European Journal of Human Genetics (2001) 9, 892–902

www.nature.com/ejhg

ARTICLE

Genomic organisation of the ~1.5 Mb Smith-Magenis syndrome critical interval: Transcription map, genomic contig, and candidate gene analysis

Rebecca E Lucas[1], Christopher N Vlangos[1], Parimal Das[4], Pragna I Patel[4] and Sarah H Elsea*[,1,2,3]

[1]*Genetics Graduate Program, Michigan State University, East Lansing, Michigan, MI 48824, USA;* [2]*Department of Zoology, Michigan State University, East Lansing, Michigan, MI 48824, USA;* [3]*Department of Pediatrics and Human Development, Michigan State University, East Lansing, Michigan, MI 48824, USA;* [4]*Department of Neurology, Baylor College of Medicine, Houston, Texas, TX 77030, USA*

Smith-Magenis syndrome (SMS) is a multiple congenital anomalies/mental retardation syndrome associated with an interstitial deletion of chromosome 17 involving band p11.2. SMS is hypothesised to be a contiguous gene syndrome in which the phenotype arises from the haploinsufficiency of multiple, functionally-unrelated genes in close physical proximity, although the true molecular basis of SMS is not yet known. In this study, we have generated the first overlapping and contiguous transcription map of the SMS critical interval, linking the proximal 17p11.2 region near the SMS-REPM and the distal region near D17S740 in a minimum tiling path of 16 BACs and two PACs. Additional clones provide greater coverage throughout the critical region. Not including the repetitive sequences that flank the critical interval, the map is comprised of 13 known genes, 14 ESTs, and six genomic markers, and is a synthesis of Southern hybridisation and polymerase chain reaction data from gene and marker localisation to BACs and PACs and database sequence analysis from the human genome project high-throughput draft sequence. In order to identify possible candidate genes, we performed sequence analysis and determined the tissue expression pattern analysis of 10 novel ESTs that are deleted in all SMS patients. We also present a detailed review of six promising candidate genes that map to the SMS critical region.
European Journal of Human Genetics (2001) **9,** 892–902.

Keywords: Smith-Magenis syndrome; transcription map; 17p11.2; chromosome 17; microdeletion syndrome; contig; continuous gene syndrome; haploinsufficiency

Introduction

Smith-Magenis syndrome (SMS) is a multiple congenital anomalies/mental retardation syndrome associated with an interstitial deletion of chromosome 17 involving band p11.2.[1–3] SMS patients have a recognisable physical phenotype that includes characteristic facies, brachycephaly, brachydactyly, hearing loss, myopia, and hoarse voice. Patients also display a distinctive behavioural profile that includes moderate mental retardation, sleep disturbances, self-injurious behaviour, and delayed speech and motor development.[1,4,5] The birth prevalence of SMS is estimated to be approximately 1:25,000,[1] although SMS is likely underdiagnosed due to the fact that it is a recently-described syndrome and its phenotype can be subtle.

The SMS common deletion interval is estimated to be 4–5 Mb[6] and is bordered by marker D17S58 (proximal) and anonymous cosmid cCI17-498 (distal).[7,8] Extensive analyses of rare, smaller deletions through fluorescence *in situ*

*Correspondence: SH Elsea, Department of Zoology, S-320 Plant Biology Building, Michigan State University, East Lansing, Michigan, MI 48824, USA. Tel: +517 353 5597; Fax: +517 432 7131; E-mail: elsea@msu.edu
Received 18 June 2001; revised 27 September 2001; accepted 28 September 2001

hybridisation (FISH) and somatic rodent : human hybrid cell lines that retain the deleted copy of chromosome 17 have been carried out in order to determine the smallest deletion necessary to give the full SMS phenotype.[7–11] The SMS 'critical interval' within 17p11.2 is approximately 1.5 – 2 Mb, and lies between marker D17S29 (centromeric) and cosmid cCI17-638 (telomeric).[7,8] Haploinsufficiency of one or more genes that lie within the critical interval is likely to be responsible for the SMS phenotype, although the true molecular basis of SMS is not yet known.[8,12]

Recent investigation into the mechanisms for chromosomal rearrangements indicates that genomic regions that are prone to duplications, deletions and inversions are often flanked by chromosome-specific low-copy repeats called duplicons.[13] The repeated regions that flank the SMS common deletion breakpoints within 17p11.2, termed the proximal and distal SMS-REPs (SMS-REPP and SMS-REPD, respectively), span <200 kb and contain a cluster of at least four genes/pseudogenes.[14] *De novo* deletions and duplications within the SMS region are likely due to interchromosomal recombination events involving the SMS-REPs.[14,15] A middle REP (SMS-REPM), that lies within the SMS critical interval, has not yet been shown to promote chromosomal rearrangement, though it is possible that SMS-REPM may be implicated in producing smaller deletions of 17p11.2.

