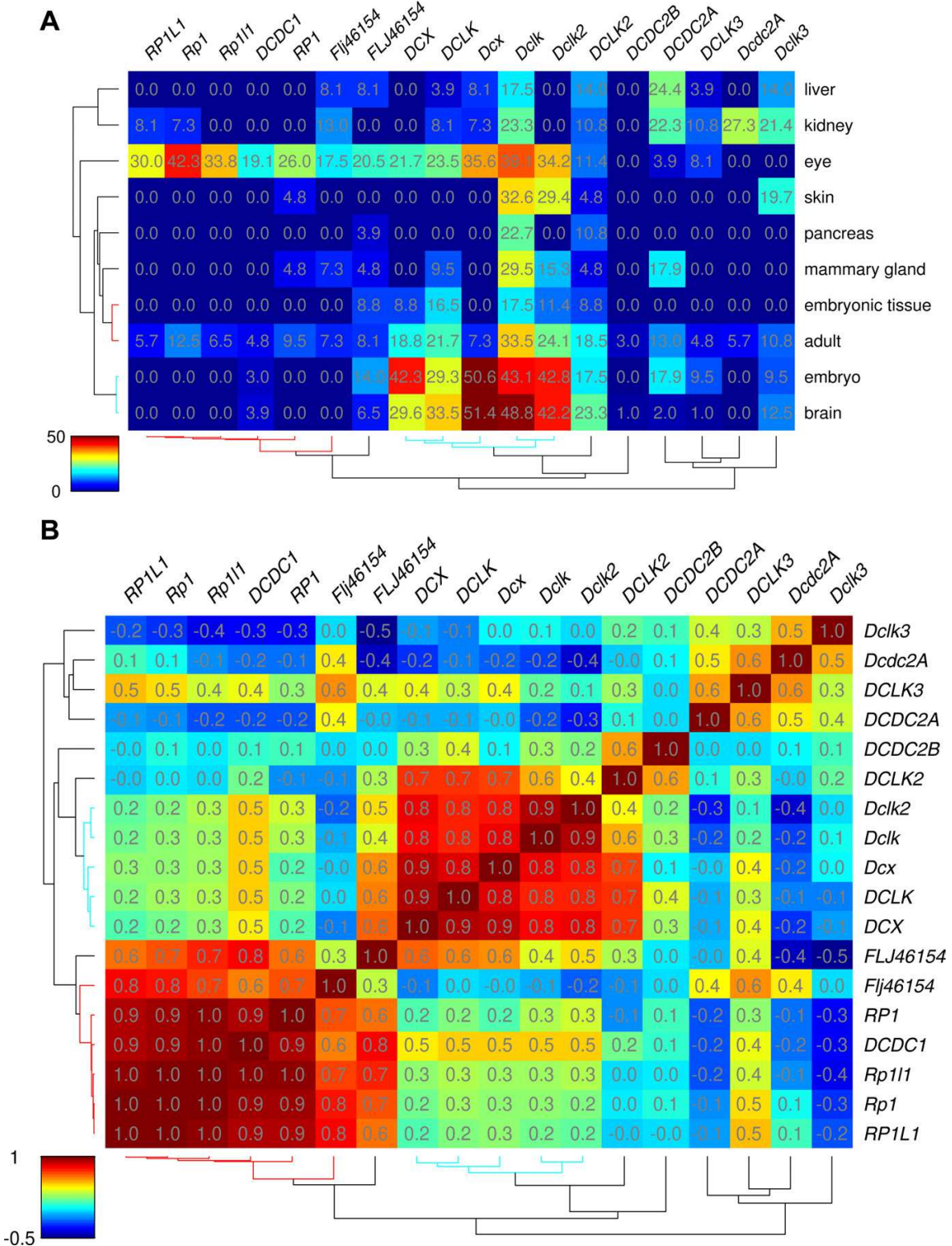


Figure 6
 A) Clustered Unigene [33] gene-tissue expression data. B) Gene-gene correlations based on Unigene expression data.

BAC26042 and *FLJ46154* transcripts present in the septum, various cell groups of the ventral thalamus, and in the posterior hypothalamus. Other sites of expression are a group of neurons at the base of the olfactory bulb (Fig. 5I,J), the pretectal area, the facial nucleus, and scattered neurons in the ventral and dorsal parts of the spinal cord (data not shown). *Dcdc2A* expression in the CNS is restricted to a group of scattered neurons in the lateral most part of the developing cerebellum (Fig. 5K). *BAC26042* and *Dcdc2A* are expressed in the choroid plexi (Fig. 5I,K).

The majority of DCX-repeat encoding genes are expressed in the developing retina. Three types of patterns emerge: *Dcx*, *Dclk*, *Dclk2* transcripts are strongly expressed in the postmitotic inner neuroblastic layer (Fig. 5L-N), whereas *BAC26042* and *FLJ46154* are also expressed in this layer, but in a more restricted fashion near and at its surface (Fig. 5P,Q). Finally *Rp111* transcripts are found in the outer neuroblastic layer that contains proliferating cells (Fig. 5O). Radially arranged *Dcx*, *Dclk* or *Dclk2*-expressing cells are detected in the outer neuroblastic layer which is reminiscent of the situation seen in the ventricular zone of the neocortex (Fig. 5E-G).

