regulation, and our understanding of human aneuploidies, such as Down's syndrome. □

Received 11 May; accepted 19 September 2002; doi:10.1038/nature01178.

1. Strachan, T., Abitbol, M., Davidson, D. & Beckmann, J. S. A new dimension for the human genome project: towards comprehensive expression maps. *Nature Genet.* **16,** 126–132 (1997).
2. Neidhardt, L. *et al.* Large-scale screen for genes controlling mammalian embryogenesis, using high-throughput gene expression analysis in mouse embryos. *Mech. Dev.* **98,** 77–94 (2000).
3. Bulfone, A. *et al.* The embryonic expression pattern of 40 murine cDNAs homologous to *Drosophila* mutant genes (Dres): a comparative and topographic approach to predict gene function. *Hum. Mol. Genet.* **7,** 1997–2006 (1998).
4. Reymond, A. *et al.* The tripartite motif family identifies cell compartments. *EMBO J.* **20,** 2140–2151 (2001).
5. Gawantka, V. *et al.* Gene expression screening in *Xenopus* identifies molecular pathways, predicts gene function and provides a global view of embryonic patterning. *Mech. Dev.* **77,** 95–141 (1998).
6. Epstein, C. J. *The Metabolic and Molecular Bases of Inherited Disease* (eds Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D.) 749–794 (McGraw Hill, New York, 1995).
7. Hattori, M. *et al.* The DNA sequence of human chromosome 21. *Nature* **405,** 311–319 (2000).
8. Davisson, M. T. *et al.* Evolutionary breakpoints on human chromosome 21. *Genomics* **78,** 99–106 (2001).
9. Pletcher, M. T., Wiltshire, T., Cabin, D. E., Villanueva, M. & Reeves, R. H. Use of comparative physical and sequence mapping to annotate mouse chromosome 16 and human chromosome 21. *Genomics* **74,** 45–54 (2001).
10. Reymond, A. *et al.* Nineteen additional unpredicted transcripts from the human chromosome 21. *Genomics* **79,** 824–832 (2002).
11. Reymond, A. *et al.* From PREDs and open reading frames to cDNA isolation: revisiting the human chromosome 21 transcription map. *Genomics* **78,** 46–54 (2001).
12. Herzig, U. *et al.* Development of high-throughput tools to unravel the complexity of gene expression patterns in the mammalian brain. *Novartis Found. Symp.* **239,** 129–146 (2001).
13. Egeo, A. *et al.* Developmental expression of the SH3BGR gene, mapping to the Down syndrome heart critical region. *Mech. Dev.* **90,** 313–316 (2000).
14. Abbott, G. W. *et al.* MiRP1 forms Ikr potassium channels with HERG and is associated with cardiac arrhythmia. *Cell* **97,** 175–187 (1999).
15. Sertie, A. L. *et al.* Collagen XVIII containing an endogenous inhibitor of angiogenesis and tumour growth, plays a critical role in the maintenance of retinal structure and in neural tube closure (Knobloch syndrome). *Hum. Mol. Genet.* **9,** 2051–2058 (2000).
16. Camacho Vanegas, O. *et al.* Ullrich scleroatonic muscular dystrophy is caused by recessive mutations in collagen type VI. *Proc. Natl. Acad. Sci. USA* **98,** 7516–7521 (2001).
17. Higuchi, I. *et al.* Frameshift mutation in the collagen VI gene causes Ulrich's disease. *Ann. Neurol.* **50,** 261–265 (2001).
18. Jobsis, G. J. *et al.* Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. *Nature Genet.* **14,** 113–115 (1996).
19. Sumarsono, S. H. *et al.* Down's syndrome-like skeletal abnormalities in Ets2 transgenic mice. *Nature* **379,** 534–537 (1996).
20. Kim, S. K. *et al.* A gene expression map for *Caenorhabditis elegans*. *Science* **293,** 2087–2092 (2001).
21. Lercher, M. J., Urrutia, A. O. & Hurst, L. D. Clustering of housekeeping genes provides a unified model of gene order in the human genome. *Nature Genet.* **6,** 6 (2002).
22. Cohen, B. A., Mitra, R. D., Hughes, J. D. & Church, G. M. A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nature Genet.* **26,** 183–186 (2000).
23. Bortoluzzi, S. *et al.* A comprehensive, high-resolution genomic transcript map of human skeletal muscle. *Genome Res.* **8,** 817–825 (1998).
24. Spellman, P. T. & Rubin, G. M. Evidence for large domains of similarly expressed genes in the *Drosophila* genome. *J. Biol.* **1,** 5 (2002).
25. Tang, H. & Lewontin, R. C. Locating regions of differential variability in DNA and protein sequences. *Genetics* **153,** 485–495 (1999).
26. Ballabio, A. The rise and fall of positional cloning? *Nature Genet.* **3,** 277–279 (1993).
27. Barlow, G. M. *et al.* Down syndrome congenital heart disease: a narrowed region and a candidate gene. *Genet. Med.* **3,** 91–101 (2001).

