[Frontiers in Bioscience 11, 2631-2646, September 1, 2006]

Identification and analysis of novel genes expressed in the mouse embryonic facial primordia

Jennifer S. Bennetts [1], Lindsay F. Fowles [1], Natalie C. Butterfield [1], Jennifer L. Berkman [1], Rohan D. Teasdale [1,2], Fiona Simpson [1], and Carol Wicking [1]

[1] Institute for Molecular Bioscience and [2] ARC Centre in Bioinformatics, The University of Queensland, QLD 4072, Australia

TABLE OF CONTENTS

1. ABSTRACT

Craniofacial anomalies are a common feature of human congenital dysmorphology syndromes, suggesting that genes expressed in the developing face are likely to play a wider role in embryonic development. To facilitate the identification of genes involved in embryogenesis, we previously constructed an enriched cDNA library by subtracting adult mouse liver cDNA from that of embryonic day (E)10.5 mouse pharyngeal arch cDNA. From this library, 273 unique clones were sequenced and known proteins binned into functional categories in order to assess enrichment of the library (1). We have now selected 31 novel and poorly characterised genes from this library and present bioinformatic analysis to predict proteins encoded by these genes, and to detect evolutionary conservation. Of these genes 61% (19/31) showed restricted expression in the developing embryo, and a subset of these was chosen for further *in silico* characterisation as well as experimental determination of subcellular localisation based on transient transfection of predicted full-length coding sequences into mammalian cell lines. Where a human orthologue of these genes was detected, chromosomal localisation was determined relative to known loci for human congenital disease.

2. INTRODUCTION

Embryonic development requires the highly coordinated expression of a host of genes in a precise spatio-temporal manner. Mutations in specific genes contributing to this program of expression can result in a range of dysmorphology syndromes in humans. Craniofacial abnormalities are a common feature of such syndromes, and those disorders that present with craniofacial defects together with abnormalities in other organs and tissues suggest that genes involved in facial development may also play a significant role in the development of other structures. For example, both craniofacial and limb malformations in humans or mice have been associated concurrently with mutations in genes such as *TWIST* (2), *fibroblast growth factor receptor 2* (*FGFR2)* (3) and *GLI3* (4). To date, genes responsible for only a fraction of human congenital abnormalities have been identified, and the discovery of novel genes involved in normal embryonic development will provide further candidates for such disease genes, as well as contributing to our understanding of basic developmental processes.

In order to identify genes with a role in embryonic development, we performed a screen to uncover

genes expressed in mouse embryonic facial primordia, and have previously reported those known genes identified in this screen (1). Briefly, a subtractive hybridisation approach was employed in order to isolate genes expressed specifically in the E10.5 first and second pharyngeal arches. Subtraction was performed against adult mouse liver cDNA to reduce representation of ubiquitously expressed genes. Analysis of the subtracted library revealed successful enrichment for developmentally expressed genes with important functional roles in embryogenesis, and hence validated our approach. We have now selected either completely novel or poorly characterised candidates for further analysis.

In total, 453 clones from the subtracted cDNA library were randomly picked and sequenced. The insert sequence from each of these clones was analysed by BLAST search for homologies against sequences within the publicly available databases, and 273 of these clones were shown to represent unique predicted transcripts, the majority of which corresponded to known genes (1). Of these 273, nine sequences were identified which were identical to novel/predicted gene transcripts, four were identical to transcripts which were annotated based on their predicted protein sequence homology to known proteins, six corresponded to identified/published molecules without characterised functional roles, ten corresponded to partially characterised genes which had not been previously shown to have a role in embryogenesis and two had homology to nucleotide sequences within the database for which a protein-coding region could not be identified. For novel and predicted transcripts, *in silico* analysis was carried out in order to extend the sequence using overlapping sequence data and to identify potential protein-coding regions. To further evaluate these 31 genes and prioritise candidates for future studies, their expression patterns in E10.5 mouse embryos were determined by whole mount *in situ* hybridisation, with 19 candidates showing spatially restricted embryonic expression. For 17 of these candidates, we present more detailed *in silico* investigation to predict protein domains, evolutionary conservation and subcellular localisation. In addition, chromosomal localisation of human orthologues was determined relative to known loci for human congenital diseases. Where appropriate, sequences corresponding to predicted protein coding regions have been amplified and cloned into epitope-tagged expression vectors, and subcellular localisation examined by transient expression in mammalian cell lines. Collectively, these approaches represent an important tool in predicting the functions of novel and uncharacterised developmental molecules, and provide a means of prioritising candidates for more extensive functional characterisation.

3. MATERIALS AND METHODS

3.1. Bioinformatics

Homology searches were performed using BLAST algorithms (5, 6) to identify homologous nucleotide sequences through the National Center for Biotechnology Information (NCBI) public database (http://www.ncbi.nlm.nih.gov/BLAST/). Additionally the NCBI HomoloGene tool was used to identify orthologues of protein sequences in a range of species (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=homologene&cmd=search&term=). Contiguous EST (expressed sequence tag) sequence assemblies were generated using Sequencher 4.2 (Gene Codes, MI, USA) and analysed as previously described to obtain predicted full-length cDNA sequences (1). In addition, the recently released FANTOM III (7) sequence database (http://fantom3.gsc.riken.jp/db) was used to determine and cross-reference sequences. Analysis of predicted protein sequences for the presence of functional protein domains was performed as described previously (1). Additionally, the databases Interpro (http://www.ebi.ac.uk/InterProScan/), Prosite (http://au.expasy.org/prosite/), PsortII (http://psort.nibb.ac.jp/form2.html) and the LOCATE Subcellular Localisation Database (http://locate.imb.uq.edu.au/) were used to determine putative domains and motifs, and only predictions which were above the default threshold value for each program are reported. Human chromosomal localisations were obtained by BLAST search against the publicly available human genome database at (http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html). Human disease loci mapping to each region were obtained using the GenAtlas (http://www.dsi.univ-paris5.fr/genatlas/) and OMIM (Online Mendelian Inheritance in Man, http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) databases. To analyse relative representation of expressed transcripts in the adult mouse, the Source database (http://smd.stanford.edu/cgi-bin/source/sourceSearch) was utilised (8).

3.2. Whole mount *in situ* hybridisation

Whole mount *in situ* hybridisation was performed as described previously (1, 9). Briefly, E10.5 embryos were dissected from pregnant CD1 female mice following timed matings, fixed overnight in 4% paraformaldehyde/PBS at 4°C, dehydrated and then rehydrated through a methanol series. Embryos were placed in prehybridisation buffer and incubated 2 hours to overnight at 65°C. Digoxygenin (DIG)-labelled riboprobes were transcribed from linearised clones using T7 or SP6 polymerase and added to prehybridised embryos at a concentration of 0.2–1.0 µg/ml. Both antisense and sense probes were examined. After overnight hybridisation, embryos were washed, blocked and incubated with anti-DIG antibody conjugated to alkaline phosphatase (Roche Diagnostics, Mannheim, Germany). After a further round of washing, embryos were incubated with the colour reagents nitro blue tetrazolium (Roche Diagnostics) and 5-bromo-4-chloro-3-indolylphosphate (Roche Diagnostics) until satisfactory colour development was achieved. Background was reduced by washing in PBS containing 1% Triton X-100 at 4°C for one to three days and the colour then fixed with 4% paraformaldehyde in PBS overnight at 4°C.

3.3. Cloning of epitope-tagged coding sequences for candidate genes

Where available, RIKEN clones containing full-length predicted coding sequences (supplied from the RIKEN mouse cDNA clone set FANTOM I and II, established by the Genome Exploration Group, RIKEN GSC 'http://genome.riken.go.jp/' and replicated and provided by K. K. Dnaform (10-13)) were obtained from the SRC microarray facility at the Institute for Molecular Bioscience. The use of these clones was sublicensed from The Institute of Physical and Chemical Research (RIKEN). If suitable clones were not available, mRNA was extracted from whole E10.5 mouse embryos using the RNeasy Kit (Qiagen, CA, USA), and cDNA was generated using MMLV-Reverse Transcriptase (Invitrogen Life Technologies, Carlsbad, CA, USA) according to the manufacturer's instructions. RIKEN clone or cDNA templates were used to amplify predicted full-length protein-coding regions of each candidate gene by polymerase chain reaction (PCR) using flanking primers (Geneworks) with engineered restriction sites. The PCR products were each cloned into pGEMTeasy (Promega, Madison, WI, USA) using standard techniques (14), sequenced (ABI PRISM BigDye^{TM} Terminator Version 3.1) and analysed (Australian Genome Research Facility). These constructs were digested using appropriate restriction enzymes (New England Biolabs Inc., Beverley, MA, USA) to release the full-length coding sequences which were then ligated into a pcDNA3 (Invitrogen) vector 5' to a HA (hemagglutinin) or myc epitope tag allowing over-expression of carboxy-terminally tagged predicted full-length proteins in cultured mammalian cells.

3.4. Cell culture and immunofluorescence

Epitope-tagged constructs were transiently transfected into BHK-21 (Syrian golden hamster kidney) and HeLa (human cervical adenocarcinoma) cells using Lipofectamine 2000 (Invitrogen). Prior to transfection, cells were cultured on glass coverslips in multi-well dishes containing DMEM (Invitrogen) supplemented with 5-10% fetal calf serum. Where necessary for visualisation, cells were incubated with 0.05% saponin in PBS for 60 seconds to wash out the cytosol. Cells were fixed 24-48 hours after transfection in 3% paraformaldehyde in PBS for 15-30 minutes. Fixed cells were permeabilised with 0.1% Triton X-100 in PBS for 5-10 minutes, washed and blocked with 2% bovine serum albumin in PBS for a minimum of 30 minutes, and then incubated with appropriate primary antibodies diluted according to the manufacturer's instructions in blocking solution for 60 minutes. Primary antibodies used were: monoclonal anti-HA 16B12 (Santa Cruz Biotechnology Inc., CA, USA), polyclonal anti-HA (Cell Signaling Technology, Beverley MA, USA), monoclonal anti-myc 9B11 (Cell Signaling), polyclonal anti-myc (Cell Signaling). For colocalisation experiments cells were incubated with MitoTracker Red (Molecular Probes, Eugene, OR, USA), monoclonal anti-human vinculin antibody V9131 (Sigma, St Louis, MO, USA), or co-transfected with the construct I-CAT-GFP (15). Cells were washed with blocking solution for 30 minutes, and incubated with anti-mouse or anti-rabbit Alexa fluor secondary antibodies (Molecular Probes) diluted in blocking solution for 60 minutes. Cells were washed for 30 minutes in blocking solution and incubated with DAPI nuclear stain, then washed with PBS and mounted with Mowiol 4-88 (Calbiochem, La Jolla, CA, USA). All incubations and washes were performed at room temperature. Cells were visualised using an Olympus AX-70 microscope (Tokyo, Japan) with a 100X UplanApo oil immersion objective, and images were captured with a DAGE-MTI CCD300-RC ccd camera (Michigan, IN, USA) using NIH Image 1.62. Merged images were produced using Adobe Photoshop 7.0. For colocalisation figures, cells were visualised using an Axiovert 200 M SP LSM 510 META confocal laser scanning microscope (Zeiss, Germany) under oil with a 63X or 100X objective, at excitation and emission wavelengths of 488 and 543, 505-530 and ≥560 nm for green and red fluorescence, respectively. Data was processed using LSM 510 META (Zeiss) software and images were assembled using Adobe Photoshop 7.0.