In addition, lung and kidney express *Dcx*, *Dclk* and *Dcdc2A*. *Dclk2* transcripts are also found in the developing ovary and weak expression is also seen throughout the kidney (data not shown).

Our analysis included most of the 11 genes listed in Table 1, the exceptions being *Dclk3*, and *Dcdc2C* for which we could not yet identify suitable templates. *Rp1* was also examined but it is not expressed at E14.5, except expression noted in some midline cells of the spinal cord (data not shown). To summarize our studies, we found that tissues destined to respond to electrical stimuli – central and peripheral nervous systems and skeletal muscles – represent the most striking sites of expression of DCX-repeat encoding genes. Outside these tissues, expression is mostly low and usually not regional, the exceptions being kidney and lung.

Expression analysis in human and mouse

The relevance of functional genomics approaches using mouse models for studying human diseases obviously depends on the similarity of gene expression in the two species. Thus, we compared the expression of the human members of the DCX gene superfamily investigated in this study with their mouse orthologs. For this purpose, we used the Unigene database of expression data website. Tissue-dependent expression profiles for both human and murine DCX repeat-containing proteins were generated from the EST count provided by UNIGENE [33]. Since the mouse-human comparison was a key feature, the analysis

was limited to tissues with a high total number of EST counts that were common to both organisms. We analyzed data for ten different human genes, and eight mouse genes. For two human genes there were no corresponding expression data in mouse: *DCDC2B*, which has a mouse ortholog that is not listed in UNIGENE, and *DCDC1*, which does not have a mouse ortholog. The clustered expression data resulting from this analysis is shown Fig. 6A and a gene-gene correlation based on this information is shown in Fig. 6B.

We tested the significance of the correlation by random permutation analysis. The correlations were re-calculated 1000 times after rescrambling for each gene independently all tissues at random. We found that all high correlation (>0.5) were significant ($p < 0.01$). Two clusters revealing very high correlation were observed. The largest group included human *RP1* and *RP1L1*, and their murine orthologs. In addition, *DCDC1*, which so far had been reported to be expressed mainly in testis, and embryonic brain [34], was included in this group. This group is characterized by high levels of expression in the eye, which is common amongst most DCX proteins, and has been noted in our *in situ* analysis. In addition to expression in the eye, these genes are expressed at lower levels only in a few other tissues. In this group there is no clear distinction in the gene-gene correlation in the expression in mouse and human. The correlation between the different members of this group is >0.9 in all cases. Both the human and the mouse *FLJ46154* are related to this group, however the correlation between the human and mouse *FLJ46154* is low (0.3). The protein-products of these two genes have also diverged, with a loss of a DCX domain in the mouse protein. Thus, it may be possible that there has been less conservation in the regulatory regions of these genes as well.

The second group exhibiting high gene-gene correlations includes the murine genes *Dcx*, *Dclk*, and *Dclk2*, and their human orthologs. Human *DCLK2* exhibited somewhat lower correlations with its mouse ortholog (0.4) than the other genes in this group. This may stem from its general overall lower expression levels (Fig. 6A). Our *in situ* data also indicated a high similarity in the co-expression of *Dcx*, *Dclk*, and *Dclk2*. Furthermore, our functional analysis [26] indicated that this group shares more properties and only they interact with the scaffold protein neurabin 2. A third group of genes with lower levels of correlation include *DCDC2A*, *DCLK3*, *Dcdc2A*, and *Dclk3*. In this group the correlation between the corresponding orthologs does not exceed 0.5. It should be noted that there are some additional high correlations between different genes, for example; *DCLK3* and *Flj46154*, or *FLJ46154* with *DCX*, *DCLK*, and *Dcx*.