Supplementary Information accompanies the paper on *Nature*'s website (◗ **http://www.nature.com/nature**).

Acknowledgements We thank M. Traditi and G. Lago for the design of the website, and B. Fischer for preparation of the specimens. We are grateful to F. Chapot, S. Deutsch, M. Guipponi, K. Hashimoto, P. Kahlem, J. Michaud, H. S. Scott and M. L. Yaspo for plasmids and reagents, and to J. Ahidan, M. Friedli, C. Rossier and the Telethon Institute of Genetics and Medicine (TIGEM) RNA *in situ* hybridization core for core assistance. This work was supported by grants from the Jérôme Lejeune Foundation to R.L. and A.R.; from the Swiss Fonds National Suisse de la Recherche Scientifique, the European Union/Office Fédéral de l'Education et de la Santé and ChildCare foundation to S.E.A.; from the German Ministry of Research to G.E.; from the EC Fifth Framework Program to A.B. and G.E.; from the Italian Telethon Foundation to TIGEM; from The National Center for Competence in Research-Frontiers in Genetics to S.E.A.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.B. (e-mail: ballabio@tigem.it), S.E.A. (e-mail: stylianos.antonarakis@medecine.unige.ch) or G.E. (e-mail: gregor.eichele@mpihan.mpg.de).

Authors' contributions The three laboratories of A.B., S.E.A. and G.E. contributed equally to this work.

A gene expression map of human chromosome 21 orthologues in the mouse

The HSA21 expression map initiative
***Group 1:Yorick Gitton†, Nadia Dahmane‡, Sonya Baik†**
Ariel Ruiz i Altaba†
***Group 2:Lorenz Neidhardt§, Manuela Scholze§, Bernhard G. Herrmann§**
***Group 3:Pascal Kahlem‖, Alia Benkahla‖, Sabine Schrinner‖,**
Reha Yildirimman‖, Ralf Herwig‖, Hans Lehrach‖ & Marie-Laure Yaspo‖

** These author groups contributed equally to the work.*
† Skirball Institute, Developmental Genetics Program and Department of Cell Biology, New York University School of Medicine, 540 First Avenue, New York, New York 10016, USA
‡ Université de la Méditerranée – CNRS UMR 6156, Institut de Biologie du Développement, Campus de Luminy, Case 907, 13 288 Marseille Cedex 09, France
§ Max Planck Institute for Immunology, Department of Developmental Biology, Stübeweg 51, D-79108 Freiburg, Germany
‖ Max Planck Institute for Molecular Genetics, Department of Vertebrate Genomics, Ihnestrasse 73, D-14195 Berlin, Germany

The DNA sequence of human chromosome 21 (HSA21)[1] has opened the route for a systematic molecular characterization of all of its genes. Trisomy 21 is associated with Down's syndrome, the most common genetic cause of mental retardation in humans. The phenotype includes various organ dysmorphies, stereotypic craniofacial anomalies and brain malformations[2]. Molecular analysis of congenital aneuploidies poses a particular challenge because the aneuploid region contains many protein-coding genes whose function is unknown. One essential step towards understanding their function is to analyse mRNA expression patterns at key stages of organism development. Seminal works in flies, frogs and mice showed that genes whose expression is restricted spatially and/or temporally are often linked with specific ontogenic processes. Here we describe expression profiles of mouse orthologues to HSA21 genes by a combination of large-scale mRNA *in situ* hybridization at critical stages of embryonic and brain development and *in silico* (computed) mining of expressed sequence tags. This chromosome-scale expression annotation associates many of the genes tested with a potential biological role and suggests candidates for the pathogenesis of Down's syndrome.

The current HSA21 gene catalogue[1] (http://chr21.molgen.mpg.de) contains 238 entries for which we have identified 168 cognate mouse orthologues (see Methods in the Supplementary Information). We isolated 187 mouse cDNA clones matching 158 unique genes (referred to as mmu21 genes; see Supplementary Table 1) whose orthology was confirmed on the basis of the known synteny between HSA21 and segments of mouse chromosomes MMU16, MMU10 and MMU17 (http://www.informatics.jax.org/) (see Supplementary Information; this database is also available at http://chr21.molgen.mpg.de/hsa21/).