4. RESULTS

4.1. Initial candidates identified by sequence analysis from the developmentally enriched library

From a total of 273 uniquely represented sequences from our subtracted cDNA library (1), 31 non-redundant sequences were selected for further study. These were identical to regions of nucleotide sequences corresponding to novel or poorly characterised molecules based on homology comparisons by BLAST search against the publicly available databases. Overlapping contiguous EST sequence assemblies were generated to obtain and confirm predicted full-length transcripts. Open reading frames were identified, and these molecules were categorised into one of five groups (Table 1). Nine represented hypothetical/predicted novel protein coding genes (*nv1-nv9*); three were novel genes with annotations based on protein sequence homology to either putative orthologues in other species (*nv10-nv11*) or to characterised proteins with similar domain structure (*Trp53inp2*); six represented molecules which had previously been published either as part of a large-scale screen or based on expression data, with no characterised functional role (*Etea, Kctd3, Cdca4, Cxx1, Hn1* and *Fto*); ten represented molecules which were poorly characterised and had not been previously described as developmentally associated (*Usp52, p15^{Paf}, Elp3, Cdca7, Limd1, C-mir, Thoc3, Maged2, Rasl11b* and *Psf2*); and two represented transcripts for which a confidently predicted open reading frame could not be identified (*nv12-nv13*), either because of a lack of available sequence data in the database, or because they may potentially represent either pseudogenes or regulatory antisense RNA transcripts. At the time that this work was carried out, for one sequence an open reading frame could not be identified, and this gene was not analysed further. However recently released FANTOM III sequence data (7) has allowed us to determine that this sequence actually corresponds to the 3' UTR of the predicted protein-coding gene *Zcchc3* (*NM_175126*).

4.2. Expression analysis by whole mount *in situ* hybridisation

To assess the likelihood of each of these genes playing a role in embryogenesis, the pattern of expression

Table 1. Novel/ poorly characterised molecules selected from subtracted cDNA library for further analysis

Name	Gene accession	Entrez gene/ MGI	Protein name/ description	Human protein	General comments	Status	*In situ*
Novel predicted protein-coding genes							
nv1	NM_134022 6330403K07Rik	103712/1918001	NP_598783 UGS148 protein	-	Novel	N	✓
nv2	AK081685 2600001B17Rik	268490/1919592	NP_766535 hypothetical protein LOC268490	NP_689557 hypothetical protein LOC124801	Novel	N	✓
nv3	AK147478 M5C1050D22	ND/ 1919847	Autism susceptibility candidate 2 (Auts2)	BC064693 AUTS2	Translocation breakpoint in a pair of autistic twins (39)	N	
nv4	NM_026455 2510049I19Rik	67922/1915172	NP_080731 hypothetical protein LOC67922	NP_054796 hypothetical protein LOC26017	Novel	N	✓
nv5	AK013921 3100002L24Rik	66376/1913626	NP_001001737 hypothetical protein LOC414758	-	Novel	N	✓
nv6	NM_025607 2500002L14Rik	66510/1913760	NP_079883 hypothetical RING finger containing protein	NP_057578 hypothetical protein LOC51255	Novel	N	✓
nv7	AK010349 2410003A14Rik	67872/1915122	Novel	NP_060085 hypothetical protein LOC54780	Novel	N	
nv8	AK132460 4732427I03	103012/3533348	BAE21179 hypothetical transmembrane protein	ND	Novel	N	
nv9	AK141733 D030047E04	99003/ 2138986	Partial coding sequence	ND	Novel	N	✓
Molecules annotated based on protein sequence homology							
nv10	NM_029571 1110001A12Rik	100087/1923547	NP_083847 hypothetical protein (yeast Ktl2 homolog)	NP_612426 KTI12 homolog	Chromatin modifier in *S. cerevisiae* (KTI12) (40)	U	
nv11	NM_133934 5830446M03Rik	57905/1923310	NP_598695 hypothetical protein	NP_065752 KIAA1160 predicted protein	mRNA splicing factor Isy1p in *S. cerevisiae* with potential role in cell cycle progression (41, 42)	U	✓
Trp53inp2	NM_178111 1110029F20Rik	68728/1915978	NP_835212 tumor protein p53 inducible nuclear protein 2	NP_067025 TRP53INP2	Sequence homology to stress induced proteins Trp53inp1a/b	U	✓
Zcchc3	NM_175126 2810406K24	67917/1915167	NP_780335 Zinc finger CCHC domain containing protein 3 (Zcchc3), 400aa	NP_149080 ZCCHC3	Novel	N	✓
Molecules with no characterised functional role							
Etea	NM_178397 2210404D11Rik	76577/1923827	NP_848484 protein expressed in T-cells and eosinophils in atopic dermatitis	NP_055428 ETEA	Possible role in cell resistance to apoptosis (43)	N	
Kctd3	NM_172650 4930438A20Rik	226823/2444629	NP_766238 potassium channel tetramerisation domain containing 3	NP_057205 KCTD3/NY-REN-45	Antigen recognised by autologous antibody in patients with renal cell carcinoma (44)	N	
Cdca4	NM_028023 2410018C03Rik	71963/1919213	NP_082299 cell division cycle associated 4	NP_060425 CDCA4	Expressed in hematopoietic progenitors (45)	N	✓
Cxx1	NM_024170 1110012O05Rik	66158/ 1913408	NP_077132 CAAX box 1	NP_003919 CXX1	Frequently methylated in colon carcinogenesis (46)	N	
Hn1	NM_008258	15374/1096361	NP_032284 hematological and neurological expressed sequence 1	NP_057269 HN1	Expressed ubiquitously in adult tissues (47, 48). Upregulated expression in epithelial ovarian cancers (49)	N	✓
Fto	NM_011936 C430017N20	26383/1347093	NP_036066 Fatso	BAB21843 FTO/ KIAA1752	Deleted in fused toes mouse mutant. Expressed in embryonic and adult mouse tissues (50). Upregulated in retina following optic nerve injury (51).	N	✓
Poorly characterised molecules with no known role in development							
Usp52	NM_133992 1200014024Rik	103135/1918984	NP_598753 ubiquitin specific protease 52	NP_055686 USP52	Involved in deadenylation and degradation of mRNA transcripts (52, 53)	U	
$p15^{PAF}$/ OEATC-1	NM_026515 2810417H13Rik	68026/1915276	NP_080791 hypothetical protein LOC68026	Q15004 $p15^{PAF}$(PCNA-associated factor)/ KIAA0101	Altered expression in tumor tissues, possible role in DNA repair, apoptosis and cell cycle regulation (16)	U	✓
Elp3	NM_028811 2610507P14Rik	74195/3529974	NP_083087 elongation protein 3 homolog	NP_060561 ELP3	Histone acetyltransferase in yeast *S. cerevisiae* and *S. pombe* (Elp3) (40, 54, 55)	U	✓
Cdca7	NM_025866 2310021G01Rik	66953/1914203	NP_080142 cell division cycle associated 7	NP_665809 CDCA7/ c-Myc target JPO1	May participate in transformation (56-58)	U	
Limd1	NM_013860	29806/1352502	NP_038888 LIM domains containing 1	NP_055055 LIMD1	Tumor suppressor, binds retinoblastoma protein and suppresses expression of E2F-dependent genes (29, 59)	U	✓
C-mir	NM_027920 1300017E09Rik	71779/1919029	NP_082196 membrane-associated ring finger (C3HC4) 8 (March-8)	NP_659458 cellular modulator of immune recognition (C-mir)	E3 ubiquitin ligase for immune recognition molecules (60)	U	
Thoc3	NM_028597 2410044K02Rik	73666/1920916	NP_082873 THO complex 3	NP_115737 THOC3	Associated with THO complex, involved in mitotic recombination, transcription elongation and mRNA nuclear export (61-63)	U	✓
Maged2	NM_030700	80884/1933391	NP_109625 melanoma antigen, family D, 2	NP_055414 MAGED2	Expressed ubiquitously in adult tissues (64). Expressed in neural tissues at late embryonic stages (65)	U	✓
Rasl11b	BC083068 1190017B18Rik	68939/1916189	AAH83068 RAS-like, family 11, member B	NP_076429 RASL11B	Negatively regulates TGF-beta transcriptional activity. Expression also induced by BMPs (66)	U	
Psf2	AK014776 5730548J20	272551/1921019	Q9D600 DNA replication complex GINS protein, Partner of SLD five, 2	NP_057179 PSF2	Involved in *S. pombe* DNA replication (67). Expressed throughout development in *X. laevis* (68). Upregulated in intrahepatic cholangiocarcinoma (69)	U	✓
Clone inserts corresponding to transcripts/ESTs without identified ORFs, potential pseudogene/ regulatory antisense RNA molecules							
nv12	AK006372 1700026B20Rik	69434/ 2144869	ND	ND	Novel	N	
nv13	NM_146002 (Rhbdl7) Antisense sequence	ND	ND	ND	Novel	N	✓

A '✓' represents molecules for which spatially-restricted expression was observed by whole mount *in situ* hybridisation in the developing mouse embryo at E10.5. Abbreviations: MGI (Mouse Genome Informatics); N, novel; ND, not detected or not able to be obtained through public databases; nv, novel molecule; ORF, open reading frame; U, uncharacterised role in embryonic development.

was determined by whole mount *in situ* hybridisation in E10.5 mouse embryos. DIG-labelled RNA probes were transcribed from the subtracted cDNA library clone templates, and sense probes were also hybridised as negative controls. Of these 31 candidates, 19 exhibited spatially restricted expression in the developing embryo, and 17 of these are shown in Figure 1. Of the remaining two, further characterisation of one of these genes ($p15^{Paf}$) has recently been reported (16), and the other (*Cdca4*) is under further investigation and will be reported at a later date. In addition to expression in the facial primordia, each of these 19 genes was also shown to be expressed in the developing limbs, providing support for the concept that conserved molecular processes govern both limb and craniofacial development (17). These genes were also expressed in a number of other regions most commonly including the heart, otic vesicle, developing neural tissues, developing tail, somites, and mesonephros (Figure 1). Expression of each of these genes in adult mouse tissues was analysed *in silico* using the SOURCE database