Discussion

Protein architecture

Number of doublecortin repeats

The DCX domain family of proteins consists of multiple members in the animal kingdom from the unicellular organisms *Plasmodium* and *Dictyostelium* (in the non-aggregated part of its life cycle) to human. Although, *Plasmodium* and *Dictyostelium* do not have a nervous system, motility is one of their important characteristics. The shift from a single to a tandem DCX repeat (or *vice versa*) is a rather simple mode to modifying protein activity from binding to tubulin and assisting MT polymerization (one repeat) to enabling it to promote MT bundling (two repeats). We have demonstrated experimental evidence for this view for DCX [12], and for eight additional members of the mouse DCX domain proteins [26]. The duplication of a single DCX motif was an ancient event because of the existence of a tandem repeat already in *Dictyostelium*, and since the similarity within the N-terminal domains is generally higher than to the C-terminal domains of the same protein. We were able to document the loss of a domain, by the loss of an exon ("domain death"). Interestingly, the group of proteins encoding a single repeat is expanded in nematodes. These single repeat proteins are usually similar to the same extent to the groups of the N- and C-terminal repeats. Our previous analysis [12], structural analysis of DCX [14,15,35], and our current analysis reveals differences between the N-terminal and the C-terminal domains. Our analysis detected a short stretch of conserved amino acids common to both N-terminal and C-terminal repeats, however the sequences diverged and the N-terminal domain has a unique block of conserved amino acids. Close inspection of the phylogenetic trees revealed higher evolutionary conservation among proteins in the N-terminal group than those in the C-terminal group. Our analysis detected that both the N-terminal and C-terminal domains of DCX promote assembly of MTs [12], however the N-terminal portion has higher activity than the C-terminal one. In addition, other proteins containing a single or tandem DCX domain were capable of promoting MT assembly [26]. Among the studied mouse proteins FLJ46154 exhibited the highest activity in promoting MT assembly [26]. This protein contains a single DCX domain, which is more similar to the C-terminal domain of DCX. Future studies directed at the structural analysis of the different DCX domains combined with the sequence analysis conducted here may provide the basis for elucidating these functional variations. Another study indicated that the N-terminal domain of DCX binds only to assembled MTs, whereas the C-terminal domain binds both to assembled MTs and to unpolymerized tubulin [14]. Therefore, the conserved amino acid block common to both repeats may be important in directing the binding to assembled MTs.

Additional domains

A significant portion of the DCX-domain containing proteins has a kinase domain attached to them. This group includes DCLK, and DCLK2, in which a tandem DCX repeat is detected, and DCLK3, which contains a single repeat more similar to the C-terminal one. Orthologs of these proteins are also found also in invertebrates including nematodes. The *C. elegans* ortholog ZYG-8 has been studied extensively [36]. In nematodes, ZYG-8 was found to be important for assembly of astral microtubules. The mutant ZYG-8 phenotype was observed with several different alleles including mutations in the DCX domain, and the kinase domain, therefore suggesting a role for kinase activity in regulating MT assembly. So far, the endogenous substrates of these kinases, besides DCLK which undergoes autophosphorylation [37], are unknown.

Addition (or reduction) of the number of domains is not always conserved in evolution. In close relationship with DCLK3, which contains a kinase domain, a branch including proteins with a different domain is found. This branch includes the *Drosophila melanogaster* protein (EMAL_DROME), which contains an additional HELP domain, and WD repeats, which can be detected in the closely related *Anopheles gambiae*, and *Apis mellifera* proteins. The HELP (Hydrophobic ELP) motif is found in EMAP and EMAP-like proteins (ELPs), and has been found to mediate binding to MTs *in vitro* [38,39]. Therefore, the combinatorial possibilities of modifying or changing functional activities will depend upon specific amino acid substitutions within the DCX domains, and addition of other functional domains.

Additional ways to add on functions

On the level of the whole organism, one of the best ways to modify protein activity is to vary the expression pattern. Indeed, even in the case of proteins with similar sequence and similar expression pattern such as in the case of DCX, DCLK, and DCLK2 some notable differences in expression pattern were noted. For example, the expression of DCLK2 in the ventricular zone of the cerebral cortex is much more obvious than that of DCLK, suggesting that the former may have a pronounced role in regulation of proliferation of neurogenic progenitors. In fact, the involvement of DCLK in regulation of mitosis has been demonstrated recently [40]. DCLK regulated the formation of bipolar mitotic spindles and the proper transition from prometaphase to metaphase during mitosis in HEK293 cells and neuronal progenitors. In cultured cortical neural progenitors, DCLK RNAi disrupted the structure of mitotic spindles and the progression of M phase, causing an increase of cell-cycle exit index and an ectopic commitment to a neuronal fate [40]. As mentioned above, the *C. elegans* ortholog of DCLK, ZYG-8 regulates assembly of

astral microtubules [36]. Taken together, these findings suggest that regulation of the MT-based mitotic spindle may represent a conserved function among DCX-domain proteins.

Two genes FLJ46154, and BAC26042, exhibited very similar expression patterns in the mouse. In the mouse genome, the distance between these two genes is only 2 kb, suggesting that they may share common regulatory sequences. Perhaps due to such a highly similar expression pattern, the loss of BAC26042 in other mammalian species is not surprising. Partial copies of this gene (containing only one exon instead of two) exist in rat and in the rhesus monkey. A proposed unified nomenclature for the doublecortin superfamily based on correspondence with the mouse gene nomenclature committee is presented in supplementary figure 7.