To identify potential candidates with a role in patterning and organ development, we have explored the expression of the 158 mmu21 genes by systematic whole-mount *in situ* hybridization[3] (WISH) at mid-gestation (embryonic day 9.5; E9.5), a stage covering a wide range of embryonic processes. We also analysed a subset of clones at two other stages (Supplementary Table 1). We found 111 of 158 genes expressed at E9.5; 78 genes showed widespread expression and 33 genes a restricted pattern (49% and 21%, respectively, of the genes examined at E9.5). In addition, 12 widespread genes also defined particular embryonic structures (for example, *Prkcbp2*; Supplementary Table 1). Among the mmu21 genes conserved in *Saccharomyces cerevisiae* (Y), *Caenorhabditis*

elegans (C) and *Drosophila melanogaster* (D) (CDY in Supplementary Table 1), 60% were widely expressed and 17% showed a restricted pattern. In contrast, 43% of the genes conserved only in multicellular organisms (CD) were widespread and 30% were patterned. The functional classification of the 45 pattern genes shows a bias for cell–cell communication and signal transduction molecules. Only six of the mmu21 genes were previously analysed by whole mounts[4–9], including *Ets2*, a gene that causes skeletal defects in transgenic mice[10]. Our analysis at E9.5 identified a number of interesting patterns (Fig. 1). *Kiaa0184* is a new gene with a striking expression profile in the central and peripheral nervous systems, marking subsets of cells in the midbrain and hindbrain, basal plate margin cells of the spinal cord, and cranial as well as dorsal root ganglia. *Samsn1*, containing SH3 and SAM domains, was found associated with blood vessel formation. In addition, we uncovered new potential roles for previously characterized genes. *Tsga2*, a component of junctional membrane complexes reported to be testis-specific[11], is restricted to rhombomeres at E9.5 and might take part in early events of hindbrain patterning. Informative patterns might be associated with organ-specific traits possibly found in Down's syndrome, although their function remains to be demonstrated. *Igsf5* and *Tff3* are candidates for male sterility. *Igsf5* is present in the mesonephros required for seminiferous tubule formation. *Tff3*, a secretory protein promoting cell migration in intestinal epithelium[12,13], was detected in a small group of cells possibly representing primordial germ cells. *Sh3bgr* is strikingly expressed in heart, correlating with the specific cardiac defects in Down's syndrome. Genes (such as *Slc19a1*, *Clic6* and *Lss*) expressed in the branchial apparatus might have a role in the stereotypic facial abnormalities. *Adamts5*, a disintegrin metalloprotease, is specifically expressed in the maxillary process of the branchial apparatus, and transiently (E8.75–9.5) in the ventral midbrain, making it a candidate for facial abnormalities and mental retardation in Down's syndrome. Of the 28 genes expressed in the brain and/or head mesenchyme at E9.5, 20 show a regionalized expression in the maturing brain at postnatal day 2 (P2) (Supplementary Table 1); examples are *Clic6* and *Tsga2*, which are restricted to the choroid plexus (Fig. 2).

We have investigated the expression of mmu21 genes in neonatal brain, the most important organ affected in Down's syndrome. At P2, the maturation of brain structures involves critical events, including a major shift from neurogenesis to gliogenesis and the elaboration of cytoarchitecture and connectivity. Under our experimental conditions and scoring criteria, 60% of the mmu21 genes were found expressed by *in situ* hybridization (ISH) in the P2 mouse brain (Supplementary Table 1). Of the clones expressed in brain, 42% gave widespread signals, half of these also displaying regionally enhanced expression (for example, *Pfkl*; Supplementary Table 1); 58% of the expressed clones defined only patterns in discrete brain territories, as illustrated in Fig. 2. Figure 3 summarizes the distribution of the mmu21 transcriptome in 11 mouse brain regions.

Overall, there was a prominent expression of mmu21 genes in post-mitotic cells (Fig. 2), whereas only 19 genes were found expressed in the mitotically active ventricular/subventricular zones (Supplementary Table 1). The largest fraction of the mmu21 transcripts expressed at P2 was detected in three major laminar structures of the brain: neocortex (41%), hippocampus (25%) and cerebellum (25%). These structures develop through partly common mechanisms[14,15] and share the expression of a

