C21orf5, a New Member of Dopey Family Involved in Morphogenesis, Could Participate in Neurological Alterations and Mental Retardation in Down Syndrome

Mohammed Rachidi,* Carmela Lopes, Maher Costantine, and Jean-Maurice Delabar

EA 3508 Université Paris Denis Diderot, Paris, France

(Received 5 November 2004; revised 20 April 2005)

Abstract

Availability of the human genome sequence promises important progress in the understanding of human pathologies, particularly for multifactorial diseases. Among these, Down syndrome (DS) is the most frequent genetic cause of mental retardation. A critical region of chromosome 21, the Down syndrome Chromosomal Region-1 (DCR-1), is responsible for many features of the DS phenotype including mental retardation. We studied DCR-1 C21orf5 as a new candidate gene for DS considering its restricted expression in key brain regions altered in DS patients and involved in learning and memory processes. To elucidate C21orf5 molecular function, we performed a comparative study of protein sequences in several species and showed that C21orf5 represents a new member of the Dopey leucine zipper-like family. The C21orf5 C-termini contains two highly conserved leucine-like zipper domains in invertebrate and vertebrate species. Evolution analysis indicated a common ancestral origin of these protein sequences also suggesting a conserved function of this gene throughout phylogenesis. Mutations of the known C21orf5 homologous genes Aspergillus nidulans DopA, Saccharomyces cerevisiae Dop1 and Caenorhabditis elegans pad1, determine morphological abnormalities. We studied transgenic mice carrying the human C21orf5 gene and we showed that this gene is overexpressed in brain regions by in situ hybridization and by real-time RT–PCR experiments. Interestingly, we also showed that these transgenic mice have an increased density of cortical cells overexpressing C21orf5. Similarly, DS patients have an altered lamination pattern in their cortex. Considering together our and previous findings, we suggest that the human dopey family member, C21orf5, could play a role in brain morphogenesis and, when overexpressed, it could participate in neurological features and mental retardation observed in DS patients.

Key words: Down syndrome; Dopey family; Leucine-like zipper domain; Cortical alteration; Mental retardation

1. Introduction

The Human genome sequencing project delivered the complete DNA sequence of our species.1,2 Numerous data mining tools allowed gene identification, prediction, annotation and precise localization. The availability of these data promises breakthroughs in the understanding of human pathologies. Important progresses could be particularly expected in regard to multifactorial pathologies, determined by interactions between a variable number of genes and characterized by complex phenotypes. Among the multifactorial diseases, Down syndrome (DS) or Trisomy 21 is the most frequent genetic cause of mental retardation. This pathology is the most common autosomal trisomy, with an incidence of 1 in 700 live births, in comparison with other trisomies such as Trisomy 18 (1 : 3500) or Trisomy 13 (1 : 7000). DS phenotypes are variable, affecting many different organs and tissues including brain (mental retardation, increased occurrence of Alzheimer’s disease), muscle, heart, skeleton, intestine and blood. Metabolic disorders are also observed as neurotransmitter alterations and hormonal
imbalances.\textsuperscript{3–7} However, only mental retardation is invariably present in DS patients. In most cases, DS results of an extra copy of chromosome 21 in all cells of the afflicted individuals.\textsuperscript{8} In some rare cases, DS results from a partial trisomy 21 showing variable phenotypes depending on triplicated region. Clinical, cytogenetic and molecular analysis of such patients allowed the identification of Down syndrome Chromosomal Region-1 (DCR-1), at 21q22.2 sub-band, responsible for many features of the DS, including mental retardation.\textsuperscript{9,10}