Expression patterns and phylogenetic correlations

One of the most exciting outcomes of the present analysis is that in some cases we observed a high correlation between orthologs and paralogs, and their expression profiles. This was most striking in case of the human and mouse *RP1*, and *RP1L1*. A possible hypothesis, based on these results is that a high correlation in the expression of the human and mouse genes may be indicative of a high probability that a mouse mutant would model the human disease (or *vice versa*). If we review this hypothesis within the limited genetic data available, we see that this hypothesis is supported. *Rp1*^{-/-} mice closely reflect very well the progressive blindness observed in human patients [16-18,41,42]. Furthermore, in these mice the JNK pathway was affected [43], and our functional analysis describe a possible connection between *RP1* and the JNK pathway [26]. An unexpected member to this group is *DCDC1*. The gene-to-gene function correlations are also very high in case of the human and mouse *DCX*, *DCLK*, and *DCLK2* genes. They are found on adjacent branches in the phylogenetic analysis. These genes are highly expressed in the central nervous system not only in human and mouse, but also in additional animals as chicken (*DCX*, *DCLK* [44], and *DCLK2* (unigene data for *Dcx* Gga.2608, *Dclk2* Gga.16742), cow (*dclk* and *dclk2*; Bt.34533, Bt.55548), rat (*dcx* Rn.121471, *dclk* Rn.155540, *dclk2* Rn.23327), and dog (*dcx* Cfa.19843, *dclk* Cfa.9988). Our analysis detected expression of some of these genes in several additional tissues such as muscle, heart, and kidney, sites which have not been previously described. Although *Dcx*^{-/-} mice have minor neuronal abnormalities [21,45], they do not adequately recapitulate the human disease. This is due to gene redundancy because the double mutant mice *Dcx*^{-/-} and *Dclk*^{-/-} exhibited perinatal lethality accompanied with multiple brain abnormalities [23,24]. However, using *in utero* electroporation in the rat and mouse embryo, reduction in the expression of *Dcx* resulted in

apparent inhibition of migration of neurons in the cerebral cortex [22,46]. In the rat, it was possible to observe heterotropic positioning of neurons in postnatal rat brains similar to SBH in human [22]. In contrast to the cases discussed, there are several genes for which the orthologs exhibit low gene-gene correlation; the lowest are *DCLK3* and *FLJ46154* (0.3), *DCLK2* (0.4), and *DCDC2A* (0.5). Among these genes variability in sequence analysis was noted as well. With future expansion of available information in databases, this analysis will be extended to include many more tissues, addition developmental stages, and additional species. We suggest that this analysis may be useful in predicting possible functional redundancy. Furthermore, it may provide useful clues in which tissues the mouse gene expression is most similar to its human ortholog, thus suggesting the possible relevance of a corresponding mouse model for a human disease.

Conclusion

In summary, the DCX-domain family of proteins has proven to be a very successful and prolific protein family involved in signal transduction and cytoskeletal regulation. Modifications in regulatory sequences allowing unique expression patterns, addition of new domains, and modifying the amino acid sequence of the DCX domains themselves explains – at least in part- why this superfamily is thriving.

Methods

Database homology search and phylogenetic analysis

Our database similarity search initiated with an NCBI blast search using the sequence of mouse doublecortin protein (NP_034155.2). In addition, all proteins with the DCX domain were retrieved from the CDD, (Conserved Domain DB). In a similar fashion, DCX proteins from the EBI InterPro db, which were not detected in CDD were extracted and added. In order not to miss related orthologs, BLAT analysis using UCSC BLAT site [31] was conducted initiating with the human and mouse genes. Definition of orthologs was based on positioning in syntenic chromosomal localizations. In addition, an exhaustive TBLASTN search was conducted to detect DCX-related transcripts in the EST database of *Ciona intestinalis* (sea squirt) [47]. The corresponding cDNAs were mapped to their corresponding genomic localizations.

For our analysis, three datasets were generated, the first dataset for the proteins with 2 DCX motifs (including both the N-term and C-term repeats) composed of 57 proteins from human, mouse, chicken, bovine, ciona, frog, fish, nematodes, chimp and flies.

The second dataset included the proteins with a single DCX motif, which exhibited higher similarity to the

mouse doublecortin C-terminal DCX motif, and the C-terminal ones of those with two DCX motifs. The single repeats exhibited which exhibited identical similarity to both the N- and C-terminal repeats of mouse DCX were included in both N-term and C-term groups. The comparison was based on two-way BLAST analysis using blast2seq. This dataset integrated 77 proteins from human, mouse, chicken, ciona, dog, bovine, frog, fish, sea squirt, nematode, ascidian, fruit fly and fungus.

The third dataset was composed of the proteins with a single DCX motif, which was more similar to the mouse doublecortin N-terminal DCX motif, and the N-terminal ones of those with two DCX motifs. This dataset included 85 proteins from the species mentioned above.

Multiple alignment program, Clustalw, was run on each of the datasets and the regions corresponding to the DCX motifs were extracted, resulting in three multiple alignments (data sets are deposited in supplementary data 1).

The Phylip package ver. 3.6 was used to build the ML trees with boot strapping values.