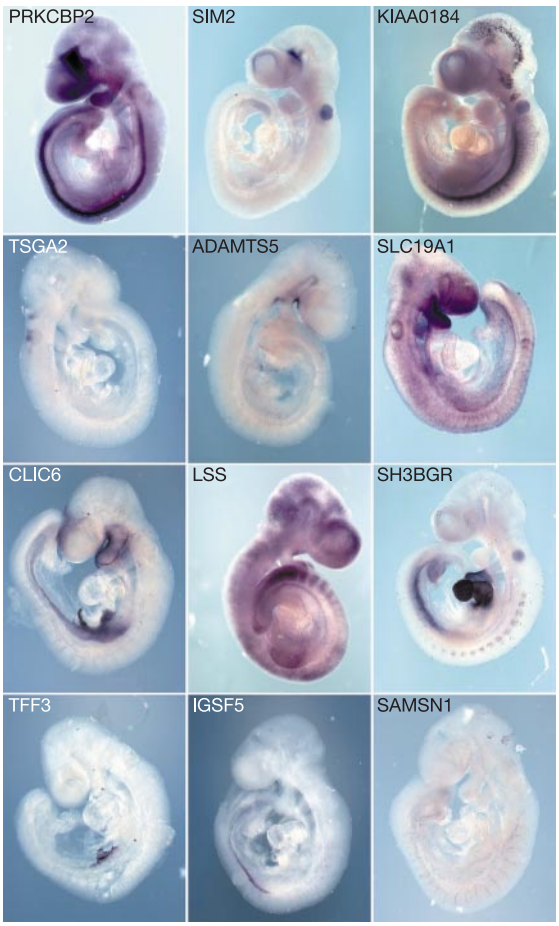

Figure 1 Examples of expression patterns of mmu21 genes in whole-mount E9.5 mouse embryos. Gene symbols of human orthologues are used; for details see Supplementary Table 1.

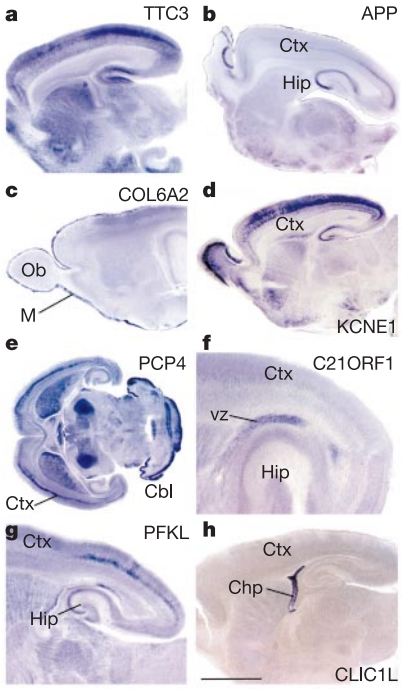

Figure 2 Expression pattern of selected genes on postnatal brain sections by RNA *in situ* hybridization. All sections are from P2 brains and are sagittal, except **e**, which is horizontal. Specific expression is denoted in blue–purple. Anterior is to the left and dorsal side is up. Hip, hippocampus; Ctx, cortex; Tct, tectum; Cb, cerebellum; Ob, olfactory bulb; Chp, choroid plexus; M, meninges; VZ/SVZ, ventricular zone/subventricular zones. Scale bar in **h** represents 200 μm in **a**–**e**, **g**, **h** and 70 μm in **f**. Gene symbols of human orthologues are used.

significant number of genes (Fig. 3). We observed a bias towards neocortical expression, although it remains unclear whether such predominance is also found for the genes encoded by other chromosomes[16,17]. Transcripts prominently expressed in this structure represent a pool of candidates for the cognitive defects of Down's syndrome. *Dscam*, *Synj1* and *Tiam1* have already been described as being involved in synaptic function, axonal guidance, cell migration and neurite outgrowth[18–20]. The hippocampus is necessary for the integrity of memory, and impaired spatial memory in Down's syndrome[21] might be associated with some of the hippocampal genes, such as the semaphorin-like oncogene *Pttg1ip* (ref. 22). A plot of the brain expression domains along the chromosome (Fig. 3a) did not show evidence for the existence of local expression clusters. We found four genes expressed in the meninges (*Wrb*, *Col18a1*, *Col6a1* and *Col6a2*), a tissue enveloping the central nervous system. The three collagen genes are physically close on 21q22.3 and might represent a functionally regulated gene cluster. Because many of the mmu21 genes are expressed in the dorsal brain at P2, our data would be compatible with the hypothesis that cumulative gene dosage effects might underlie brain malfunction in trisomy 21 (refs 23–25).

We obtained a wider overview of mmu21 gene expression by extracting *in silico* expression data from sequenced mouse cDNA libraries representing a large range of tissues and developmental stages. Screening of 632 libraries from non-cancerous tissues totalling 2,253,129 mouse expressed sequence tags (ESTs) identified 5,454 ESTs (on average 34 EST hits per mmu21 gene) displayed in Supplementary Table 1 (see Methods in the Supplementary Information), allowing a detailed analysis of mmu21 gene expression. Although EST mining shows limitations for measuring differential gene expression levels[26–29], we can exploit EST counts to reveal patterns of genes with correlated expression profiles by using a correlation-based clustering method[30]. Figure 4 shows the expression profiles for the mmu21 genes, many of which are weakly or moderately expressed in more than one tissue. Interestingly, 19 genes seemed prominently expressed in single tissues (red boxes in Fig. 4), suggesting that they might have a crucial role in a tissue-specific mechanism (for example, *Tff* (refs 12, 13)). Gene clusters identified here provide information on possible roles of these genes in common biological mechanisms. For instance, five genes (*Runx1*, *Aire*, *Ifngr2*, *Itgb2* and *Ubash3a*) from cluster no. 8 showing thymus-specific expression are linked to immunity or haematopoiesis.