DS may be considered as a multifactorial disease with an unusual etiology determined by the overexpression of several chromosome 21 genes. The number of the genes involved in the DS pathogenesis remains undetermined and candidate genes should be searched on whole chromosome 21, particularly in the DCR-1. Extensive efforts have been carried out to construct transcriptional map of the DCR-1 region.\textsuperscript{11} The complete sequencing of the human chromosome 21,\textsuperscript{12} its transcriptional activity analysis\textsuperscript{13} and the frequent updatings of the genomic sequence identified >300 genes, of which at least 28 are localized in the DCR-1. It is from these that etiological genes for DS will be identified, since all of them may be potential candidate genes. To date, only few genes in DCR-1 have been studied as candidate genes, DYRK1A,\textsuperscript{14} and SIM2.\textsuperscript{15,16} These two human genes are orthologs of the Drosophila minibrain and single minded genes\textsuperscript{17,18} both involved in Drosophila central nervous system development. Interestingly, overexpression of the mammalian Dyrk1A and Sim2 genes determines some impairments in behavior, memory and learning capacities.\textsuperscript{19,20}

We have previously identified \textit{C21orf5} gene as a new transcriptional unit in DCR-1 region.\textsuperscript{11} More recently, we have investigated the expression pattern of the \textit{C21orf5} mouse ortholog and identified an interesting specific transcriptional profile in the medial temporal-lobe system,\textsuperscript{21} a key system involved in memory storage.\textsuperscript{22} On the basis of these findings, we studied \textit{C21orf5} as a new candidate gene for DS and proposed a potential role in brain morphogenesis, and overexpression of this gene could participate in neurological alterations and mental retardation seen in DS patients.

2. Materials and Methods

2.1. Bioinformatic sequence studies

Sequences with similarity to \textit{C21orf5} were identified by BLAST (Basic Local Alignment Search Tool) family of program (http://www.ncbi.nlm.nih.gov/blast/index.html) searches for similarity to sequences contained in GeneBank.

The sequence data were accessed through NCBI (http://www.ncbi.nlm.nih.gov/); Human Genome Browser Gateway of the Santa Cruz University (http://genome.ucsc.edu/index.html) and Ensembl Project (http://www.ensembl.org/).

Complete genome and Expressed Sequence Tag (EST) data were also accessed through the following specialized sites: \textit{Caenorhabditis elegans} (http://elegans.swmed.edu/); Fugu (http://www.fugu-sg.org/); Mouse (http://www.ensembl.org/Mus_musculus/); Rat (http://www.ensembl.org/Rattus_norvegicus/); \textit{Drosophila} (http://flybase.bio.indiana.edu/).

Multiple sequence alignments were performed with BLAST 2 sequence program, for sequence pairs (http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html) and the alignment CLUSTALW algorithm (http://www.infobiogen.fr/services/analyseq/cgi-bin/clustalw_in.pl). The output files were visualized using the Boxshade program (http://www.ch.embnet.org/software/BOX_form.html). Percent identities were derived from the CLUSTALW output.

SMART (Simple Modular Architecture Research Tool; http://smart.embl-heidelberg.de/) and PFSCAN (http://hits.isb-sib.ch/cgi-bin/PFSCAN) programs were used to study protein domain structures, to identify Prosite profiles and Pfam domains in the protein sequences.

The hydrophobicity profiles and α-helix propensity of the proteins were analyzed using algorithms available at ProtScale (http://www.expasy.org/cgi-bin/protscale.pl).

Phylogenetic relationships were analyzed using the Phylogenetic Tree Prediction program (http://www.genebee.msu.su/services/phtree_reduced.html).

2.1. Normal and transgenic animals and tissues

Normal and transgenic littersmates were used for \textit{in situ} hybridization using a \textit{C21orf5} riboprobe. Independent lines of transgenic mice containing separate YAC DNAs have been created. Together, the transgenic YACs cover ~2 Mb of contiguous human sequence from DCR-1 that represented an \textit{in vivo} library for phenotypic screening.\textsuperscript{23} Among these independent transgenic mouse lines, we studied two lines both carrying one copy of the human 230E8 YAC, containing \textit{C21orf5} gene.\textsuperscript{11} By amplification of several fragments of the gene sequence, we determined that \textit{C21orf5} is entirely contained in the transgenic human YAC integrated into the mouse genome and by RT–PCR we determined that the human gene was expressed in the transgenic mice.\textsuperscript{21}

Adult brains were dissected, freshly embedded in OCT and frozen quickly over liquid nitrogen. Sections (10 μm) were cut in cryostat (Leika) and stored at −20°C until hybridization.