One hundred datasets were generated using the program SEQBOOT from the original data (multiple alignment done by clustalw) for each of the 3 groups of motifs (2-repeats, C-terminal repeats and N-terminal repeats). This was followed by running the program ProML (protein Maximum Likelihood) on each of the datasets in the group, using the JTT model. A consensus tree (from all the 100 trees) was generated using the program CONSENSE.

The trees were drawn using NJPLOT program or TreeView [48]. All sequences of the three original datasets are available in supplementary Fig. 1.

Separate datasets were generated for the human and mouse proteins, the first dataset for the proteins that have two DCX motifs (N-terminal and C-terminal) composed of 13 proteins. The second dataset for the proteins with a single DCX motif more similar to the mouse doublecortin C-term DCX motif, and the C-terminal ones of those with two DCX motifs. This dataset included 20 proteins. A third dataset included the proteins with a single DCX motif more similar to the mouse doublecortin N-term DCX motif, and the N-terminal ones of those with two DCX motifs. This dataset has 21 proteins. These datasets were subject to all the programs as mentioned above, clustalw, seqboot, proml, consense and njplot. CONSENSE and NJPLOT were used to draw the human-mouse phylogenetic trees.

Multiple alignment was done by Block Maker web server at the Fred Hutchinson Cancer Research Center in Seattle,

Washington, USA [49]. The Logoshown represents the Lawrence Gibbs sampler motif-finding algorithm, similar (but not identical) results were obtained using Smith's MOTIF motif-finding algorithm.

Extraction of public large-scale expression data and cluster analysis

EST counts from different human and murine tissue samples were extracted for 10 human and 8 mouse genes from the UNIGENE website [33]. We discarded all human samples with less than 150,000 total EST counts and retained only the murine samples that could be matched to these samples. The murine 'late gestation' sample was matched to the human 'embryo'. Expression value are given as "counts per million of total ESTs". We applied a scaling transformation ' $x \rightarrow t \cdot \operatorname{arcsinh}(x/t)$ ' to the raw expression values x to scale down (logarithmically) values above the threshold $t = 10$.

In order to identify similarly expressed genes the combined set of expression data was clustered using Pearson correlations over all tissues as similarity measures and single linkage for the generation of the dendrograms. Similarly, the tissues were clustered based on similar expression across the genes. Alternative clustering methods were tested and yielded similar results. We used the standard Matlab clustering algorithm.

In situ hybridization

The antisense RNA templates for *in situ* hybridization were generated by *in vitro* transcription using PCR products (using the appropriate combinations of T7, T3, and SP6 primers) from the corresponding genes, which were cloned in BSIKS+ or pGEMT. The sizes of the probes ranged between 0.6 – 3 kb. The sequences of the templates can be requested from the author. In several cases more than one probe per gene was tested. Hybridizations and data collection were done as described [50].

Authors' contributions

OR, SB, GE planned the experiments, collected and analyzed the data, and wrote the manuscript. FMC, PB, TL, AK, TS, NB, conducted experiments, analyzed the data, and participated in manuscript writing.

Additional material

Additional File 1

Supplementary Fig. 1: Detailed list and sequences of all the DCX proteins used in this study. The list contains the set of proteins with two domains. In addition, the N-terminal domains including proteins with single domains more similar to the N-terminal domain of DCX, and the C-terminal domains including proteins with single domains more similar to the C-terminal domain of DCX are presented in this set.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-188-S1.doc>]

Additional File 2

Supplementary Fig. 2: Synthetic positions of mouse and human DCX domain proteins.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-188-S2.doc>]

Additional File 3

Supplementary Fig. 3: CLUSTAL W (1.83) multiple sequence alignment of the DCX proteins with tandem domains, N-terminal domains and C-terminal domains.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-188-S3.doc>]

Additional File 4

Supplementary Fig. 4: ML tree of the N-terminal DCX domain proteins from different species. Bootstrap values are indicated.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-188-S4.jpeg>]

Additional File 5

Supplementary Fig. 5: ML tree of the C-terminal DCX domain proteins from different species. Bootstrap values are indicated.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-188-S5.jpeg>]

Additional File 6

Supplementary Fig. 6: Western blot analysis using anti-FLJ46154 antibodies. The specificity of the antibodies was verified using cells transfected with Flag-FLJ46154 (A) compared with the preimmune serum. (B) Anti-FLJ46154 antibodies recognize a protein of approximate size 78 kDa in mouse brain extract.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-188-S6.jpeg>]

Additional File 7

Supplementary Fig. 7: Proposed nomenclature for the doublecortin superfamily.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-188-S7.doc>]

Acknowledgements

We thank Prof. Shmuel Pietrekovski for helpful discussions. The work has been supported in part by the Fritz Thyssen Stiftung Foundation, the Israeli Science Foundation (grant no. 270/04), Foundation Jérôme Lejeune, Minerva foundation with funding from the Federal German Ministry for Education and Research, the German – Israeli collaboration grant Gr-1905, a grant from the Paul Godfrey Research Foundation in Childrens' diseases, the Kekst center, the Forcheimer center, the Jewish communal fund Albert Einstein College of Medicine of Yeshiva University, and the David and Fela Shapell Family Center for Genetic Disorders Research. O.R. is an Incumbent of the Berstein-Mason professorial chair of Neurochemistry. F. C. was supported by a Post-Doctoral Fellowship from the Association pour la Recherche sur le Cancer (Villejuif, France), and by the Sir Charles Clore Post-Doctoral Fellowship, the Weizmann Institute of Science (Rehovot, Israel).