We sought to compare ISH data with expression profiles generated by other approaches. Genes that were scored as not expressed (or not interpretable) on brain sections were tested by polymerase chain reaction with reverse transcription (RT–PCR) on whole P2 brain. As expected, RT–PCR was more sensitive: almost 50% of these ISH-negative genes were found expressed in whole P2 brain by this method (Supplementary Table 1). It is likely that most of those are either expressed ubiquitously at very low levels or are present within a few cells that were absent from the tested sections, which

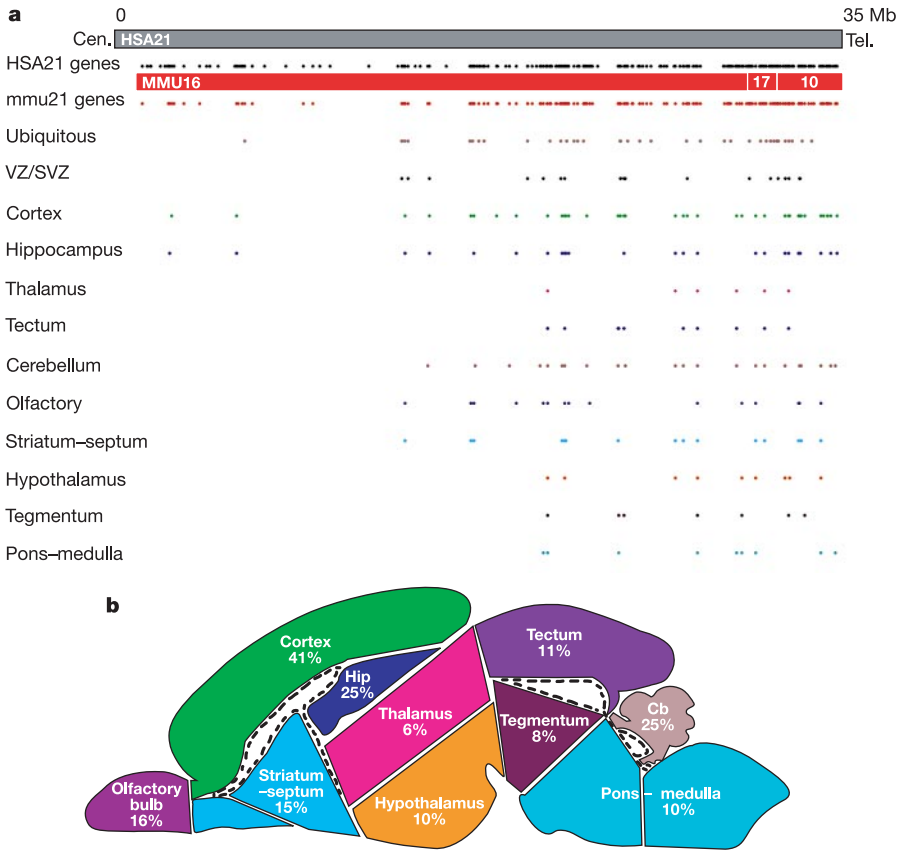

Figure 3 Expression of mmu21 genes in P2 mouse brain. **a**, Chromosomal distribution of the HSA21 genes (top row, black), of their mouse orthologues (second row, red) and of the expression domains of mmu21 genes in different brain regions. The reference scale is that of HSA21 genomic sequence; horizontal bars represent the mouse syntenic chromosomal segments (MMU16, MMU17 and MMU10). Cen, centromere; Tel, telomere; VZ/SVZ, ventricular/subventricular zones. **b**, Diagram summarizing the expression of mmu21 genes in different regions of the P2 mouse brain determined by ISH as reported in **a**. The brain is shown as a sagittal section divided into its major developmental and functional domains. Percentages in each domain refer to the proportion of clones showing expression in each area out of the expressed clones. The ventricular zone is shown by black dashed line. Cb, cerebellum; Hip, hippocampus.