Extensive work by our group and others has identified genes and marker loci that map to the SMS critical and common deletion regions within 17p11.2, in order to recognise potential candidate genes for SMS.[8,16–21] This region of 17p11.2 is very gene-rich. Within the critical deletion region, eleven genes have previously been reported: the human homologue of the *Drosophila flightless-I* gene (*FLII*); cytosolic serine hydroxymethyltransferase (*SHMT1*);[21,23] the human homologue of *Drosophila lethal 2 giant larva* (*LLGL1*);[18] topoisomerase IIIα (*TOP3A*);[19,24,25] elongation factor 1α (*EEF1A3*);[8] sterol regulatory element binding protein 1 (*SREBF1*);[8] myosin 15 (*MYO15A*);[26,27] phosphatidylethanolamine-N-methyltransferase 2 (*PEMT2*);[8,20] signalosome subunit 3 of the COP9 signalosome (*COP53, SGN3*);[28,29] developmentally-regulated GTP-binding protein 2 (*DRG2*);[30] and deoxyribonucleotidase-2 (*NT5M*, formerly *dNT-2*).[31] Many of these genes have been well-characterised and the effect of hemizygosity of several genes has been assessed in a variety of systems; yet there still remains no single excellent candidate gene for the SMS phenotype.[19,21,28] A breakpoint region for medulloblastomas (MB) has also been identified within 17p11.2, overlapping the SMS deletion region.[16,17] This chromosomal region also contains aberrantly methylated CpG islands in association with MBs, although no expression changes due to methylation were found in several ESTs and cDNAs.[32]

We now report a physical and transcription map of the SMS critical interval that links the proximal 17p11.2 region near the SMS-REPM and the distal region near D17S740. Not including the repetitive elements that flank the SMS critical interval, this map contains 13 known genes, 14 ESTs, and six anonymous genetic markers.

Materials and methods

Contig construction and analysis

BAC/PAC DNA preparation Human California Institute of Technology BAC libraries B-D (hCITB-D) or Roswell Park Cancer Institute library-11 (RPCI-11) BAC and RPCI-1 PAC clones were obtained (Research Genetics, Huntsville, AL, USA or BAC/PAC Resources, Oakland, CA, USA). DNA was isolated using a modification of the Qiagen very low copy protocol (Qiagen Inc., Valencia, CA, USA). Bacteria were grown overnight at 37°C on selective antibiotic (12.5 μg/ml chloramphenicol for BACs and 100 μg/ml kanamycin for PACs), then pelleted by spinning at $6000 \times g$ for 15 min at 4°C. Subsequently, 60 ml each of Qiagen Buffers P1, P2, and P3 were added to the pellet and mixed. Tubes were gently inverted, placed on ice for 30 min, and then centrifuged at 4°C at $\geqslant 20\,000 \times g$. Supernatant containing BAC or PAC DNA was promptly filtered, precipitated with 0.7 volumes of room temperature isopropanol, and centrifuged at $\geqslant 1\,5000 \times g$ for 30 min at 4°C. The pelleted DNA was resuspended in 500 μl sterile $1 \times$ TE, and 5 ml of QBT was added to the BAC/PAC DNA. The BAC/PAC DNA was applied to an equilibrated Qiagen Tip 500 and the Qiagen low copy protocol was then followed according to manufacturer's instructions. Purified BAC/PAC DNA was electrophoresed on a 1% TBE gel, stained with ethidium bromide, visualised on a transilluminator, and quantitated using a UV spectrophotometer (Ultrospec 2000, Pharmacia Biotech, Piscataway, NJ, USA). The SP6 and T7 ends of bc989B15, at the distal end of the critical interval (Figure 1), were sequenced by MWG Biotech (MWG Biotech, High Point, NC, USA).