References

- des Portes V, Pinard JM, Billuart P, Vinet MC, Koulakoff A, Carrie A, Gelot A, Dupuis E, Motte J, Berwald-Netter Y, Catala M, Kahn A, Beldjord C, Chelly J: **A novel CNS gene required for neuronal migration and involved in X-linked subcortical laminar heterotopia and lissencephaly syndrome.** *Cell* 1998, **92**:51-61.
- Gleeson JG, Allen KM, Fox JW, Lamperti ED, Berkovic S, Scheffer I, Cooper EC, Dobyns WB, Minnerath SR, Ross ME, Walsh CA: **doublecortin, a brain-specific gene mutated in human X-linked lissencephaly and double cortex syndrome, encodes a putative signaling protein.** *Cell* 1998, **92**:63-72.
- Reiner O, Coquelle FM: **Missense mutations resulting in type I lissencephaly.** *Cell Mol Life Sci* 2005, **62**(4):425-434.
- Dobyns WB, Reiner O, Carrozzo R, Ledbetter DH: **Lissencephaly: a human brain malformation associated with deletion of the LIS1 gene located at chromosome 17p13.** *J Am med Ass* 1993, **270**:2838-2842.
- Barkovich AJ, Guerrini R, Battaglia G, Kalifa G, N'Guyen T, Parmegiani A, Santucci M, Giovanardi-Rossi P, Granata T, D'Incerti L: **Band heterotopia: correlation of outcome with magnetic resonance imaging parameters.** *Ann Neurol* 1994, **36**(4):609-617.
- Horesh D, Sapir T, Francis F, Caspi M, Grayer Wolf S, Elbaum M, Chelly J, Reiner O: **Doublecortin, a stabilizer of microtubules.** *Hum Mol Genet* 1999, **8**:1599-1610.
- Gleeson JG, Lin PT, Flanagan LA, Walsh CA: **Doublecortin is a microtubule-associated protein and is expressed widely by migrating neurons.** *Neuron* 1999, **23**(2):257-271.
- Francis F, Koulakoff A, Boucher D, Chafey P, Schaar B, Vinet MC, Fricourt G, McDonnell N, Reiner O, Kahn A, McConnell SK, Berwald-Netter Y, Denoulet P, Chelly J: **Doublecortin is a developmentally regulated, microtubule-associated protein expressed in migrating and differentiating neurons.** *Neuron* 1999, **23**(2):247-256.
- Yoshiura K, Noda Y, Kinoshita A, Niikawa N: **Colocalization of doublecortin with the microtubules: An ex vivo colocalization study of mutant doublecortin.** *J Neurobiol* 2000, **43**(2):132-139.
- Burgess HA, Reiner O: **Doublecortin-like Kinase, is a Microtubule-Associated Protein Kinase Expressed in Growth Cones.** *MCN* 2000, **16**:529-541.
- Lin PT, Gleeson JG, Corbo JC, Flanagan L, Walsh CA: **DCAMKLI encodes a protein kinase with homology to doublecortin that regulates microtubule polymerization.** *J Neurosci* 2000, **20**(24):9152-9161.
- Sapir T, Horesh D, Caspi M, Atlas R, Burgess HA, Grayer Wolf S, Francis F, Chelly J, Elbaum M, Pietrekovski S, Reiner O: **Doublecortin Mutations Cluster in Evolutionary Conserved Functional Domains.** *Hum Mol Genet* 2000, **5**:703-712.
- Taylor KR, Holzer AK, Bazan JF, Walsh CA, Gleeson JG: **Patient mutations in doublecortin define a repeated tubulin-binding domain.** *J Biol Chem* 2000, **275**(44):34442-34450.
- Kim MH, Cierpicki T, Derewenda U, Krowarsch D, Feng Y, Devedjiev Y, Dauter Z, Walsh CA, Otlewski J, Bushweller JH, Derewenda ZS: **The DCX-domain tandems of doublecortin and doublecortin-like kinase.** *Nat Struct Biol* 2003, **10**(5):324-333.