cannot cover the whole brain even though median sagittal sections span a maximum of structures. ISH and RT–PCR are complementary, because ISH aims at identifying patterned genes that are often highly expressed in a given territory, whereas RT–PCR lacks topographical resolution. In addition, the efficiency of riboprobes depends in part on their sequence composition or on the occurrence of splice variants that might show differential expression. It was difficult to correlate our expression map of the P2 brain with the EST information, owing to the heterogeneity in size, tissue region and developmental stages of the original material for the different brain EST libraries analysed. Of the genes found expressed by section ISH, 60% did identify brain ESTs (interrogating 251,426 brain ESTs pooled for all brain regions and ages ranging from neonatal to adult, excluding embryo; see Supplementary Information); 65% of the ISH^{-}/PCR^{+} genes hit brain ESTs, which originated from stages older than P2 with a few exceptions, for example *Ifnrg2* and *Cryzl1* at neonatal stage. Of the genes scoring ISH^{-}/PCR^{-} at P2, 37% hit brain ESTs originating mostly from adult stages, except *Cbs*, which is present in the neonatal cortex. Finally, 16 genes are not, or are very weakly, expressed in neonatal brain because they scored negative by all three methods (*Cldn8*, *Kcne1*, *Runx1*, *Sim2*, *Wdr9*, *C21orf11*, *Mx2*, *Mx1*, *C21orf25*, *Tff3*, *Tff2*, *Ubash3a*, *Wdr4*, *Kiaa0958*, *Ftcd* and *Hsf2bp*).

As expected, the correlation between data from WISH and EST mining was very good because both techniques are able to detect transcripts in most or all embryonic tissues (620,471 ESTs ranging from stage 2 cells to E19.5). This correlation was quantified statistically with Fisher's exact test to refute the hypothesis of statistical independence of the two techniques. For each gene we denoted whether it had positive or negative expression in both methods. Out of 158 genes tested, 123 showed an agreement either with expression (100) or lack of expression (23) in both techniques, whereas in 35 cases we found a disagreement ($P = 2.0496 \times 10^{-7}$). Of the 47 genes not detected at E9.5 by WISH, 23 are unlikely to be expressed at early stages of development because they do not hit any embryonic EST either. One-quarter are found only expressed in libraries originating from later stages, starting from E12.5–E14.5. Ten genes hit ESTs encompassing stage E9.5 (*Tmprss2*, *Slc37a1*, *Cbs*, *Tmem1*, *Dnmt3l*, *C21orf19*, *Ifnar2*, *Cldn8*, *Usp16* and *Usp25*) but were not detected, possibly because of very low expression levels in E9.5 embryos. Overall, most of the transcripts detected at E9.5 by WISH did hit ESTs and also showed sustained expression until stages E14.5–E15.5 or later. Nevertheless, embryonic EST libraries are often derived from the whole fetus and therefore lack organ information. Where information is available, EST mining is correlated with the WISH results (for example, *Sh3bgr* in heart, or *Kiaa0184* and *Pcbp3* in brain).