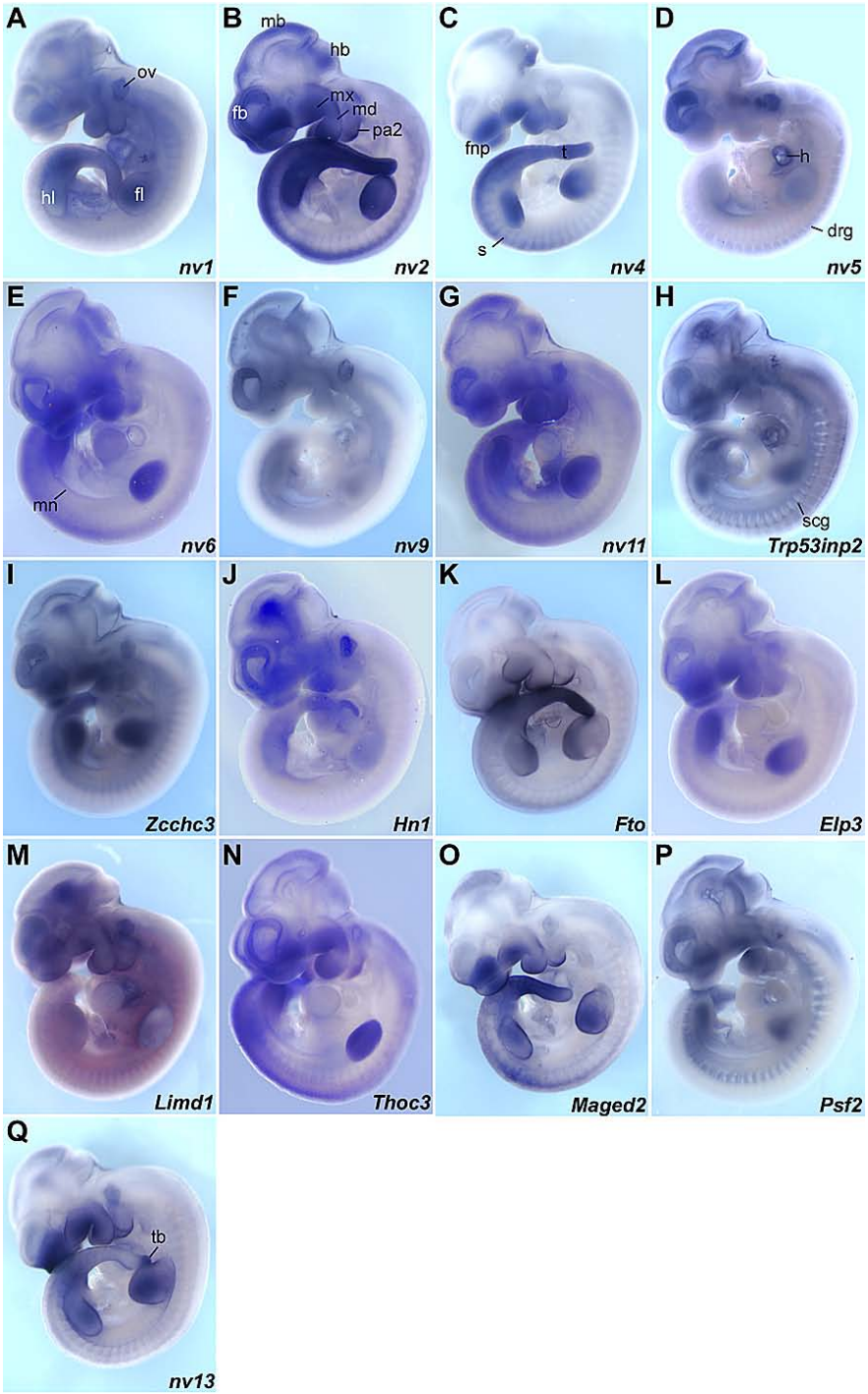

Figure 1. A selection of novel/poorly characterised genes from the subtracted cDNA library which showed restricted expression in E10.5 mouse embryos. Expression patterns for the following genes are shown: *nv1* (A), *nv2* (B), *nv4* (C), *nv5* (D), *nv6* (E), *nv9* (F), *nv11* (G), *Trp53inp2* (H), *Zcchc3* (I), *Hn1* (J), *Fto* (K), *Elp3* (L), *Limd1* (M), *Thoc3* (N), *Maged2* (O), *Psf2* (P) and *nv13* (Q). Common sites of expression to all genes examined include the maxillary and mandibular processes, the second pharyngeal arch, the frontonasal process and the developing fore- and hindlimb buds (A-Q). Other sites of expression include the otic vesicle (developing ear) (A, B, D-G, I, J, L-N, Q), the developing fore-, mid- and hindbrain (A-J, L-P), the somites (B, C, K, L, P, Q), the developing heart (A, D, E, H, J, M), the dorsal root ganglia (D, H) and the spinal chain ganglia (H), the developing tail (A-C, K, O, I), the tailbud (Q) and the mesonephros (G, I, M, P). Abbreviations: drg, dorsal root ganglion; fb, forebrain; fl, forelimb; fnp, frontonasal process; h, heart; hb, hindbrain; hl, hindlimb; mb, midbrain; md, mandibular process; mn, mesonephros; mx, maxillary process; ov, otic vesicle; pa2, 2nd pharyngeal arch; s, somite; scg, spinal chain ganglion; t, developing tail; tb, tailbud.

(http://smd.stanford.edu/cgi-bin/source/sourceSearch) and the majority were found to have widespread expression in the adult (data not shown). This is consistent with a number of known developmental genes, such as those encoding certain members of the transforming growth factor and Wnt signalling molecule families (18, 19), which play very specific and restricted roles in the embryo, but are expressed more widely in the adult and may carry out different functions.

4.3. In depth *in silico* analysis

For each of the expression patterns shown in Figure 1, the corresponding gene was analysed more extensively using *in silico* bioinformatics techniques, except for *nv13*, for which a confidently predicted open reading frame was not identified, and *nv9* (*AK141733*), for which there was insufficient sequence data in the public databases to extend the 5' end of the transcript, and obtain a full-length sequence. Protein sequences were analysed using publicly available domain and motif prediction tools to identify putative functional regions and subcellular distribution (Table 2).

Molecules with important roles in embryonic development are likely to be conserved across species during evolution (20). For each molecule, BLAST and HomoloGene searches were used to identify potential orthologous amino acid sequences in a range of commonly annotated species within the NCBI sequence database (Table 2). Additionally, for each gene for which a human orthologue was identified the human chromosomal localisation was determined by BLAST search against the human genome database at NCBI. To investigate potential association of genes with human hereditary disorders and developmental phenotypes, diseases which show linkage to each chromosomal region were established using the GenAtlas and OMIM databases (Table 2).

4.4. Cloning into expression vectors and transfection into mammalian cells

In order to assess the subcellular localisation of selected proteins, the mouse predicted open reading frame for each molecule was cloned into a C-terminally HA or myc epitope-tagged expression vector. Constructs were transiently transfected into the mammalian cell lines HeLa and BHK-21 and analysed by immunofluorescence with an anti-HA or anti-myc antibody. Immunofluorescence analysis of fixed transfected cells revealed the subcellular localisation of each transfected protein, and showed a high correlation of localisation in both cell lines analysed (Figure 2, Figure 3). Nv4 (Figure 2E,F), nv5 (Figure 2G,H) and Fatso (Figure 3C,D) showed a predominantly nuclear localisation, and nv2 (Figure 2C,D), nv6 (Figure 2I,J), Hn1 (Figure 3A,B) and Psf2 (Figure 3M,N) showed a nuclear and diffuse cytoplasmic distribution. Nv1 (Figure 2A,B) was localised to what appeared to be endoplasmic reticulum, while nv11 (Figure 2K,L) and Elp3 (Figure 3E,F) were localised to the nucleus as well as what appeared to be mitochondria. Trp53inp2 (Figure 2M,N) was distributed within the nucleus and in puncta surrounding the nucleus, and Limd1 (Figure 3G,H) localised throughout the cytoplasm and also at focal adhesions at the cell surface. Overexpression of Thoc3 (Figure 3I,J) caused aggregation in structures presumed to be aggresomal as well as localisation to vesicular-like structures within the cell, and Maged2 (Figure 3K,L) localised to the nucleus (specifically within nucleoli), endosomal-like structures and at the plasma membrane. To confirm preliminary data, the localisation of a subset of these molecules was explored further by colocalisation with known subcellular markers using confocal microscopy (Figure 4). Nv1 was co-transfected with I-CAT-GFP and colocalisation confirmed an endoplasmic reticulum distribution (Figure 4A-C). Nv11 and Elp3 staining colocalised with MitoTracker Red in mitochondria (Figure 4D-I), and Limd1 colocalised with vinculin at focal adhesions (Figure 4J-L). Subcellular localisation data is summarised in Table 2 alongside PsortII predicted localisations.

5. DISCUSSION

5.1. Whole mount *in situ* expression analysis and bioinformatics to identify novel/poorly characterised genes with developmental roles

We previously reported the generation and analysis of a subtracted cDNA library which was shown to be significantly enriched for molecules with functions in embryogenesis (1). Of a total of 273 non-redundant sequences analysed from this library, 31 of these (11.4%) were identified as novel/poorly characterised candidates and prioritised for further analysis to investigate potential developmental roles.

Of these 31 genes, 19 were shown to be expressed in a restricted pattern in the developing mouse embryo and of these, 17 were selected for further characterisation in the current study. For these genes, a range of expression patterns was observed. Notably, each exhibited staining in both the craniofacial region and the limbs, which lends further support for the concept that genes involved in craniofacial development are also likely to be implicated in limb development. Furthermore, within these structures a variety of expression patterns were observed. For example, while some genes appeared to be expressed generally throughout the mesenchymal core of the pharyngeal arches, others showed more restricted expression to distal mesenchyme and ectoderm (eg. *Fto*, Figure 1K). This may reflect the fact that patterning and outgrowth of the facial primordia require epithelial-mesenchymal signalling interactions (reviewed in (21)), and those more restricted genes may potentially be involved in these processes. Also, some genes appeared to show restricted expression within the first pharyngeal arch, and were expressed at lower levels in the second arch (eg. *Hn1,* Figure 1J). These differences in expression may signify molecular differences between the arches, and may also suggest potential pathways in which these uncharacterised molecules could function. In the limb a range of restricted expression domains were also observed, consistent with known reciprocal molecular interactions governing patterning and development in all three axes of the limb (22).

Table 2. Summary of further analysis of predicted proteins encoded by developmentally restricted novel/ poorly characterised genes