15. Moores CA, Perderiset M, Francis F, Chelly J, Houdusse A, Milligan RA: **Mechanism of microtubule stabilization by doublecortin.** *Mol Cell* 2004, **14(6)**:833-839.
16. Pierce EA, Quinn T, Meehan T, McGee TL, Berson EL, Dryja TP: **Mutations in a gene encoding a new oxygen-regulated photoreceptor protein cause dominant retinitis pigmentosa.** *Nat Genet* 1999, **22(3)**:248-254.
17. Sullivan LS, Heckenlively JR, Bowne SJ, Zuo J, Hide WA, Gal A, Denton M, Inglehearn CF, Blanton SH, Daiger SP: **Mutations in a novel retina-specific gene cause autosomal dominant retinitis pigmentosa.** *Nat Genet* 1999, **22(3)**:255-259.
18. Gao J, Cheon K, Nusinowitz S, Liu Q, Bei D, Atkins K, Azimi A, Daiger SP, Farber DB, Heckenlively JR, Pierce EA, Sullivan LS, Zuo J: **Progressive photoreceptor degeneration, outer segment dysplasia, and rhodopsin mislocalization in mice with targeted disruption of the retinitis pigmentosa-1 (Rpl) gene.** *Proc Natl Acad Sci USA* 2002, **99(8)**:5698-5703.
19. Meng H, Smith SD, Hager K, Held M, Liu J, Olson RK, Pennington BF, Defries JC, Gelernter J, O'Reilly-Pol T, Somlo S, Skudlarski P, Shaywitz SE, Shaywitz BA, Marchione K, Wang Y, Paramasivam M, Loturco JJ, Page GP, Gruen JR: **DCDC2 is associated with reading disability and modulates neuronal development in the brain.** *Proc Natl Acad Sci USA* 2005.
20. Schumacher J, Anthoni H, Dahdouh F, Konig IR, Hillmer AM, Kluck N, Manthey M, Plume E, Warnke A, Remschmidt H, Hulsmann J, Cichon S, Lindgren CM, Propping P, Zucchelli M, Ziegler A, Peyrard-Janvid M, Schulte-Körne G, Nothen MM, Kere J: **Strong Genetic Evidence of DCDC2 as a Susceptibility Gene for Dyslexia.** *Am J Hum Genet* 2006, **78(1)**:52-62.
21. Corbo JC, Deuel TA, Long JM, LaPorte P, Tsai E, Wynshaw-Boris A, Walsh CA: **Doublecortin is required in mice for lamination of the hippocampus but not the neocortex.** *J Neurosci* 2002, **22(17)**:7548-7557.
22. Bai J, Ramos RL, Ackman JB, Thomas AM, Lee RV, LoTurco JJ: **RNAi reveals doublecortin is required for radial migration in rat neocortex.** *Nat Neurosci* 2003, **6(12)**:1277-1283.
23. Koizumi H, Tanaka T, Gleeson JG: **doublecortin-like kinase Functions with doublecortin to Mediate Fiber Tract Decussation and Neuronal Migration.** *Neuron* 2006, **49(1)**:55-66.
24. Deuel TA, Liu JS, Corbo JC, Yoo SY, Rorke-Adams LB, Walsh CA: **Genetic Interactions between Doublecortin and Doublecortin-like Kinase in Neuronal Migration and Axon Outgrowth.** *Neuron* 2006, **49(1)**:41-53.
25. Edelman AM, Kim WY, Higgins D, Goldstein EG, Oberdoerster M, Sigurdson W: **Doublecortin kinase-2, a novel doublecortin-related protein kinase associated with terminal segments of axons and dendrites.** *J Biol Chem* 2005, **280(9)**:8531-8543.
26. Coquelle FM, Levy T, Bergmann S, Wolf SG, Bar-El D, Sapir T, Brody Y, Orr I, Barkai N, Eichele G, Reiner O: **Common and divergent roles for members of the mouse DCX superfamily.** *Cell Cycle* 2006, **5(9)**:976-983.
27. Weimer JM, Anton ES: **Doubling Up on Microtubule Stabilizers: Synergistic Functions of Doublecortin-like Kinase and Doublecortin in the Developing Cerebral Cortex.** *Neuron* 2006, **49(1)**:3-4.
28. MacMillan LB, Bass MA, Cheng N, Howard EF, Tamura M, Strack S, Wadzinski BE, Colbran RJ: **Brain actin-associated protein phosphatase I holoenzymes containing spinophilin, neurabin, and selected catalytic subunit isoforms.** *J Biol Chem* 1999, **274(50)**:35845-35854.
29. Satoh A, Nakanishi H, Obaishi H, Wada M, Takahashi K, Satoh K, Hirao K, Nishioka H, Hata Y, Mizoguchi A, Takai Y: **Neurabin-III/spinophilin. An actin filament-binding protein with one pdz domain localized at cadherin-based cell-cell adhesion sites.** *J Biol Chem* 1998, **273(6)**:3470-3475.
30. Liu Y, Chirino AJ, Misulovin Z, Leteux C, Feizi T, Nussenzweig MC, Bjorkman PJ: **Crystal structure of the cysteine-rich domain of mannose receptor complexed with a sulfated carbohydrate ligand.** *J Exp Med* 2000, **191(7)**:1105-1116.