We observed local clusters of gene expression in several chromo-

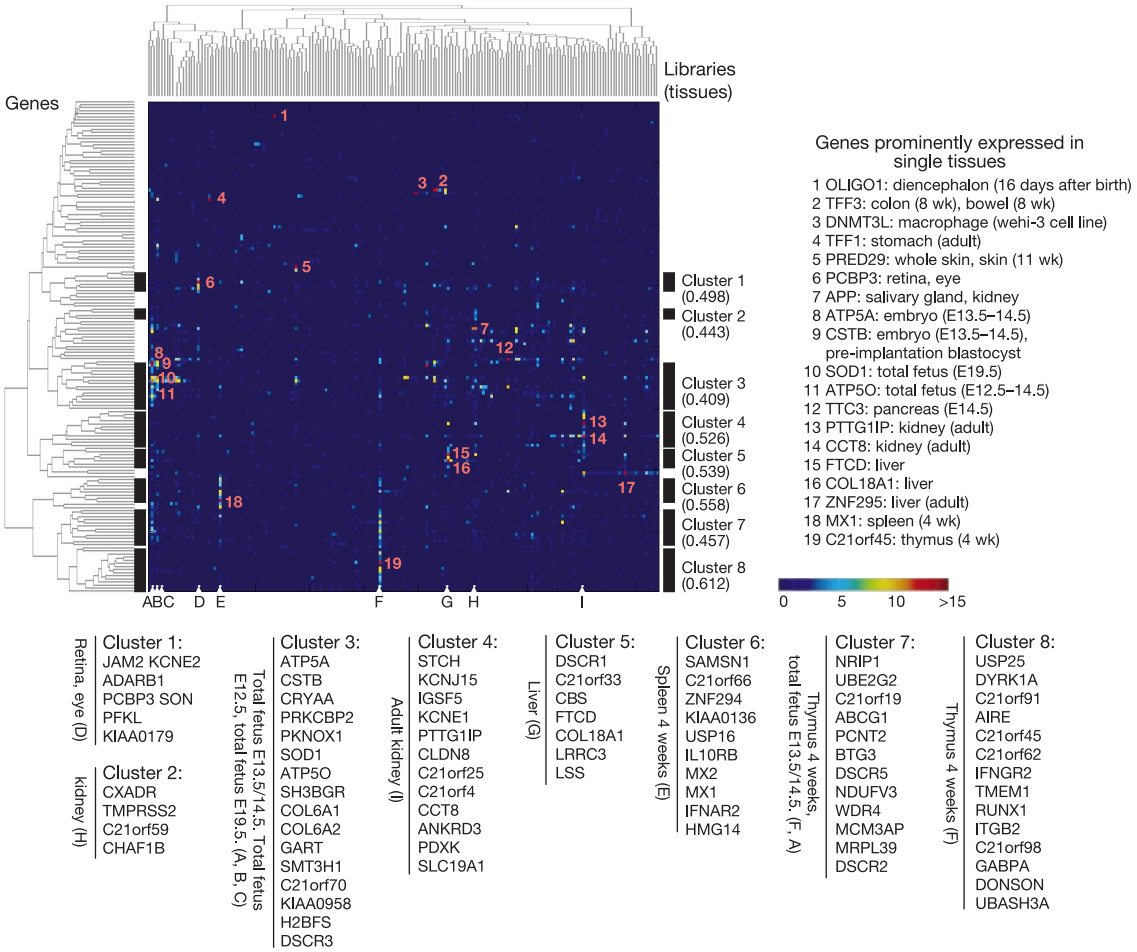

Figure 4 EST analysis. Matrix displaying the expression profiles of 159 mmu21 genes (rows) within 190 cDNA libraries (columns). Dendrograms used to reorder libraries (top) and genes (left) are shown, together with eight significant gene clusters (solid bars, left and right of the matrix together with correlation coefficients). The cluster composition is shown with corresponding libraries (A-I and indicated by white arrows at the bottom of the matrix). Coloured dots represent the number of ESTs found in a given library for each gene (see scale). Numbers in red refer to genes prominently expressed in single tissues, listed at the right. Interactive figure and details are given in the Supplementary Information. Gene symbols of human orthologues are used.

somal regions. Significance was judged by computing upper-tail *P* values (<0.01 in all cases) for observed numbers of expressed genes in groups of neighbouring genes of given sizes in accordance with a hypergeometric distribution. WISH data suggest co-expression from *Pdxk* to *Mcm3ap* in head mesenchyme and extra-embryonic component, and from *Tmprss2* to *Abcg1* in the nose. EST mining data give evidence for a significant group from *H2bfs* to *Pfkl* co-expressed in haematopoietic tissues, kidney, liver and mammary gland. In addition, a large silenced region from *Kcne2* to *Cryaa* was identified in brain, ovary, retina and testis.

Our analysis of the pattern of gene expression for HSA21 is based on complementary techniques, combining the thorough coverage of developmental stages by EST mining with the high anatomical resolution of WISH and brain ISH. We have identified new expression patterns for a large number of genes, including many with highly localized expression, and generated information on the possible functions for a large fraction of the HSA21 genes. This study provides tools and insights for the further analysis of developmental processes and identifies a pool of gene candidates for phenotypes specific to Down's syndrome and other pathologies linked to HSA21. □

Received 11 June; accepted 12 September 2002; doi:10.1038/nature01270.