Somatic cell hybrids The somatic cell hybrids utilised for regional mapping to human chromosome 17, including MH22-6, 88H5, LS-1, and 357-2D and the hamster and mouse parental cell lines a23 and Cl-1D (respectively), as well as hybrid cell lines that carry the deleted chromosome 17 from SMS patients, including 147-20D and 484-2D, have been described previously.[9] Hybrid 540-1D, which carries the deleted chromosome 17 from the SMS patient with the smallest reported deletion and hybrid 765-18D, which carries a chromosome 17 with a deletion overlapping the distal portion of the SMS region 17p, have been previously described.[8]

Southern analysis For Southern analysis, 6 μg of BAC or PAC DNA or 9 μg of human genomic or rodent-human hybrid DNA were digested overnight with 4 U/μg of *Eco*RI (Gibco, BRL Life Technologies, Rockville, MD, USA), with 2.5 mM spermidine added to genomic DNAs. All samples were electrophoresed in 1% Tris-acetate (TAE) agarose gels with buffer recirculation. Southern transfer and hybridisa-

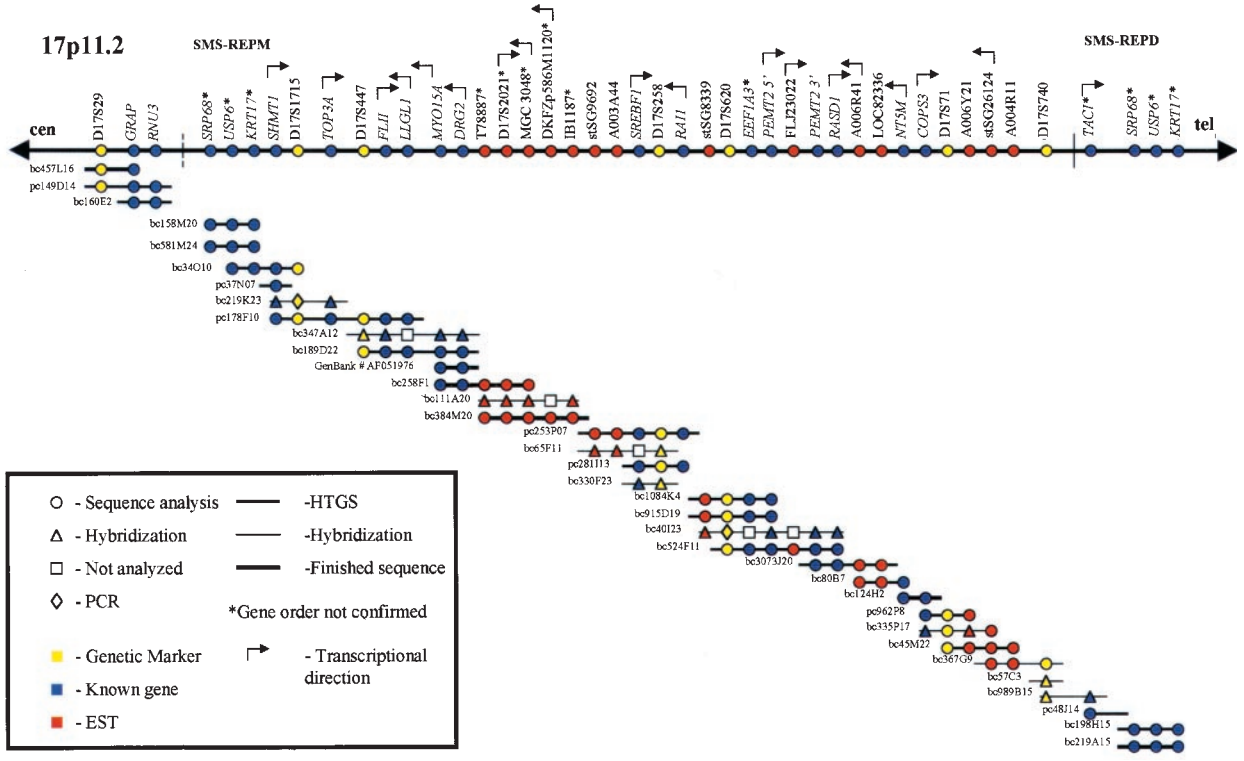

Figure 1 SMS critical interval transcription map. The SMS critical interval spans between the two vertical hatch marks. The dotted hatch marker at the proximal end indicates that we have not yet determined the particular BAC that spans the breakpoint for that end of the critical interval. The transcription map is comprised of an overlapping contig of BACs and PACs from the proximal SMS-REPM region to the distal genomic marker D17S740. Markers, including ESTs and known genes, were mapped to the contig through a combination of hybridization to *Eco*RI-digested BAC and PAC DNA, PCR, and database sequence analysis from the human genome project. All ESTs and genes were also hybridised to the somatic cell hybrid mapping panel (shown in Figure 2) to confirm HTGS sequence analysis and position within the SMS critical interval. BAC/PAC-end sequence overlaps confirmed the position of clones where markers are absent.

tions were performed essentially as described.[7] The blots were exposed to Hyperfilm (Amersham, Piscataway, NJ, USA), with one or two intensifying screens at −80°C.