Name, length (aa)	Functional Domains/ Motifs (amino acid position), function	Psort prediction (amino acid position)	Subcellular distribution of construct	Orthologue Ac. ID, *species* (% aa identity)	Human locus	Disease/abnormality and genomic region implicated
nv1 (121aa)	Transmembrane region (96-118)	69.6% nuclear	Endoplasmic reticulum	ND	17p13	ND
nv2 (195aa)	ND	47.8% nuclear	Nucleus, cytoplasm	CAG31028 *gg* (93%) NP_998313 *dr* (85%) CAG06569 *tn* (84%)	17q21	Autism, pervasive development disorder,susceptibility locus 21 (AUTS21) 17q11-17q21 (70).
nv4 (112aa)	ND	NLS (24) 73.9 % nuclear	Nucleus	XP_573878 *rn* (99%) NP_001002203 *dr* (67%) CAG07958 *tn* (63%)	19p13	Ectrodactyly, ectodermal dysplasia, cleft lip/palate, and urogenital defects (EEC2) 19p13-q12 (71); Postaxial polydactyly type A/B (PAPA/B) in chinese kindreds (PAPA3) 19p13.2-p13.1 (72); Dominant intermediate Charcot-Marie-Tooth neuropathy 1 (DICMT1) 19p13.2-p12 (73); Subtelomeric deletion of 19p13.3-pter. Phenotype includes cleft palate, distinctive facial appearance, heart defects, hearing impairment, (DEL19PD) (74)
nv5 (64aa)	Kruppel-associated box (KRAB) domain (4-44) present in C2H2 zinc finger proteins, involved in protein-protein interactions	47.8% cytoplasmic, 26.1% nuclear	Nucleus	ND	ND	ND
nv6 (165aa)	Zinc finger, C3HC4 type (RING) (88-128)- zinc-finger like motif, ubiquitin E3 ligase activity	39.1% nuclear, 26.1% cytoplasmic	Nucleus, cytoplasm	NP_001007648 *rn* (98%) NP_956600 *dr* (63%) AAH73002 *xl* (63%) CAG02493 *tn* (57%)	2p11.2	ND
nv9 partial transcript (1203aa)	NA	NA	NA	NA	11p13	Spinocerebellar ataxia 20 (SCA20), 11p13-q11 (75)
nv11 (285aa)	Coiled coil regions (67-95)(175-214); Glutamic acid-rich region (174-233)	34.8% cytoplasmic, 30.4% mitochondrial, 17.4% nuclear	Nucleus, mitochondria	NP_001014210 *rn* (98%) XP_414311 *gg* (89%) AAH72869 *xl* (83%) CAF93584 *tn* (71%) AAF54193 *dm* (61%) NP_505803 *ce* (50%)	3q22.1	Mobius Syndrome 2 (MBS2) 3q21-22 (76)
Trp53inp2 (221aa)	ND	4 residue NLS (202) 52.2% nuclear	Nucleus, perinuclear region	CAC82596 *rn* (98%) XP_852549 *cf* (84%)	20q11	Nonprogressive cerebellar ataxia 3 (CLA3) 20q11-q13 (77)
Zcchc3 (400aa)	Retrovirus zinc finger-like domains (318-366)(350-388)	65.2% cytoplasmic, 34.8% nuclear	NA	ND	20p12-13	ND
Hn1 (154aa)	ND	7 residue NLS (41) 65.2% nuclear	Nucleus, cytoplasm	AAH79311 *rn* (91%) XP_533120 *cf* (91%) CAG31139 *gg* (70%) NP_991176 *dr* (61%)	17q25.3	Desbuquois syndrome (DSBQS) 17q25.3 (78); Hereditary neuralgic amyotrophy (HNA) 17q25 (79)
Fto (502aa)	ND	Bipartite NLS (2)	Nucleus	XP_535301 *cf* (87%) AAH79990 *xl* (53%) CAG05424 *tn* (40%)	16p11.2-16p12	ND
Elp3 (547aa)	N-acetyltransferase domain (439-536) Radical S-adenosylmethionine domain (93-305)	30.4% cytoplasmic, 26.1% mitochondrial	Nucleus, mitochondria	XP_534563 *cf* (98%) CAG32722 *gg* (88%) AAH88956 *xl* (89%) CAI12039 *dr* (89%) NP_608834 *dm* (81%) NP_506055 *ce* (78%) NP_015239 *sc* (73%) NP_568725 *at* (74%)	8p21.1	ND
Limd1 (668aa)	LIM domains (464-522)(529-586)(589-655)- zinc-binding domain, binds protein partners via tyrosine-containing motifs, found in many key regulators of developmental pathways	60.9% nuclear, 21.7% cytoplasmic	Cytoplasm, focal adhesions	XP_236734 *rn* (92%) XP_541912 cf (78%) XP_418800 *gg* (53%) XP_691259 *dr* (65%)	3p21.3	ND
Thoc3 (351aa)	WD40 repeats (46-83)(90-128)(133-169)(215-252)(257-294)(299-335)- beta-propeller structure involved in assembly of multi-protein complexes	43.5% cytoplasmic, 17.4% nuclear	Aggresomes, vesicular structures	XP_537899 cf (99%) AAH79495 *dr* (93%) CAG32595 *gg* (92%) CAG04532 *tn* (89%) XP_237957 *rn* (98%) NP_649784 *dm* (63%), NP_200424 *at* (48%)	5q35.2	Deletion 5q35, distal to the Sotos syndrome associated deletion . (DEL5q35) 5q35.3 (80)
Maged2 (616aa)	Melanoma-associated antigen (MAGE) family domain (44-401)- expressed in tumor cells, testis and developing embryo	7 residue NLS (248), bipartite NLS (203)(215)(239) 78.3% nuclear	Nucleus, nucleoli, cytoplasmic endosomal-like structures, plasma membrane	NP_536727 *rn* (93%) XP_864804 cf (86%)	Xp11.2	X-linked mental retardation, investigated but no causative mutation established (64)
Psf2 (185aa)	ND	56.5% cytoplasmic, 26.1% nuclear	Nucleus, cytoplasm	XP_851786 cf (94%) AAH73044 *xl* (85%) CAB07195 *ce* (37%) AC069474 *at* (40%) NP_012463 *sc* (24%)	16q24	Progressive eye movement disorder with variable expression of bilateral ptosis and restrictive external ophthalmoplegia (FEOM3) 16q24.2-q24.3 (81)
nv13	NA	NA	NA	NA	7q11.23	Deafness, congenital neurosensory, autosomal recessive 39 (DFNB39) 7q11.22-q21.12 (82); Cerebellar atrophy with progressive microcephaly (CLAM) 7q11-q21 (83)

Abbreviations: aa, amino acid; Ac., accession; *at, Arabidopsis thaliana*; *ce, Caenorhabditis elegans* (nematode worm); *cf, Canis familiaris* (dog), *dr, Danio rerio* (zebrafish); *dm, Drosophila melanogaster* (fruitfly); *gg, Gallus gallus* (chick); NA, not analysed; ND, none detected; NLS, nuclear localisation signal; *rn, Rattus norvegicus* (rat); *sc, Sacchyromyces cerevisiae* (budding yeast); *tn, Tetraodon nigroviridus* (pufferfish); *xl, Xenopus laevis* (frog).

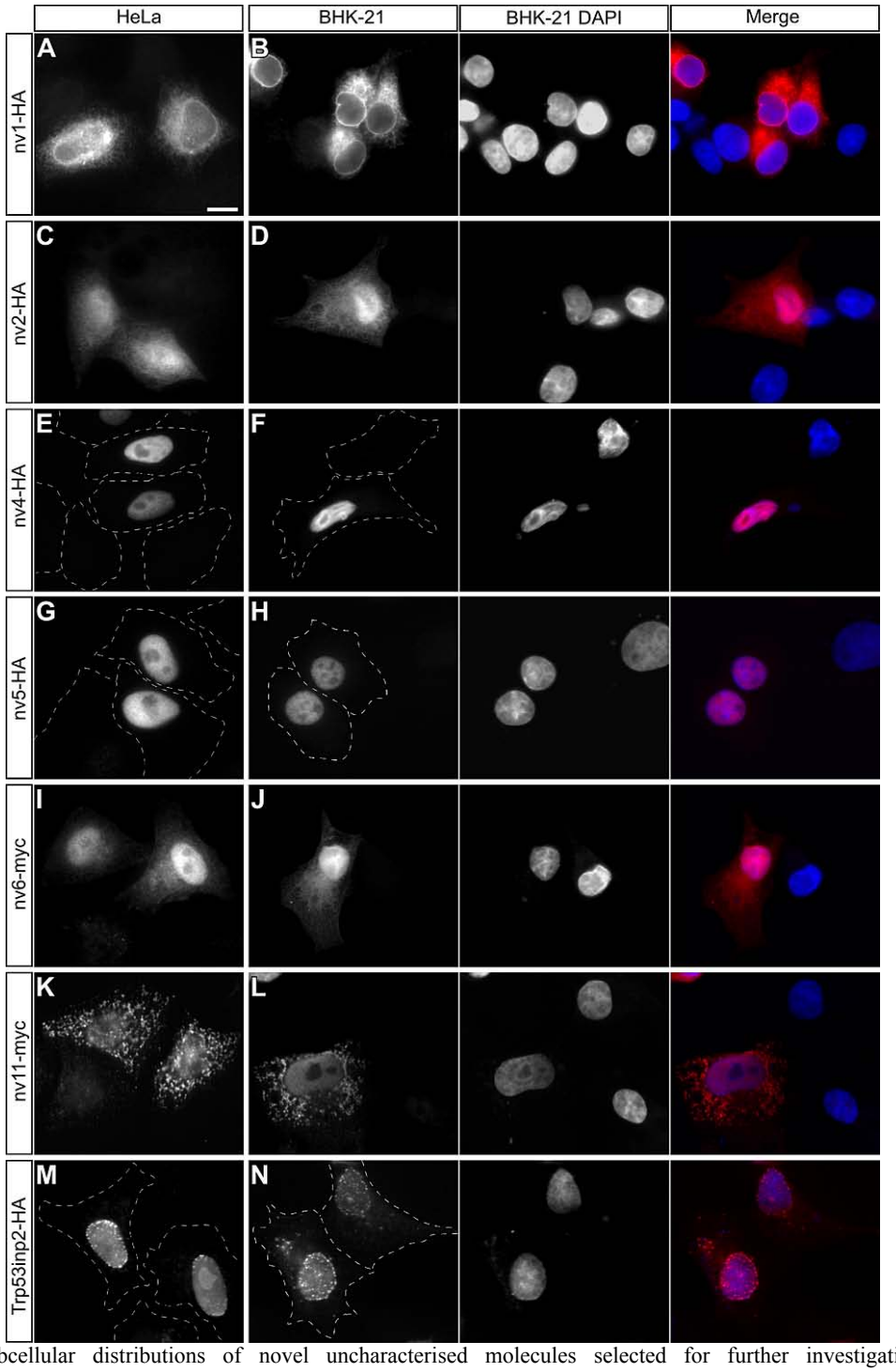

Figure 2. Subcellular distributions of novel uncharacterised molecules selected for further investigation from the developmentally enriched cDNA library: nv1 (A, B), nv2 (C, D), nv4 (E, F), nv5 (G, H), nv6 (I, J), nv11 (K, L), Trp53inp2 (M, N). Cultured BHK-21 and HeLa cells were transiently transfected with mouse HA or myc epitope-tagged full-length predicted coding sequence constructs for each molecule investigated. For BHK-21 transfections, DAPI stained and merged images are shown alongside images of transfected cells detected with monoclonal anti-HA or anti-myc antibody. nv1-HA appears to localise to endoplasmic reticulum (A, B), nv2-HA (C, D) and nv6-myc (I, J) show nuclear and cytoplasmic distribution, nv4-HA (E, F) and nv5-HA (G, H) appear to be predominantly nuclear, nv11-myc appears to be localised to the nucleus and mitochondria (K, L), and Trp53inp2 appears to be distributed to the nucleus and in puncta in the perinuclear region (M, N). All images are resized to equivalent scale, scale bar in (A) equals 10 microns.