2.3. Non-radioactive in situ hybridization.

Cryosections were air-dried, washed in PBS, and fixed in 4% paraformaldehyde (w/v). Sections were incubated in PBS for 1 h and washed three times with 1× TE, pH 8.0,
for 2 min each. Finally, they were washed twice in 2× SSC for 5 min. Prehybridization was done in 50% formamide, 4× SSC, 5% dextran sulfate, 1× Denhardt solution, and 500 ng/ml salmon sperm DNA for 30 min at 55 °C. Human cDNA (1600–2800 bp) were used as template to generate sense and antisense riboprobes for in situ hybridization, by Sp6 and T7 RNA polymerases (Promega) incorporating digoxigenin-UTP (Boehringer). Hybridization was done overnight at 65 °C with digoxigenin riboprobe (200 ng/ml). After washes in SSC solutions at increasing stringency, hybridized probes were visualized with alkaline phosphatase-coupled anti-digoxigenin Fab fragment (Boehringer), as described.21 The negative controls underwent the same procedure using the corresponding sense riboprobe.

2.4. RT–PCR quantification

Real-time PCR was performed using the LightCycler thermal cycler system (Roche Diagnostics) with the Amplification kit SYBR Green mix. Poly A+ RNA was prepared from cerebrum of adult mice using the Fast track kit (Invitrogen). After a DNase purification step (Qiagen), RNA concentrations were spectrophotometrically assessed. Reverse transcription was performed using the Retroscript kit with random decamers (Ambion). From a total volume of 50 μl per cDNA, 8 μl were used to prepare five dilutions which were used in the real-time PCR reaction. For gene expression assays, the reference oligonucleotides were chosen at crossing point (Cp) levels equivalent to those observed for the target: ppb (MMU9) fwd: GGATTTGGCTACAAA- AACAGCAA; rev: AGCCAGGCCCCGTAGTGCTTC; the target oligonucleotides for C21orf5 mouse and human mRNAs were fwd: CCTACCTCTCGGCTGAGCTTCC; rev: CTCTCCTCAAGAGGCTGGTCC. No primer dimers were observed during amplification. The slopes of the standard curves were between C3.3 and C3.7. Ratios between transgenic and wild-type values were obtained by using the δ (δ Cp) method with reference curves established for each experiment.

3. Results and Discussion

Previously, we have performed a transcriptional map of the 2 Mb DCR-1 critical region for DS, located at 21q22.2 sub-band, using exon trapping and cDNA selection.11 Among the 27 transcriptional units identified, we isolated human cDNA clones corresponding to a new gene, C21orf5 that we localized between CBR324 and KIAA013625 genes. The genomic sequence of C21orf5, spanning ~130 kb, consists in 37 exons. The 8.5 kb C21orf5 mRNA encodes a putative protein of 2298 amino acids (NM_005128).

We have also previously identified the C. elegans homolog, pad-1, involved in morphogenesis control.26 To identify other C21orf5 homologs and analyze sequence conservation during evolution, we searched translated genomic sequences showing homologies with the C21orf5 protein sequence in specialized databases for different species (see Materials and Methods). We identified two other known genes showing significant homologies with the human C21orf5: Aspergillus nidulans DopA27 and Saccharomyces cerevisiae DOP128 genes. In addition to these known genes, we found deduced homolog protein sequences in mouse, rat, fugu and Drosophila genomes. We also identified several ESTs (Expressed Sequence Tags) for these species, suggesting that the putative genomic sequences are transcribed and correspond to C21orf5 homolog genes.