1. The chromosome 21 mapping and sequencing consortium The DNA sequence of human chromosome 21. *Nature* **405,** 311–319 (2000).
2. Epstein, C. J., *et al. The Metabolic and Molecular Bases of Inherited Disease*, 7th edn (ed. Scriver, C. R.) 749–794 (McGraw-Hill, New York, 1995).
3. Neidhardt, L. *et al.* Large-scale screen for genes controlling mammalian embryogenesis, using high-throughput gene expression analysis in mouse embryos. *Mech. Dev.* **98,** 77–94 (2000).
4. Maroulakou, I. G., Papas, T. S. & Green, J. E. Differential expression of *ets-1* and *ets-2* proto-oncogenes during murine embryogenesis. *Oncogene* **9,** 1551–1565 (1994).
5. Lin, B. *et al.* Prostate-localized and androgen-regulated expression of the membrane-bound serine protease *TMPRSS2. Cancer Res.* **59,** 4180–4184 (1999).
6. Egeo, A. *et al.* Developmental expression of the *SH3BGR* gene, mapping to the Down syndrome heart critical region. *Mech. Dev.* **90,** 313–316 (2000).
7. Epstein, D. J. *et al.* Members of the bHLH-PAS family regulate Shh transcription in forebrain regions of the mouse CNS. *Development* **127,** 4701–4709 (2000).
8. Orti, R. *et al.* Characterization of a novel gene, C21orf6, mapping to a critical region of chromosome 21q22.1 involved in the monosomy 21 phenotype and of its murine ortholog, *orf5. Genomics* **64,** 203–210 (2000).
9. Casas, C. *et al. Dscr1*, a novel endogenous inhibitor of calcineurin signaling, is expressed in the primitive ventricle of the heart and during neurogenesis. *Mech. Dev.* **101,** 289–292 (2001).
10. Sumarsono, S. H. *et al.* Down's syndrome-like skeletal abnormalities in *Ets2* transgenic mice. *Nature* **379,** 534–537 (1996).
11. Taketo, M. M. *et al.* Mapping of eight testis-specific genes to mouse chromosomes. *Genomics* **46,** 138–142 (1997).
12. Mashimo, H., Wu, D. C., Podolsky, D. K. & Fishman, M. C. Impaired defense of intestinal mucosa in mice lacking intestinal trefoil factor. *Science* **274,** 262–265 (1996).
13. Otto, W. R. & Patel, K. Trefoil factor family (TFF)-domain peptides in the mouse: embryonic gastrointestinal expression and wounding response. *Anat. Embryol.* **199,** 499–508 (1999).
14. Altman, J. & Bayer, S. A. *Atlas of Prenatal Rat Brain Development* (CRC Press, Boca Raton, 1995).
15. Dahmane, N. *et al.* The Sonic Hedgehog–Gli pathway regulates dorsal brain growth and tumorigenesis. *Development* **128,** 5201–5212 (2001).
16. Sandberg, R. *et al.* Regional and strain-specific gene expression mapping in the adult mouse brain. *Proc. Natl Acad. Sci. USA* **97,** 11038–11043 (2000).
17. Miki, R. *et al.* Delineating developmental and metabolic pathways *in vivo* by expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays. *Proc. Natl Acad. Sci. USA* **98,** 2199–2204 (2001).
18. Leeuwen, F. N. *et al.* The guanine nucleotide exchange factor Tiam1 affects neuronal morphology; opposing roles for the small GTPases Rac and Rho. *J. Cell Biol.* **139,** 797–807 (1997).
19. Yamakawa, K. *et al.* DSCAM: a novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system. *Hum. Mol. Genet.* **7,** 227–237 (1998).
20. Saito, T. *et al.* Mutation analysis of *SYNJ1*: a possible candidate gene for chromosome 21q22-linked bipolar disorder. *Mol. Psychiat.* **6,** 387–395 (2001).
21. Sylvester, P. E. The hippocampus in Down's syndrome. *J. Ment. Defic. Res.* **27,** 227–236 (1983).
22. Pei, L. Identification of c-*myc* as a down-stream target for pituitary tumour- transforming gene. *J. Biol. Chem.* **276,** 8484–8491 (2001).
23. Delabar, J. M. *et al.* Molecular mapping of twenty-four features of Down syndrome on chromosome 21. *Eur. J. Hum. Genet.* **1,** 114–124 (1993).
24. Korenberg, J. R. *et al.* Down syndrome phenotypes: the consequences of chromosomal imbalance. *Proc. Natl Acad. Sci. USA* **91,** 4997–5001 (1994).
25. Kahlem, P. & Yaspo, M. L. Human chromosome 21 sequence: impact for the molecular genetics of Down syndrome. *Gene Funct. Dis.* **1,** 175–183 (2000).
26. Schmitt, A. O. *et al.* Exhaustive mining of EST libraries for genes differentially expressed in normal and tumour tissues. *Nucleic Acids Res.* **27,** 4251–4260 (1999).
27. Vingron, M. & Hoheisel, J. Computational aspects of expression data. *J. Mol. Med.* **77,** 3–7 (1999).
28. Scheurle, D. *et al.* Cancer gene discovery using digital differential display. *Cancer Res* **60,** 4037–4043 (2000).
29. Krause, A., Haas, S. A., Coward, E. & Vingron, M. SYSTERS, GeneNest, SpliceNest: exploring sequence space from genome to protein. *Nucleic Acids Res.* **30,** 299–300 (2002).
30. Ewing, R. M. *et al.* Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Res.* **9,** 950–959 (1999).

Supplementary Information accompanies the paper on *Nature*'s website (http://www.nature.com/nature).

Acknowledgements We thank M. Sultan, H.-J. Warnatz, G. Teltow, B. Eppens, D. Balzereit, C. Bail, V. Palma, P. Sánchez, M. Beyna, J. Mullor and M. Rallu for their contribution; S. Antonarakis for providing clones; P. Avner for the mouse somatic cell hybrids; and the Resource Center of the German Genome Project (RZPD, Berlin) for providing clones and filters. Y.G., N.D., L.N., P.K. and A.B.K. contributed equally. This work was supported by grants from the Hirscht Trust and the NIH-NINDS and NCI to A.R.A.; from an ATIPE grant to N.D.; from the German Human Genome Project to B.G.H.; from the National Genome Forschung Netz (NGFN), the Ministry of Education and Research (BMBF) and the Max Planck Society to M.-L.Y.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.R.A. (e-mail: ria@saturn.med.nyu.edu), B.G.H. (e-mail: herrmann@immunbio.mpg.de) or M.-L.Y. (e-mail: yaspo@molgen.mpg.de).