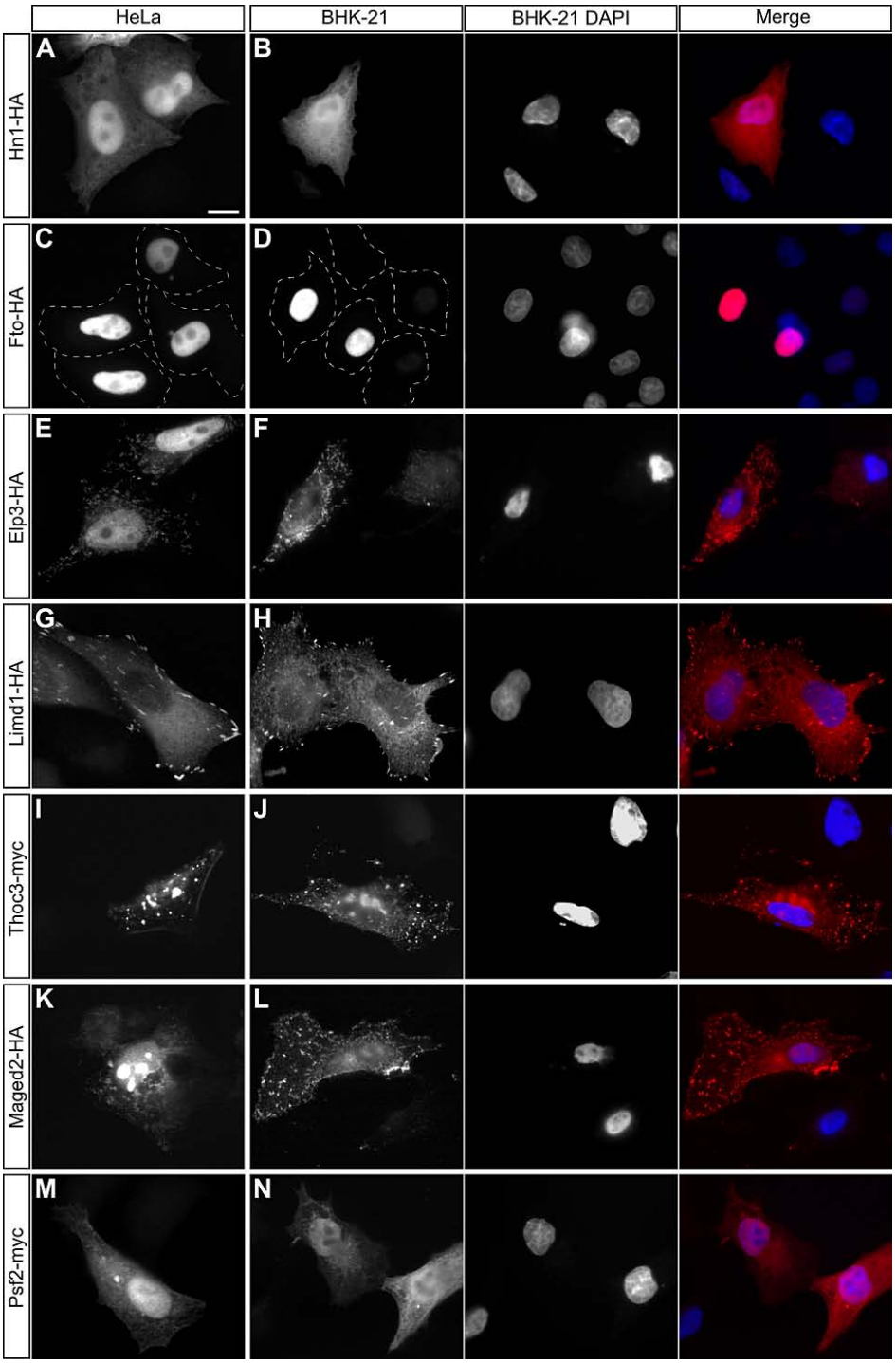

Figure 3. Subcellular localisations of poorly characterised molecules selected for further investigation from the developmentally enriched cDNA library: Hn1 (A, B), Fatso (C, D), Elp3 (E, F), Limd1 (G, H), Thoc3 (I, J), Maged2 (K, L), Psf2 (M, N). Cultured BHK-21 and HeLa cells were transiently transfected with mouse HA or myc epitope-tagged full-length predicted coding sequence constructs for each molecule investigated. For BHK-21 transfections, DAPI stained and merged images are shown alongside images of transfected cells detected with monoclonal anti-HA or anti-myc antibody. Hn1 (A, B) and Psf2 (M, N) show nuclear and cytoplasmic distribution, Fatso is predominantly nuclear (C, D), Elp3 appears to have a nuclear and mitochondrial localisation (E, F), Limd1 appears to be localised across the cytoplasm and at focal adhesions (G, H), Thoc3 exhibits aggresomal and vesicle-like distribution (I, J), and Maged2 is localised to the nucleus (concentrated in nucleoli) and cytoplasmic endosomal-like structures (K, L), and also appears to be membrane-associated (L). All images are resized to equivalent scale, scale bar in (A) equals 10 microns.

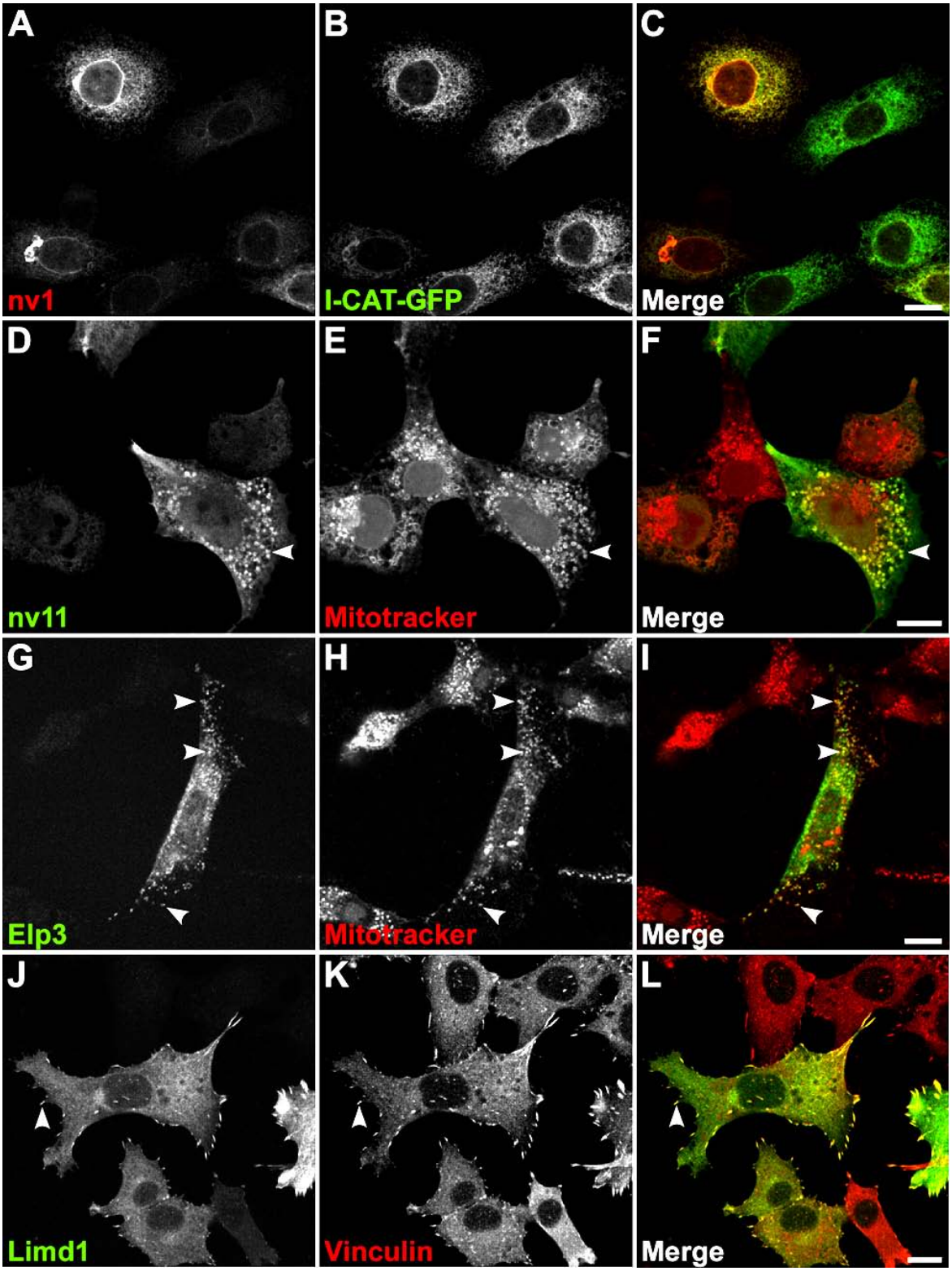

Figure 4. Co-localisations of a subset of transiently transfected novel/uncharacterised molecules with known markers for subcellular compartments and structures. Nv1-HA (A) colocalises with transfected I-CAT-GFP (B) within the endoplasmic reticulum (A-C). Nv11-myc (D) and Elp3-HA (G) colocalise with the mitochondrial marker MitoTracker Red (D-I). Limd1-HA (J) colocalises with vinculin (K) at focal adhesions (J-L). Arrows indicate observed colocalisation. Scale bars are equal to 10 microns.

Eleven of the 17 genes selected for further analysis represent molecules which are either novel or had no previously characterised functional role (*nv1/NM_134022, nv2/AK081685, nv4/NM_026455, nv5/AK013921, nv6/NM_025607, nv9/AK141733, nv11/NM_133934, Trp53inp2, Zcchc3, Hn1, Fto*). For one of these (*nv9/AK141733)* a partial protein is predicted, however sequence data was insufficient to determine a putative full-length transcript and corresponding protein. Five of the 17 candidates had published functional roles but had not been previously described in embryonic development (*Elp3, Limd1, Thoc3, Maged2, Psf2*). One molecule (*nv13*) was identified for which originally there were no corresponding proteins identified in the NCBI public database, although our embryonic expression data (Figure 1Q) suggested that it was likely to represent an expressed RNA transcript. The most recent mouse genome annotation (FANTOM III) suggests that *nv13* may represent a potential protein-coding transcript *AJ290943* (although this is not confidently predicted), but also identifies it as a complementary antisense sequence for the mRNA transcript of Rhbdl7 (Rhomboid, veinlet-like 7) (23). Rhbdl7 is a novel predicted transmembrane protein with similarity to *Drosophila* Rhomboid protein, a known integral membrane serine protease important in dorso-ventral axis establishment and development of the nervous system (24-26). Non-coding RNA transcripts could potentially play a role in regulation and modulation of expression of their complementary genes, thereby having a relatively unique and unexplored functional role in embryogenesis (reviewed in (27)). Alternatively this novel transcript could represent an expressed pseudogene, and this still remains to be determined.

5.2. Prediction of domains, subcellular localisation and transfection of tagged constructs

While many key developmental genes are known to be involved in signal transduction and transcriptional regulation, there is increasing evidence that other classes of molecules, such as those involved in cellular trafficking, play important roles in regulating the processes of embryogenesis (1, 28). Analysis of subcellular localisation of uncharacterised proteins has the potential to shed light on their function at the cellular level, and thereby provide insight into the role of these proteins in the developing embryo. For the 17 candidates selected, prediction tools were utilised to analyse amino acid sequences for known or predicted mouse proteins (summarised in Table 2). Additionally, tagged constructs for 14 protein-coding genes were generated and transiently transfected into mammalian cell lines to reveal subcellular localisation (Figure 2, Figure 3). Nine of the 14 proteins analysed showed a subcellular localisation which was consistent with that predicted by PsortII (nv2/NP_766535, nv4/NP_080731, nv6/NP_079883, nv11/NP_598695, Trp53inp2, Hn1, Fto, Maged2 and Psf2) and three of the remaining five had localisations that were consistent with the known functions of their predicted domains (nv5/NP_001001737, Elp3 and Limd1; Table 2). This correlation between predicted and experimental localisations supports the value of the utilisation of prediction tools in analysing protein sequences to glean informative data. For the remaining two proteins, predicted localisations did not agree with experimental localisations observed for over-expressed proteins (nv1/NP_598783 and Thoc3). This could be due to the fact that the myc/HA epitope tag may cause the protein to misfold and/or mislocalise, or because the cell type used was not physiologically relevant and correct localisation required the presence of interacting partners. Alternatively, the prediction databases used may not detect domains or motifs which have not yet been well characterised or for which amino acid identity of the query sequence is below threshold level.

Of the molecules analysed, four contain domains typical of and/or are known to be involved in transcriptional regulation, including nv5/NP_001001737 (Kruppel-associated box), nv6/NP_079883 (C3HC4 RING zinc finger), Elp3 (histone acetyltransferase), and Limd1 (Lim domains, known to repress E2F-driven transcription (29)). Transcriptional activation and repression via molecules such as transcription factors and chromatin modifiers is known to be a pivotal mechanism for co-ordination of pathways involved in embryonic development. Also, two proteins contain motifs typical of molecules involved in protein-protein interactions, nv11/NP_598695 (coiled-coils) and Thoc3 (WD40 repeats).