Using global alignment of these homologous protein sequences, we identified two highly conserved regions (Fig. 1): an N-terminal region, corresponding to the first 300 amino acids of the human sequence; and a C-terminal region, corresponding to the 1944–2178 interval of the human sequence. Although a large variable central region, the protein sequences showed significant identity percent between sequence pairs (Table 1). The high phylogenetic distance between the considered species could have determined accumulation of neutral mutations during evolution that decreases the identity percents in sequence alignments.

Venkatarajan and Braun29 performed a cluster analytical study using a 5D descriptors measuring differences by physical–chemical properties of the 20 amino acids involved in protein synthesis and defined a hierarchic classification in six functional amino acid groups. We applied this functional classification to C21orf5 homolog sequences and the identity percents were calculated again on the entire sequences. An increase of identity percents was observed particularly between the human and lower species sequences (Table 1). This finding suggested that some protein sequence differences correspond to conserved mutations. Homology analysis of the more conserved N- and C-terminal regions showed high identity percent between sequence pairs that are greatly increased (Table 1) using the Venkatarajan and Braun’s functional classification.29

Interestingly, we found that the C-terminal region of C21orf5 protein sequence showed very high homology with a leucine zipper-like motif, called Dopey. Firstly identified in the A. nidulans DopA protein,27 the Dopey motif consists in two adjacent leucine zipper-like domains. Leucine zipper domains are normally characterized by repeated leucines, separated by seven amino acids. In these Dopey leucine zipper-like motifs, some leucines were substituted with functionally similar amino acids (valine, isoleucine, methionine or phenylalanine) and less frequently with other hydrophobic amino acids (Fig. 2A and B). In addition to the homolog C21orf5 complete sequences, we found partial sequences corresponding to the Dopey zipper-like domain in ESTs of other
Figure 1. Multiple sequence alignment of homologous C2orf5 proteins. Identity in amino acids was indicated in black. Similar conserved amino acids were indicated in grey. The two leucine-ZIP domains, identified in the C-terminal regions, are boxed. Hs: human; Mm: mouse; Rn: rat; Fr: fugu; Dm: Drosophila; Ce: C. elegans.

Table 1. Comparison of the homologous C2orf5 sequences.

| Comparison | Complete | N-terminal | C-terminal | ZIP 1 | ZIP 2 |
|------------|---------|-----------|-----------|------|------|
| Hs/Mm      | 85.4% (90.8%) | 94.0% (95.8%) | 91.5% (92.3%) | 96.6% (100%) | 87.5% (95.6%) |
| Hs/Rn      | 83.6% (86.9%) | 87.7% (91.1%) | 86.8% (92.8%) | 96.0% (96.6%) | 87.5% (100%) |
| Hs/Fr      | 47.2% (66.3%) | 71.8% (87.3%) | 65.8% (91.7%) | 79.3% (96.6%) | 60.9% (86.9%) |
| Hs/Dm      | 36.4% (48.4%) | 51.9% (71.1%) | 43.4% (66.9%) | 65.5% (89.7%) | 39.1% (73.9%) |
| Hs/Ce      | 29.7% (45.5%) | 43.8% (64.4%) | 41.0% (64.7%) | 55.2% (82.8%) | 37.2% (65.2%) |
| Hs/An      | 17.1% (34.7%) | 32.1% (58.2%) | 19.9% (44.3%) | 31.0% (82.8%) | 35.3% (69.5%) |

Identity percent using protein sequences and the Venkatarajan’s and Braun’s functional classification, in brackets, were indicated for the complete sequence, the N- and C-terminal regions, and for the two leucine-ZIP domains. Hs: human; Mm: mouse; Rn: rat; Fr: fugu; Dm: Drosophila; Ce: C. elegans; An: A. nidulans.
species for which the complete cDNA is unknown (Cow, AW462239; Marmot, BI416456 and Chicken, BU460448).