31. **UCSC BLAT Site** [<http://genome.ucsc.edu/cgi-bin/hgBlat?command=start>]
32. Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Yamashita RA, Yin JJ, Zhang D, Bryant SH: **CDD: a Conserved Domain Database for protein classification.** *Nucleic Acids Res* 2005, **33(Database):**D192-196.
33. **Unigene** [<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=unigene>]
34. Zeng L, Gu S, Li Y, Zhao E, Xu J, Ye X, Wu Q, Wang L, Xie Y, Mao Y: **Identification of a novel human doublecortin-domain-containing gene (DCDC1) expressed mainly in testis.** *J Hum Genet* 2003, **48(7)**:393-396.
35. Kim MH, Derewenda U, Devedjiev Y, Dauter Z, Derewenda ZS: **Purification and crystallization of the N-terminal domain from the human doublecortin-like kinase.** *Acta Crystallogr D Biol Crystallogr* 2003, **59(Pt 3)**:502-505.
36. Gonczy P, Bellanger JM, Kirkham M, Pozniakowski A, Baumer K, Phillips JB, Hyman AA: **zyg-8, a gene required for spindle positioning in C. elegans, encodes a doublecortin-related kinase that promotes microtubule assembly.** *Dev Cell* 2001, **1(3)**:363-375.
37. Burgess HA, Reiner O: **Cleavage of Doublecortin-like Kinase by Calpain Releases an Active Kinase Fragment from a Microtubule Anchorage Domain.** *J Biol Chem* 2001, **276(39)**:36397-36403.
38. Eichenmuller B, Ahrens DP, Li Q, Suprenant KA: **Saturable binding of the echinoderm microtubule-associated protein (EMAP) on microtubules, but not filamentous actin or vimentin filaments.** *Cell Motil Cytoskeleton* 2001, **50(3)**:161-172.
39. Eichenmuller B, Everley P, Palange J, Lepley D, Suprenant KA: **The human EMAP-like protein-70 (ELP70) is a microtubule destabilizer that localizes to the mitotic apparatus.** *J Biol Chem* 2002, **277(2)**:1301-1309.
40. Shu T, Tseng HC, Sapir T, Stern P, Zhou Y, Sanada K, Fischer A, Coquelle FM, Reiner O, Tsai LH: **Doublecortin-like Kinase Controls Neurogenesis by Regulating Mitotic Spindles and M Phase Progression.** *Neuron* 2006, **49(1)**:25-39.
41. Liu Q, Zhou J, Daiger SP, Farber DB, Heckenlively JR, Smith JE, Sullivan LS, Zuo J, Milam AH, Pierce EA: **Identification and subcellular localization of the RPI protein in human and mouse photoreceptors.** *Invest Ophthalmol Vis Sci* 2002, **43(1)**:22-32.
42. Liu Q, Zuo J, Pierce EA: **The retinitis pigmentosa I protein is a photoreceptor microtubule-associated protein.** *J Neurosci* 2004, **24(29)**:6427-6436.
43. Liu J, Huang Q, Higdon J, Liu W, Xie T, Yamashita T, Cheon K, Cheng C, Zuo J: **Distinct gene expression profiles and reduced JNK signaling in retinitis pigmentosa caused by RPI mutations.** *Hum Mol Genet* 2005, **14(19)**:2945-2958.
44. Capes-Davis A, Tolhurst O, Dunn JM, Jeffrey PL: **Expression of doublecortin (DCX) and doublecortin-like kinase (DCLK) within the developing chick brain.** *Dev Dyn* 2005, **232(2)**:457-467.
45. Kappeler C, Saillour Y, Baudoin JP, Tuy FP, Alvarez C, Houbbron C, Gaspar P, Hamard G, Chelly J, Metin C, Francis F: **Branching and nucleokinesis defects in migrating interneurons derived from doublecortin knockout mice.** *Hum Mol Genet* 2006, **15(9)**:1387-1400.
46. Ramos RL, Bai J, Loturco JJ: **Heterotopia Formation in Rat but Not Mouse Neocortex after RNA Interference Knockdown of DCX.** *Cereb Cortex* 2005.
47. **EST database of Ciona intestinalis (sea squirt)** [<http://ghost.zool.kyoto-u.ac.jp/index1.html>]
48. Page RD: **TreeView: an application to display phylogenetic trees on personal computers.** *Comput Appl Biosci* 1996, **12(4)**:357-358.
49. Henikoff S, Henikoff JG, Alford WJ, Pietrokovski S: **Automated construction and graphical presentation of protein blocks from unaligned sequences.** *Gene* 1995, **163**:17-26.
50. Yaylaoglu MB, Titmus A, Visel A, Alvarez-Bolado G, Thaller C, Eichele G: **Comprehensive expression atlas of fibroblast growth factors and their receptors generated by a novel robotic in situ hybridization platform.** *Dev Dyn* 2005, **234(2)**:371-386.
51. **UCSC site** [<http://genome.ucsc.edu>]