Concurrent with our study, Limd1 was described as a tumor suppressor which binds the retinoblastoma protein (pRB) and represses E2F-driven transcription (29). In the same study, it was also shown to shuttle between the nucleus and cytosol in U2OS cells (human osteosarcoma), and our study showed that in addition, this protein is localised to focal adhesions in BHK-21 and HeLa cells. This is consistent with the observed localisations of members of the same protein family, Ajuba, Zyxin, Lipoma Preferred Partner, Migfilin, Trip6 and Zrp1 (30-35), and with the findings of a very recent study published while our manuscript was in preparation, which also showed localisation of LimD1 to focal adhesions (36). Furthermore, a number of these related proteins have also been shown to shuttle between the nucleus, cytosol and focal adhesions, in a manner dependent on their interactions with various other molecules (reviewed in (37)). Focal adhesions function in cell migration, cell-substratum adhesion and intracellular signalling, processes essential to embryogenesis as well as a range of pathological processes including tumorigenesis and metastasis (reviewed in (38)).

5.3. Potential association of novel and poorly characterised molecules with human hereditary disorders

A focus of this study is the potential association of novel genes with human developmental disorders. To facilitate this, human orthologues for genes which showed restricted embryonic expression were mapped to human chromosomal loci using *in silico* techniques. Databases were searched to determine any disorders previously mapped to the same chromosomal regions which were characterised by appropriate phenotypes based primarily on sites of expression in mouse embryos. Of the 17 that were

investigated, ten genes were shown to map at or nearby chromosomal regions to which relevant human disorders have also been linked. These genes represent putative candidate genes responsible for these disorders, and a number of candidates are currently under investigation to search for potential disease-causing mutations. In other cases, further fine mapping of disease loci relative to these genes is required to assess the likelihood of their involvement.

6. CONCLUSIONS AND SUMMARY

Based on large-scale screening of a subtracted developmentally enriched pharyngeal arch library, we have identified a number of novel molecules expressed in specific regions of the developing mouse embryo, as well as molecules which had not been previously associated with developmental processes. A number of these molecules represent positional candidates for known developmental disorders based on their chromosomal localisation and sites of expression in the developing embryo. The use of a predictive *in silico* approach in combination with techniques to investigate gene expression alongside subcellular localisation provides a powerful tool to prioritise molecules for further characterisation, and can also hint at potential functional roles which are worthy of future investigation.

7. ACKNOWLEDGMENTS

The authors would like to thank Vicki Metzis for technical assistance and staff of the Physiology and Pharmacology SPF and Queensland Biosciences Precinct (The University of Queensland) animal houses for help with mouse husbandry and breeding. JSB is a recipient of an Australian Postgraduate Award, FS is a UQ Postdoctoral Fellow, CW is an Australian National Health and Medical Research Council (NHMRC) Senior Research Fellow and RDT is supported by an NHMRC R. Douglas Wright Career Development Award. This work was supported by an NHMRC project grant (CW). The IMB incorporates the Centre for Functional and Applied Genomics, a Special Research Centre of the Australian Research Council. Confocal microscopy was performed at the ACRF/IMB Dynamic Imaging Facility for Cancer Biology, established with funding from the Australian Cancer Research Foundation.

8. REFERENCES

1. Fowles, L. F., J. S. Bennetts, J. L. Berkman, E. Williams, P. Koopman, R. D. Teasdale & C. Wicking: Genomic screen for genes involved in mammalian craniofacial development. *Genesis* 35, 73-87 (2003)

2. Howard, T. D., W. A. Paznekas, E. D. Green, L. C. Chiang, N. Ma, R. I. Ortiz de Luna, C. Garcia Delgado, M. Gonzalez-Ramos, A. D. Kline & E. W. Jabs: Mutations in TWIST, a basic helix-loop-helix transcription factor, in Saethre-Chotzen syndrome. *Nat Genet* 15, 36-41 (1997)

3. Wilkie, A. O., S. F. Slaney, M. Oldridge, M. D. Poole, G. J. Ashworth, A. D. Hockley, R. D. Hayward, D. J. David, L. J. Pulleyn, P. Rutland & et al.: Apert syndrome results from localized mutations of FGFR2 and is allelic with Crouzon syndrome. *Nat Genet* 9, 165-72 (1995)

4. Kang, S., J. M. Graham, Jr., A. H. Olney & L. G. Biesecker: GLI3 frameshift mutations cause autosomal dominant Pallister-Hall syndrome. *Nat Genet* 15, 266-8 (1997)

5. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller & D. J. Lipman: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25, 3389-402 (1997)

6. Altschul, S. F., W. Gish, W. Miller, E. W. Myers & D. J. Lipman: Basic local alignment search tool. *J Mol Biol* 215, 403-10 (1990)

7. Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. C. Frith, N. Maeda, R. Oyama, T. Ravasi, B. Lenhard, C. Wells, R. Kodzius, K. Shimokawa, V. B. Bajic, S. E. Brenner, S. Batalov, A. R. Forrest, M. Zavolan, M. J. Davis, L. G. Wilming, V. Aidinis, J. E. Allen, A. Ambesi-Impiombato, R. Apweiler, R. N. Aturaliya, T. L. Bailey, M. Bansal, L. Baxter, K. W. Beisel, T. Bersano, H. Bono, A. M. Chalk, K. P. Chiu, V. Choudhary, A. Christoffels, D. R. Clutterbuck, M. L. Crowe, E. Dalla, B. P. Dalrymple, B. de Bono, G. Della Gatta, D. di Bernardo, T. Down, P. Engstrom, M. Fagiolini, G. Faulkner, C. F. Fletcher, T. Fukushima, M. Furuno, S. Futaki, M. Gariboldi, P. Georgii-Hemming, T. R. Gingeras, T. Gojobori, R. E. Green, S. Gustincich, M. Harbers, Y. Hayashi, T. K. Hensch, N. Hirokawa, D. Hill, L. Huminiecki, M. Iacono, K. Ikeo, A. Iwama, T. Ishikawa, M. Jakt, A. Kanapin, M. Katoh, Y. Kawasawa, J. Kelso, H. Kitamura, H. Kitano, G. Kollias, S. P. Krishnan, A. Kruger, S. K. Kummerfeld, I. V. Kurochkin, L. F. Lareau, D. Lazarevic, L. Lipovich, J. Liu, S. Liuni, S. McWilliam, M. Madan Babu, M. Madera, L. Marchionni, H. Matsuda, S. Matsuzawa, H. Miki, F. Mignone, S. Miyake, K. Morris, S. Mottagui-Tabar, N. Mulder, N. Nakano, H. Nakauchi, P. Ng, R. Nilsson, S. Nishiguchi, S. Nishikawa, F. Nori, O. Ohara, Y. Okazaki, V. Orlando, K. C. Pang, W. J. Pavan, G. Pavesi, G. Pesole, N. Petrovsky, S. Piazza, J. Reed, J. F. Reid, B. Z. Ring, M. Ringwald, B. Rost, Y. Ruan, S. L. Salzberg, A. Sandelin, C. Schneider, C. Schonbach, K. Sekiguchi, C. A. Semple, S. Seno, L. Sessa, Y. Sheng, Y. Shibata, H. Shimada, K. Shimada, D. Silva, B. Sinclair, S. Sperling, E. Stupka, K. Sugiura, R. Sultana, Y. Takenaka, K. Taki, K. Tammoja, S. L. Tan, S. Tang, M. S. Taylor, J. Tegner, S. A. Teichmann, H. R. Ueda, E. van Nimwegen, R. Verardo, C. L. Wei, K. Yagi, H. Yamanishi, E. Zabarovsky, S. Zhu, A. Zimmer, W. Hide, C. Bult, S. M. Grimmond, R. D. Teasdale, E. T. Liu, V. Brusic, J. Quackenbush, C. Wahlestedt, J. S. Mattick, D. A. Hume, C. Kai, D. Sasaki, Y. Tomaru, S. Fukuda, M. Kanamori-Katayama, M. Suzuki, J. Aoki, T. Arakawa, J. Iida, K. Imamura, M. Itoh, T. Kato, H. Kawaji, N. Kawagashira, T. Kawashima, M. Kojima, S. Kondo, H. Konno, K. Nakano, N. Ninomiya, T. Nishio, M. Okada, C. Plessy, K. Shibata, T. Shiraki, S. Suzuki, M. Tagami, K. Waki, A. Watahiki, Y. Okamura-Oho, H. Suzuki, J. Kawai & Y. Hayashizaki: The transcriptional landscape of the mammalian genome. *Science* 309, 1559-63 (2005)

8. Diehn, M., G. Sherlock, G. Binkley, H. Jin, J. C. Matese, T. Hernandez-Boussard, C. A. Rees, J. M. Cherry, D. Botstein, P. O. Brown & A. A. Alizadeh: SOURCE: a unified genomic resource of functional annotations,