Comparison of these sequences (Fig. 2A and B) showed that the Dopey motifs are very highly conserved, particularly the first leucine zipper-like domain, from Human to Fungi. Homology analysis of these two zipper domain showed high identity percents between sequence pairs (Table 1) that greatly increased using the

Figure 2. Sequence analyses of the Dopey leucine zipper-like motifs. (A and B) Multiple sequence alignment of the two leucine-ZIP domains of homologous C21orf5 proteins. Identity in amino acids was indicated in black. Similar conserved amino acids were indicated in grey. Arrows indicated the canonical leucine positions. (C) Linear evolution of the two leucine-ZIP domains of homologous C21orf5 proteins showed by Phylogenetic Tree Prediction Program. (D-G) Analysis of the propensity to form α-helix structures (D and E) and the hydrophobic profile (F and G) of the sequences of the two zipper-like motifs in different species, using algorithms available at ProtScale.30,31 Hs: human; Bt: cow; Mm: mouse; Rn: rat; Mar: marmot; Gg: chicken; Fr: fugu; Dm: Drosophila; Ce: C. elegans; An: A. nidulans.
It is known that the leucine zipper motifs form α-helix structures. To test the functionality of the Dopey leucine zipper-like domains, we analyzed the hydrophobic profile of the two zipper-like motifs and their propensity to form α-helix structures in different species. Interestingly, very similar profiles were observed (Fig. 2D–G), confirming the high functional conservation of the Dopey domains. Protein sequence scanning was also performed to search other functional domains, but, although some transmembrane domains and phosphorylation sites were found, none of them appeared significantly identified. We also analyzed the evolution of these sequences and we observed a linear evolution following the phylogenesis (Fig. 2C), suggesting a common origin of the ancestral sequence.

Independent lines of transgenic mice containing human YACs have been created, which together cover the DCR-1, and represent an in vivo library for phenotypic screening. We studied two of these independent transgenic lines carrying the human YAC 230E8. This YAC, ~600 kb in length, is localized between RH103382 and SHGC-51957 STS markers. We demonstrated that C21orf5 sequence is entirely contained in the transgenic 230E8 YAC and that they expressed the transgenic human C21orf5 gene, in addition to the normally expressed mouse ortholog, determining a global overexpression for this gene.

To assess the level of C21orf5 mRNA in transgenic animals carrying the YAC 230E8, we have performed quantitative real-time RT–PCR experiments, using the sybr green technology, on poly A⁺ RNA cerebrum samples from wild-type and transgenic adult littermates using oligonucleotides from two reference genes mouse ppiB (MMU9) and mouse lst1 (MMU8) (data not shown) and oligonucleotides co-amplifying human and mouse cDNA for C21orf5. With both references, the C21orf5 mRNAs were found significantly increased: by 1.83 ± 0.18 (P = 0.0083) compared with ppiB and by 1.97 ± 0.24 (P = 0.01) compared with lst1, indicating that the human gene is expressed at the same level as the mouse gene, demonstrating an overexpression in brain tissue (Fig. 3).

On the basis of the expression of C21orf5 murine ortholog in specific brain regions, we investigated whether it could play a role in cerebral development. In situ hybridization was performed using a human C21orf5 probe on brain sections of normal and transgenic 230E8 mice (Fig. 4). Interestingly, we found that transgenic mice showed increased cortical cells that overexpressed C21orf5 compared with their wild-type littermates (Fig. 4).

Our genomic analysis showed high homologies suggesting that C21orf5 homologs exist probably in all species and they appeared conserved throughout the evolution. Interestingly, application of a functional hierarchic classification of the amino acid increases the identity percent, suggesting that sequence variations between species are conservative mutations, in agreement to a conserved functional role of the Dopey family.

Dopey leucine zipper-like family included three members, all of the Fungi phylum: A. nidulans DopA, S. cerevisiae DOP1 and the Schizosaccharomyces pombe Dop1 (NP_592953). Here, we extended this protein family to C. elegans, Drosophila, Fugu, Mouse, Rat and Human C21orf5 proteins, such as to all other putative C21orf5 homologs. Members of the Dopey family have two large highly conserved domains in N- and C-termini. The N-terminus, the higher conserved region of the Dopey proteins, is leucine rich.