ontologies, and gene expression data. *Nucleic Acids Res* 31, 219-23 (2003)
9. Christiansen, J. H., C. L. Dennis, C. A. Wicking, S. J. Monkley, D. G. Wilkinson & B. J. Wainwright: Murine Wnt-11 and Wnt-12 have temporally and spatially restricted expression patterns during embryonic development. *Mech Dev* 51, 341-50 (1995)
10. Carninci, P. & Y. Hayashizaki: High-efficiency full-length cDNA cloning. *Methods Enzymol* 303, 19-44 (1999)
11. Kawai, J., A. Shinagawa, K. Shibata, M. Yoshino, M. Itoh, Y. Ishii, T. Arakawa, A. Hara, Y. Fukunishi, H. Konno, J. Adachi, S. Fukuda, K. Aizawa, M. Izawa, K. Nishi, H. Kiyosawa, S. Kondo, I. Yamanaka, T. Saito, Y. Okazaki, T. Gojobori, H. Bono, T. Kasukawa, R. Saito, K. Kadota, H. Matsuda, M. Ashburner, S. Batalov, T. Casavant, W. Fleischmann, T. Gaasterland, C. Gissi, B. King, H. Kochiwa, P. Kuehl, S. Lewis, Y. Matsuo, I. Nikaido, G. Pesole, J. Quackenbush, L. M. Schriml, F. Staubli, R. Suzuki, M. Tomita, L. Wagner, T. Washio, K. Sakai, T. Okido, M. Furuno, H. Aono, R. Baldarelli, G. Barsh, J. Blake, D. Boffelli, N. Bojunga, P. Carninci, M. F. de Bonaldo, M. J. Brownstein, C. Bult, C. Fletcher, M. Fujita, M. Gariboldi, S. Gustincich, D. Hill, M. Hofmann, D. A. Hume, M. Kamiya, N. H. Lee, P. Lyons, L. Marchionni, J. Mashima, J. Mazzarelli, P. Mombaerts, P. Nordone, B. Ring, M. Ringwald, I. Rodriguez, N. Sakamoto, H. Sasaki, K. Sato, C. Schonbach, T. Seya, Y. Shibata, K. F. Storch, H. Suzuki, K. Toyo-oka, K. H. Wang, C. Weitz, C. Whittaker, L. Wilming, A. Wynshaw-Boris, K. Yoshida, Y. Hasegawa, H. Kawaji, S. Kohtsuki & Y. Hayashizaki: Functional annotation of a full-length mouse cDNA collection. *Nature* 409, 685-90 (2001)
12. Okazaki, Y., M. Furuno, T. Kasukawa, J. Adachi, H. Bono, S. Kondo, I. Nikaido, N. Osato, R. Saito, H. Suzuki, I. Yamanaka, H. Kiyosawa, K. Yagi, Y. Tomaru, Y. Hasegawa, A. Nogami, C. Schonbach, T. Gojobori, R. Baldarelli, D. P. Hill, C. Bult, D. A. Hume, J. Quackenbush, L. M. Schriml, A. Kanapin, H. Matsuda, S. Batalov, K. W. Beisel, J. A. Blake, D. Bradt, V. Brusic, C. Chothia, L. E. Corbani, S. Cousins, E. Dalla, T. A. Dragani, C. F. Fletcher, A. Forrest, K. S. Frazer, T. Gaasterland, M. Gariboldi, C. Gissi, A. Godzik, J. Gough, S. Grimmond, S. Gustincich, N. Hirokawa, I. J. Jackson, E. D. Jarvis, A. Kanai, H. Kawaji, Y. Kawasawa, R. M. Kedzierski, B. L. King, A. Konagaya, I. V. Kurochkin, Y. Lee, B. Lenhard, P. A. Lyons, D. R. Maglott, L. Maltais, L. Marchionni, L. McKenzie, H. Miki, T. Nagashima, K. Numata, T. Okido, W. J. Pavan, G. Pertea, G. Pesole, N. Petrovsky, R. Pillai, J. U. Pontius, D. Qi, S. Ramachandran, T. Ravasi, J. C. Reed, D. J. Reed, J. Reid, B. Z. Ring, M. Ringwald, A. Sandelin, C. Schneider, C. A. Semple, M. Setou, K. Shimada, R. Sultana, Y. Takenaka, M. S. Taylor, R. D. Teasdale, M. Tomita, R. Verardo, L. Wagner, C. Wahlestedt, Y. Wang, Y. Watanabe, C. Wells, L. G. Wilming, A. Wynshaw-Boris, M. Yanagisawa, I. Yang, L. Yang, Z. Yuan, M. Zavolan, Y. Zhu, A. Zimmer, P. Carninci, N. Hayatsu, T. Hirozane-Kishikawa, H. Konno, M. Nakamura, N. Sakazume, K. Sato, T. Shiraki, K. Waki, J. Kawai, K. Aizawa, T. Arakawa, S. Fukuda, A. Hara, W. Hashizume, K. Imotani, Y. Ishii, M. Itoh, I. Kagawa, A. Miyazaki, K. Sakai, D. Sasaki, K. Shibata, A. Shinagawa, A. Yasunishi, M. Yoshino, R. Waterston, E. S. Lander, J. Rogers, E. Birney & Y. Hayashizaki: Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* 420, 563-73 (2002)
13. Kawai, J. & Y. Hayashizaki: DNA book. *Genome Res* 13, 1488-95 (2003)
14. Sambrook, J., E. F. Fritsch & T. Maniatis: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, (1989)
15. Dai, Q., E. Choy, V. Chiu, J. Romano, S. R. Slivka, S. A. Steitz, S. Michaelis & M. R. Philips: Mammalian prenylcysteine carboxyl methyltransferase is in the endoplasmic reticulum. *J Biol Chem* 273, 15030-4 (1998)
16. Simpson, F., K. Lammerts van Bueren, N. Butterfield, J. S. Bennetts, J. Bowles, C. Adolphe, L. A. Simms, J. Young, M. D. Walsh, B. Leggett, L. F. Fowles & C. Wicking: The PCNA-associated factor KIAA0101/p15^{PAF} binds the potential tumor suppressor product p33ING1b. *Exp Cell Res* 312, 73-85 (2006)
17. Schneider, R. A., D. Hu & J. A. Helms: From head to toe: conservation of molecular signals regulating limb and craniofacial morphogenesis. *Cell Tissue Res* 296, 103-9 (1999)
18. Gordon, M. D., M. S. Dionne, D. S. Schneider & R. Nusse: WntD is a feedback inhibitor of Dorsal/NF-kappaB in Drosophila development and immunity. *Nature* 437, 746-9 (2005)
19. Dunker, N. & K. Krieglstein: Targeted mutations of transforming growth factor-beta genes reveal important roles in mouse development and adult homeostasis. *Eur J Biochem* 267, 6982-8 (2000)
20. Rudel, D. & R. J. Sommer: The evolution of developmental mechanisms. *Dev Biol* 264, 15-37 (2003)
21. Helms, J. A., D. Cordero & M. D. Tapadia: New insights into craniofacial morphogenesis. *Development* 132, 851-61 (2005)
22. Niswander, L.: Interplay between the molecular signals that control vertebrate limb development. *Int J Dev Biol* 46, 877-81 (2002)
23. Katayama, S., Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, M. Nakamura, H. Nishida, C. C. Yap, M. Suzuki, J. Kawai, H. Suzuki, P. Carninci, Y. Hayashizaki, C. Wells, M. Frith, T. Ravasi, K. C. Pang, J. Hallinan, J. Mattick, D. A. Hume, L. Lipovich, S. Batalov, P. G. Engstrom, Y. Mizuno, M. A. Faghihi, A. Sandelin, A. M. Chalk, S. Mottagui-Tabar, Z. Liang, B. Lenhard & C. Wahlestedt: Antisense transcription in the mammalian transcriptome. *Science* 309, 1564-6 (2005)
24. Bier, E., L. Y. Jan & Y. N. Jan: rhomboid, a gene required for dorsoventral axis establishment and peripheral nervous system development in Drosophila melanogaster. *Genes Dev* 4, 190-203 (1990)
25. Urban, S., J. R. Lee & M. Freeman: Drosophila rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell* 107, 173-82 (2001)
26. Sturtevant, M. A., M. Roark & E. Bier: The Drosophila rhomboid gene mediates the localized formation of wing veins and interacts genetically with components of the EGF-R signaling pathway. *Genes Dev* 7, 961-73 (1993)
27. Mattick, J. S. & I. V. Makunin: Small regulatory RNAs in mammals. *Hum Mol Genet* 14 Spec No 1, R121-32 (2005)
28. Seto, E. S., H. J. Bellen & T. E. Lloyd: When cell biology meets development: endocytic regulation of signaling pathways. *Genes Dev* 16, 1314-36 (2002)

29. Sharp, T. V., F. Munoz, D. Bourboulia, N. Presneau, E. Darai, H. W. Wang, M. Cannon, D. N. Butcher, A. G. Nicholson, G. Klein, S. Imreh & C. Boshoff: LIM domains-containing protein 1 (LIMD1), a tumor suppressor encoded at chromosome 3p21.3, binds pRB and represses E2F-driven transcription. *Proc Natl Acad Sci U S A* 101, 16531-6 (2004)
30. Kanungo, J., S. J. Pratt, H. Marie & G. D. Longmore: Ajuba, a cytosolic LIM protein, shuttles into the nucleus and affects embryonal cell proliferation and fate decisions. *Mol Biol Cell* 11, 3299-313 (2000)
31. Petit, M. M., S. M. Meulemans & W. J. Van de Ven: The focal adhesion and nuclear targeting capacity of the LIM-containing lipoma-preferred partner (LPP) protein. *J Biol Chem* 278, 2157-68 (2003)
32. Tu, Y., S. Wu, X. Shi, K. Chen & C. Wu: Migfilin and Mig-2 link focal adhesions to filamin and the actin cytoskeleton and function in cell shape modulation. *Cell* 113, 37-47 (2003)
33. Li, B., L. Zhuang, M. Reinhard & B. Trueb: The lipoma preferred partner LPP interacts with alpha-actinin. *J Cell Sci* 116, 1359-66 (2003)
34. Wang, Y. & T. D. Gilmore: LIM domain protein Trip6 has a conserved nuclear export signal, nuclear targeting sequences, and multiple transactivation domains. *Biochim Biophys Acta* 1538, 260-72 (2001)
35. Sanz-Rodriguez, F., M. Guerrero-Esteo, L. M. Botella, D. Banville, C. P. Vary & C. Bernabeu: Endoglin regulates cytoskeletal organization through binding to ZRP-1, a member of the Lim family of proteins. *J Biol Chem* 279, 32858-68 (2004)
36. Petit, M. M., K. R. Crombez, H. B. Vervenne, N. Weyns & W. J. Van de Ven: The tumor suppressor Scrib selectively interacts with specific members of the zyxin family of proteins. *FEBS Lett* 579, 5061-8 (2005)
37. Kadrmas, J. L. & M. C. Beckerle: The LIM domain: from the cytoskeleton to the nucleus. *Nat Rev Mol Cell Biol* 5, 920-31 (2004)
38. Sastry, S. K. & K. Burridge: Focal adhesions: a nexus for intracellular signaling and cytoskeletal dynamics. *Exp Cell Res* 261, 25-36 (2000)
39. Sultana, R., C. E. Yu, J. Yu, J. Munson, D. Chen, W. Hua, A. Estes, F. Cortes, F. de la Barra, D. Yu, S. T. Haider, B. J. Trask, E. D. Green, W. H. Raskind, C. M. Disteche, E. Wijsman, G. Dawson, D. R. Storm, G. D. Schellenberg & E. C. Villacres: Identification of a novel gene on chromosome 7q11.2 interrupted by a translocation breakpoint in a pair of autistic twins. *Genomics* 80, 129-34 (2002)
40. Petrakis, T. G., T. M. M. Sogaard, H. Erdjument-Bromage, P. Tempst & J. Q. Svejstrup: Physical and functional interaction between elongator and the chromatin-associated Kti12 protein. *The Journal of Biological Chemistry* 280, 19454-60 (2005)
41. Ben-Yehuda, S., I. Dix, C. S. Russell, M. McGarvey, J. D. Beggs & M. Kupiec: Genetic and physical interactions between factors involved in both cell cycle progression and pre-mRNA splicing in Saccharomyces cerevisiae. *Genetics* 156, 1503-17 (2000)
42. Dix, I., C. Russell, S. B. Yehuda, M. Kupiec & J. D. Beggs: The identification and characterization of a novel splicing protein, Isy1p, of Saccharomyces cerevisiae. *Rna* 5, 360-8 (1999)
43. Imai, Y., A. Nakada, R. Hashida, Y. Sugita, T. Tanaka, G. Tsujimoto, K. Matsumoto, A. Akasawa, H. Saito & T. Oshida: Cloning and characterization of the highly expressed ETEA gene from blood cells of atopic dermatitis patients. *Biochemical and Biophysical Research Communications*1282-1290 (2002)
44. Scanlan, M. J., J. D. Gordan, B. Williamson, E. Stockert, N. H. Bander, V. Jongeneel, A. O. Gure, D. Jager, E. Jager, A. Knuth, Y. T. Chen & L. J. Old: Antigens recognized by autologous antibody in patients with renal-cell carcinoma. *Int J Cancer* 83, 456-64 (1999)
45. Abdullah, J. M., X. Jing, D. S. Spassov, R. G. Nachtman & R. Jurecic: Cloning and characterization of Hepp, a novel gene expressed preferentially in hematopoietic progenitors and mature blood cells. *Blood Cells Mol Dis* 27, 667-76 (2001)
46. Xu, X. L., J. Yu, H. Y. Zhang, M. H. Sun, J. Gu, X. Du, D. R. Shi, P. Wang, Z. H. Yang & J. D. Zhu: Methylation profile of the promoter CpG islands of 31 genes that may contribute to colorectal carcinogenesis. *World J Gastroenterol* 10, 3441-54 (2004)
47. Zhou, G., J. Wang, Y. Zhang, C. Zhong, J. Ni, L. Wang, J. Guo, K. Zhang, L. Yu & S. Zhao: Cloning, expression and subcellular localization of HN1 and HN1L genes, as well as characterization of their orthologs, defining an evolutionarily conserved gene family. *Gene* 331, 115-23 (2004)
48. Tang, W., Y. H. Lai, X. D. Han, P. M. Wong, L. L. Peters & D. H. Chui: Murine Hn1 on chromosome 11 is expressed in hemopoietic and brain tissues. *Mamm Genome* 8, 695-6 (1997)
49. Lu, K. H., A. P. Patterson, L. Wang, R. T. Marquez, E. N. Atkinson, K. A. Baggerly, L. R. Ramoth, D. G. Rosen, J. Liu, I. Hellstrom, D. Smith, L. Hartmann, D. Fishman, A. Berchuck, R. Schmandt, R. Whitaker, D. M. Gershenson, G. B. Mills & R. C. Bast, Jr.: Selection of potential markers for epithelial ovarian cancer with gene expression arrays and recursive descent partition analysis. *Clin Cancer Res* 10, 3291-300 (2004)
50. Peters, T., K. Ausmeier & U. Ruther: Cloning of Fatso (Fto), a novel gene deleted by the Fused toes (Ft) mouse mutation. *Mamm Genome* 10, 983-6 (1999)
51. Wang, A. G., C. H. Chen, C. W. Yang, M. Y. Yen, W. M. Hsu, J. H. Liu & M. J. Fann: Change of gene expression profiles in the retina following optic nerve injury. *Brain Res Mol Brain Res* 101, 82-92 (2002)
52. Uchida, N., S. Hoshino & T. Katada: Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein. *J Biol Chem* 279, 1383-91 (2004)
53. Boeck, R., S. Tarun, Jr., M. Rieger, J. A. Deardorff, S. Muller-Auer & A. B. Sachs: The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. *J Biol Chem* 271, 432-8 (1996)
54. Chinenov, Y.: A second catalytic domain in the Elp3 histone acetyltransferases: a candidate for histone demathylase activity? *TRENDS in Biochemical Sciences* 27, 115-117 (2002)
55. Li, F., J. Lu, Q. Han, G. Zhang & B. Huang: The Elp3 subunit of human Elongator complex is functionally similar to its counterpart in yeast. *Molecular Genetics and Genomics* 273, 264-72 (2005)