![Figure 4](image-url)
of either the homologous A. nidulans DopA or S. cerevisiae Dop1 N-terminal domains causes the loss of correct growth polarity and budding pattern in yeasts suggesting a functional role of this domain. In the C-terminus, two leucine-like zipper domains were observed. Several indications suggest that they are functional leucine zipper domains: (i) Amino acid changes observed in leucine canonical positions correspond generally to conservative variations. (ii) We showed that hydrophobic profiles and propensity to form α-helix structures were similar in sequences in different species for the two leucine zipper domains; (iii) Temperature-sensitive Dop1 mutation, which alters one of the leucine zippers, showed similar phenotype than null mutant ΔdopA, in A. nidulans.

The Dopey leucine-like zipper domain, such as the other leucine zipper domains, could be involved in multiple protein–protein interactions of several transcription factors, suggesting that the proteins of the Dopey family might act as transcriptional co-activators. Moreover, this domain appeared highly conserved and its evolution, according to phylogenesis, suggests a common origin of the ancestral sequence. In agreement to a transcriptional function of this protein family, it is known that A. nidulans DopA modulates expression of the key transcriptional regulators of the asexual and sexual cycle, brlA, abaA and steA. Their expression patterns are altered in DopA mutants, suggesting that DopA functions upstream in the developmental pathway.

At this time, functional roles are known for three members of the Dopey family: the A. nidulans DopA, the S. cerevisiae DOP1 and the C. elegans pad-1. DopA mutants determine drastically reduced conidiation in A. nidulans caused by aberrant cell morphology and asynchronous cell differentiation. The budding yeast homolog DOP1 is also involved in establishing cellular morphogenesis, and its inactivation was lethal. The C. elegans ortholog, pattering defective 1 (pad-1), when inactivated by dsRNA-mediated interference (RNAi), showed a lethal phenotype characterized by dramatic alteration of gastrulation and morphogenesis during embryonic worm development. Therefore, mutations of all these three C21orf5 homolog genes determine morphological abnormalities both in two very far species such as the Fungi A. nidulans and S. cerevisiae, and in a complex multicellular organism, the nematode worm C. elegans.

We have found that transgenic mice carrying human C21orf5 overexpressed this gene, by in situ hybridization and real-time RT–PCR, and showed an altered cortex with increased C21orf5-expressing cells compared with their wild-type littermates. We observed the same abnormalities in cortical layers in the two independent transgenic mouse lines produced by Smith et al., suggesting that the cortical phenotype is specific for the presence of the 230E8 YAC. In addition, we performed morphological analysis of brain regions in other transgenic mouse lines carrying three human YAC localized in DCR-1 critical region and not containing C21orf5 gene. These transgenic lines did not show alterations in the cortex similar to the phenotype observed in the two 230E8 transgenic lines (unpublished data) and only behavioural alterations were observed in the transgenic lines carrying 152F7 YAC, containing the human minibrain gene, DYRK1A.

Ours and previous findings suggest that C21orf5 is a good candidate gene for the morphological phenotype in cortical lamination observed in the transgenic line over-expressing this gene:

(i) C21orf5 shows a restricted expression in mouse cortex, hippocampus and cerebellum. Interestingly, these three brain regions play key roles in learning and memory, and they correspond to altered brain structures observed in DS patients.

(ii) The analyzed transgenic mice overexpressed the C21orf5 gene in the cortex and cerebrum at similar level than found in DS human lymphoblasts.

(iii) Mutations of some members of the Dopey family, A. nidulans DopA, S. cerevisiae Dop1 and C. elegans pad1, determine morphological abnormalities. Protein sequence conservation in invertebrate and vertebrate Dopey family members also suggested a conserved function in the control of morphogenesis throughout phylogenesis until the human member, C21orf5.

The transgenic mice overexpressing C21orf5 showed also mild impairment in learning and memory that could be caused by the cortical abnormalities. Interestingly, these mouse phenotypes correspond well to DS neurological alterations characterized by mental retardation and neuromorphological alterations, particularly the abnormal lamination pattern found in the cortex of DS patients. On these bases, we proposed that the human C21orf5 could play a role in human morphogenesis of the cortex and that the overexpression of this gene in DS could participate in morphological features seen in trisomic brains and in mental retardation pathogenesis.

Acknowledgements: This work was supported by the Pasteur Institute, Paris 7 University and Centre National de la Recherche Scientifique (CNRS). We thank Dr E.M. Rubin for providing the transgenic mice. Mohammed Rachidi thanks his colleagues of the Department of Molecular Biology-Jacques Monod at the Pasteur Institute (Paris) for their continuous support.

References
1. Lander, E. S., Linton, L. M., Birren, B., et al. 2001, Initial sequencing and analysis of the human genome, Nature, 409, 860–921.
2. Venter, J. C., Adams, M. D., Myers, E. W., et al. 2001, The sequence of the human genome, Science, 291, 1304–1351.
3. Puffenberger, E. G., Kauffman, E. R., Bolk, S., et al. 1994, Identity-by-descent and association mapping of a recessive gene for Hirschsprung disease on human chromosome 13q22, *Hum. Mol. Genet.*, 3, 1217–1225.

4. Wisniewski, K. E. and Bobinski, M. 1991, Hypothalamic abnormalities in Down syndrome, *Prog. Clin. Biol. Res.*, 373, 153–167.

5. Epstein, C. J. 1995, Down syndrome (trisomy 21), In: Scrivner, C. R., Beaudet, A. L., Sly, W. S., and Valle, D. (eds), The Metabolic and Molecular Bases of Inherited Disease, McGraw-Hill, New York, pp. 749–794.

6. Schneider, C., Risser, D., Kirchner, L., et al. 1997, Similar deficits of central histaminergic system in patients with Down syndrome and Alzheimer disease, *Neurosci. Lett.*, 222, 183–186.

7. Risser, D., Lubec, G., Cairns, N., and Herrera-Marschitz, M. 1995, Single-minded nervous system development, *Proc. Natl Acad. Sci. USA*, 92, 9191–9195.

8. Chen, H., Chrast, R., Rossier, C., et al. 1995, Single-minded and Down syndrome? *Nature Genet.*, 10, 9–10.

9. Crews, S. T., Thomas, J. B., and Goodman, C. S. 1988, The Drosophila singleminded gene encodes a nuclear protein with sequence similarity to the per gene product, *Cell*, 52, 143–151.

10. Tejedor, F., Zhu, X. R., Kaltenbach, E., et al. 1995, Minibrain: a new protein kinase family involved in postembryonic neurogenesis in Drosophila, *Neuron*, 14, 287–301.

11. Dahmane, N., Ghezala, G. A., Gosset, P., et al. 1998, Critical role of the D21S55 region on chromosome 21 in the pathogenesis of Down syndrome, *Proc. Natl Acad. Sci. USA*, 95, 5058–5062.

12. Delabar, J. M., Theophile, D., Rahmani, Z., et al. 1993, Molecular mapping of twenty-four features of Down syndrome on chromosome 21, *Eur. J. Hum. Genet.*, 1, 114–124.

13. Dahmane, N., Ghezala, G. A., Gosset, P., et al. 1998, Transcriptional map of the 2.5-Mb CBR–ERG region of chromosome 21 involved in Down syndrome, *Genomics*, 48, 12–23.

14. The Chromosome 21 Mapping and Sequencing Consortium. 2000, The DNA sequence of human chromosome 21, *Nature*, 405, 311–319.

15. Kapranov, P., Cawley, S. E., Drenkow, J., et al. 2002, Large-scale transcriptional activity in chromosomes 21 and 22, *Science*, 296, 916–919.