56. Osthus, R. C., B. Karim, J. E. Prescott, B. D. Smith, M. McDevitt, D. L. Huso & C. V. Dang: The Myc target gene JPO1/CDCA7 is frequently overexpressed in human tumors and has limited transforming activity in vivo. *Cancer Res* 65, 5620-7 (2005)
57. Prescott, J. E., R. C. Osthus, L. A. Lee, B. C. Lewis, H. Shim, J. F. Barrett, Q. Guo, A. L. Hawkins, C. A. Griffin & C. V. Dang: A novel c-Myc-responsive gene, JPO1, participates in neoplastic transformation. *J Biol Chem* 276, 48276-84 (2001)
58. Haggerty, T. J., K. I. Zeller, R. C. Osthus, D. R. Wonsey & C. V. Dang: A strategy for identifying transcription factor binding sites reveals two classes of genomic c-Myc target sites. *Proc Natl Acad Sci U S A* 100, 5313-8 (2003)
59. Kiss, H., D. Kedra, Y. Yang, M. Kost-Alimova, C. Kiss, K. P. O'Brien, I. Fransson, G. Klein, S. Imreh & J. P. Dumanski: A novel gene containing LIM domains (LIMD1) is located within the common eliminated region 1 (C3CER1) in 3p21.3. *Hum Genet* 105, 552-9 (1999)
60. Goto, E., S. Ishido, Y. Sato, S. Ohgimoto, K. Ohgimoto, M. Nagano-Fujii & H. Hotta: c-MIR, a human E3 ubiquitin ligase, is a functional homolog of herpesvirus proteins MIR1 and MIR2 and has similar activity. *J Biol Chem* 278, 14657-68 (2003)
61. Gwizdek, C., M. Hobeika, B. Kus, B. Ossareh-Nazari, C. Dargemont & M. S. Rodriguez: The mRNA nuclear export factor Hpr1 is regulated by Rsp5-mediated ubiquitylation. *J Biol Chem* 280, 13401-5 (2005)
62. Rehwinkel, J., A. Herold, K. Gari, T. Kocher, M. Rode, F. L. Ciccarelli, M. Wilm & E. Izaurralde: Genome-wide analysis of mRNAs regulated by the THO complex in Drosophila melanogaster. *Nat Struct Mol Biol* 11, 558-66 (2004)
63. Jimeno, S., A. G. Rondon, R. Luna & A. Aguilera: The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *Embo J* 21, 3526-35 (2002)
64. Langnaese, K., D. U. Kloos, M. Wehnert, B. Seidel & P. Wieacker: Expression pattern and further characterization of human MAGED2 and identification of rodent orthologues. *Cytogenet Cell Genet* 94, 233-40 (2001)
65. Bertrand, M., I. Huijbers, P. Chomez & O. De Backer: Comparative expression analysis of the MAGED genes during embryogenesis and brain development. *Dev Dyn* 230, 325-34 (2004)
66. Piek, E., M. Van Dinther, W. T. Parks, J. M. Sallee, E. P. Bottinger, A. B. Roberts & P. Ten Dijke: RLP, a novel Ras-like protein, is an immediate-early transforming growth factor-beta (TGF-beta) target gene that negatively regulates transcriptional activity induced by TGF-beta. *Biochem J* 383, 187-99 (2004)
67. Takayama, Y., Y. Kamimura, M. Okawa, S. Muramatsu, A. Sugino & H. Araki: GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev* 17, 1153-65 (2003)
68. Walter, B. E. & J. J. Henry: Embryonic expression of pre-initiation DNA replication factors in Xenopus laevis. *Gene Expr Patterns* 5, 81-9 (2004)
69. Obama, K., K. Ura, S. Satoh, Y. Nakamura & Y. Furukawa: Up-regulation of PSF2, a member of the GINS multiprotein complex, in intrahepatic cholangiocarcinoma. *Oncol Rep* 14, 701-6 (2005)
70. Cantor, R. M., N. Kono, J. A. Duvall, A. Alvarez-Retuerto, J. L. Stone, M. Alarcon, S. F. Nelson & D. H. Geschwind: Replication of autism linkage: fine-mapping peak at 17q21. *Am J Hum Genet* 76, 1050-6 (2005)
71. O'Quinn, J. R., R. C. Hennekam, L. B. Jorde & M. Bamshad: Syndromic ectrodactyly with severe limb, ectodermal, urogenital, and palatal defects maps to chromosome 19. *Am J Hum Genet* 62, 130-5 (1998)
72. Zhao, H., Y. Tian, G. Breedveld, S. Huang, Y. Zou, J. Y, J. Chai, H. Li, M. Li, B. A. Oostra, W. H. Lo & P. Heutink: Postaxial polydactyly type A/B (PAP-A/B) is linked to chromosome 19p13.1-13.2 in a Chinese kindred. *Eur J Hum Genet* 10, 162-6 (2002)
73. Kennerson, M. L., D. Zhu, R. J. Gardner, E. Storey, J. Merory, S. P. Robertson & G. A. Nicholson: Dominant intermediate Charcot-Marie-Tooth neuropathy maps to chromosome 19p12-p13.2. *Am J Hum Genet* 69, 883-8 (2001)
74. Archer, H. L., S. Gupta, S. Enoch, P. Thompson, A. Rowbottom, I. Chua, S. Warren, D. Johnson, D. H. Ledbetter, C. Lese-Martin, P. Williams & D. T. Pilz: Distinct phenotype associated with a cryptic subtelomeric deletion of 19p13.3-pter. *Am J Med Genet A* 136, 38-44 (2005)
75. Knight, M. A., R. J. Gardner, M. Bahlo, T. Matsuura, J. A. Dixon, S. M. Forrest & E. Storey: Dominantly inherited ataxia and dysphonia with dentate calcification: spinocerebellar ataxia type 20. *Brain* 127, 1172-81 (2004)
76. Kremer, H., L. P. Kuyt, B. van den Helm, M. van Reen, J. A. Leunissen, B. C. Hamel, C. Jansen, E. C. Mariman, R. R. Frants & G. W. Padberg: Localization of a gene for Mobius syndrome to chromosome 3q by linkage analysis in a Dutch family. *Hum Mol Genet* 5, 1367-71 (1996)
77. Tranebjaerg, L., T. M. Teslovich, M. Jones, M. M. Barmada, T. Fagerheim, A. Dahl, D. M. Escolar, J. M. Trent, E. M. Gillanders & D. A. Stephan: Genome-wide homozygosity mapping localizes a gene for autosomal recessive non-progressive infantile ataxia to 20q11-q13. *Hum Genet* 113, 293-5 (2003)
78. Faivre, L., M. Le Merrer, L. I. Al-Gazali, M. G. Ausems, P. Bitoun, D. Bacq, P. Maroteaux, A. Munnich & V. Cormier-Daire: Homozygosity mapping of a Desbuquois dysplasia locus to chromosome 17q25.3. *J Med Genet* 40, 282-4 (2003)
79. Meulemann, J., G. Kuhlenbaumer, A. Schirmacher, M. Wehnert, P. De Jonghe, E. De Vriendt, P. Young, E. Airaksinen, A. Pou-Serradell, J. M. Prats, B. Ringelstein, F. Stogbauer, C. Van Broeckhoven & V. Timmerman: Genetic refinement of the hereditary neuralgic amyotrophy (HNA) locus at chromosome 17q25. *Eur J Hum Genet* 7, 920-7 (1999)
80. Rauch, A., M. Beese, E. Mayatepek, H. G. Dorr, D. Wenzel, A. Reis & U. Trautmann: A novel 5q35.3 subtelomeric deletion syndrome. *Am J Med Genet A* 121, 1-8 (2003)
81. Doherty, E. J., M. E. Macy, S. M. Wang, C. P. Dykeman, M. T. Melanson & E. C. Engle: CFEOM3: a new extraocular congenital fibrosis syndrome that maps to 16q24.2-q24.3. *Invest Ophthalmol Vis Sci* 40, 1687-94 (1999)

82. Wajid, M., A. A. Abbasi, M. Ansar, T. L. Pham, K. Yan, S. Haque, W. Ahmad & S. M. Leal: DFNB39, a recessive form of sensorineural hearing impairment, maps to chromosome 7q11.22-q21.12. *Eur J Hum Genet* 11, 812-5 (2003)
83. Rajab, A., G. H. Mochida, A. Hill, V. Ganesh, A. Bodell, A. Riaz, P. E. Grant, Y. Y. Shugart & C. A. Walsh: A novel form of pontocerebellar hypoplasia maps to chromosome 7q11-21. *Neurology* 60, 1664-7 (2003)

Abbreviations: BLAST, basic local alignment search tool; DIG, digoxygenin; E10.5, embryonic day 10.5; GFP, green fluorescent protein; HA, hemagglutinin; nv, novel molecule; ORF, open reading frame; PCR, polymerase chain reaction; UTR, untranslated region.

Key Words: Mouse, Gene, Craniofacial Development, Pharyngeal Arches, Novel Genes, *In Situ* Hybridisation, Subcellular Localisation

Send correspondence to: Dr Carol Wicking, Institute for Molecular Bioscience, The University of Queensland, St Lucia 4072, Queensland, Australia. Tel: 617-3346-2052, Fax: 617-3346-2101, E-mail: c.wicking@imb.uq.edu.au

http://www.bioscience.org/current/vol11.htm