16. Guimerà, J., Casas, C., Pucharcos, C., et al. 1996, A human homologue of Drosophila minibrain (MNB) is expressed in the neuronal regions affected in Down syndrome and maps to the critical region, *Hum. Mol. Genet.*, 5, 1305–1310.

17. Dahmane, N., Charron, G., Lopes, C., et al. 1995, Down syndrome critical region contains a gene homologous to Drosophila sim expressed during rat and human central nervous system development, *Proc. Natl Acad. Sci. USA*, 92, 9191–9195.

18. Chen, H., Chrast, R., Rossier, C., et al. 1995, Single-minded and Down syndrome? *Nature Genet.*, 10, 9–10.

19. Ema, M., Ikegami, S., Hosoya, T., et al. 1999, Mild impairment of learning and memory in mice overexpressing the mSin2 gene located on chromosome 16: an animal model of Down’s syndrome, *Hum. Mol. Genet.*, 8, 1409–1415.

20. Altafaj, X., Dierssen, M., Baamonde, C., et al. 2001, Neurodevelopmental delay, motor abnormalities and cognitive deficits in transgenic mice overexpressing Dyrk1A (mini-brain), a murine model of Down’s syndrome, *Hum. Mol. Genet.*, 10, 1915–1923.

21. Lopes, C., Chettouh, Z., Delabar, J. M., and Rachidi, M. 2003, The differentially expressed C2orf5 gene in the medial temporal-lobe system could play a role in mental retardation in Down syndrome and transgenic mice, *Biochem. Biophys. Res. Commun.*, 305, 915–924.

22. Milner, B., Squire, L. R., and Kundel, E. R. 1998, Cognitive neuroscience and the study of memory, *Neuron*, 20, 445–468.

23. Smith, D. J., Stevens, M. E., Sudanagunta, S. P., et al. 1997, Functional screening of 2Mb of human chromosome 21q22.2 in transgenic mice implicates minibrain in learning defects associated with Down syndrome, *Nature Genet.*, 16, 28–36.

24. Watanabe, K., Sugawara, C., Ono, A., et al. 1998, Mapping of a novel human carbonyl reductase, CBR3, and ribosomal pseudogenes to human chromosome 21q22.2, *Genomics*, 52, 95–100.

25. Nagase, T., Seki, N., Tanaka, A., Ishikawa, K. I., and Nomura, N. 1995, Prediction of the coding sequences of unidentified human genes. IV. The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1, *DNA Res.*, 2, 167–174.

26. Guipponi, M., Brunschwig, K., Chamoun, Z., et al. 2000, C2orf5, a novel human chromosome 21 gene, has a Caenorhabditis elegans ortholog (pad-1) required for embryonic patterning, *Genomics*, 68, 30–40.

27. Pascon, R. C. and Miller, B. L. 2000, Morphogenesis in *Aspergillus nidulans* requires Dopey (DopA), a member of a novel family of leucine zipper-like proteins conserved from yeast to humans, *Mol. Microbiol.*, 36, 1250–1264.

28. Dujon, B. 1996, The yeast genome project: what did we learn? *Trends Genet.*, 12, 263–270.

29. Venkatarajan, M. and Braun, W. 2001, New quantitative descriptors of amino acids based on multidimensional scaling of a large number of physical-chemical properties, *J. Mol. Biol.*, 7, 445–453.

30. Kyte, J. and Doolittle, R. F. 1982, A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.*, 157, 105–132.

31. Deleage, G. and Roux, B. 1987, An algorithm for protein secondary structure prediction based on class prediction, *Protein Eng.*, 1, 289–294.

32. Yager, L. N., Kurtz, M. B., and Champe, S. P. 1982, *Aspergillus nidulans* requires Dopey (DopA), a member of a novel family of leucine zipper-like proteins conserved from yeast to humans, *Mol. Microbiol.*, 36, 1250–1264.

33. Golden, J. A. and Hyman, B. T. 1994, Development of the nervous system development, *Proc. Natl Acad. Sci. USA*, 91, 287–301.