Database searches Extensive information about the genes/ESTs in the region (Table 2) was obtained through the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov), including Unigene and the EST and STS databases. Draft assembly of the human high-throughput genome sequence (HTGS) was found at the Golden Path (http://genome.ucsc.edu/GoldenPath), and NCBI. Supplementary information was derived from the Genome Database (http://www.gdb.org). BLAST and BLAST +BEAUTY searches were carried out at the NCBI website (http://www.ncbi.nlm.nih.gov/BLAST) or through the Baylor Search Launcher (http://searchlauncher.bcm.tmc.edu). Proteomic analysis was carried out at the Expasy site (http://www.expasy.ch/) and through the Baylor Search Launcher. Conserved sequence domains were identified through NCBI (www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

EST characterisation and analysis

DNA preparation DNA was isolated from I.M.A.G.E. EST clones (Research Genetics, Huntsville, AL, USA, or the German Genome Project) using the QIAprep Miniprep Kit (Qiagen Inc., Valencia, CA, USA) according to manufacturer's instructions. Prepared DNA was electrophoresed and quantified as described above.

PCR The polymerase chain reaction (PCR) was performed in 25 μl volumes with primers listed in Table 1. PCR reactions were carried in an MJ Research thermocycler (MJ Research, Waltham, MA, USA) essentially as described.[8] EST clone inserts were amplified from vector primers (as indicated in the IMAGE database; http://image.llnl.gov), purified, and used as probes.

Northern analysis Human multiple tissue, foetal, and brain II Northern blots were purchased from Clontech and handled according to manufacturer's instructions. Prehybridisations and hybridisations were performed in 5 × SSPE (0.75 M NaCl,

0.05 M NaH_2PO_4, 5 mM EDTA), 10× Denhardt's solution (0.2% BSA, 0.2% Ficoll (Sigma, St. Louis, MO, USA), 0.2% Polyvinylpyrrolidone (Sigma, St. Louis, MO, USA)), 100 μg/ml herring sperm DNA, 2.0% SDS, 50% deionised formamide at 42°C. All probes were labelled with the Rediprime II kit (Amersham, Piscataway, NJ, USA), and unincorporated nucleotides were removed with a spermine precipitation. The blots were exposed to Hyperfilm (Amersham, Piscataway, NJ, USA) with one or two intensifying screens at −80°C for 2–5 days.

Results

Generation of the SMS critical interval

Figure 1 depicts the physical and transcription map of the SMS critical interval. The BACs and PACs within the contig were assembled through a combination of Southern hybridisation, PCR, and HTGS analysis. Initially, known genes, ESTs, and genomic markers that mapped to the critical interval[8] were used to screen a gridded hCITB BAC library in order to form a preliminary contig (data not shown). Overlapping BACs were identified and grouped according to *Alu*-PCR fingerprinting and restriction digests (data not shown). In the interest of space, many redundant BACs are not represented in the current contig (data not shown). In order to map markers to the initial contig, known genes and ESTs were hybridised to *Eco*RI-digested BACs or PCR-amplified from BAC DNA. This preliminary contig had many gaps and could not be completed with the BAC resources that were available.

We therefore took advantage of human genome project data to improve the coverage of our contig. To date, there remains little finished human genome sequence data for the SMS critical interval except for bc124H2 (which contains WI-11472/*NT5M* and a portion of *COPS3*) and sequence generated through *MYO15A* analysis (GenBank AF051976). This region of 17p is very gene rich but also contains numerous low-copy repeats that may complicate HTGS sequence alignment (Table 2 and Figure 1). We used BLAST searches with gene and EST sequence data to identify HTGS RPCI BAC/PAC and CIT BAC data that were pertinent to the SMS critical interval. We subsequently obtained several of these clones for analysis. We also utilised draft sequence alignments from the Golden Path (http://genome.ucsc.edu/GoldenPath). We searched with BAC-end sequence information from GenBank to establish a minimum tiling path of 16 BACs and two PACs (Figure 1). Several overlapping clones are included in the map to provide greater coverage (Figure 1). Gene order was determined through HTGS sequence analysis and Southern hybridisation for the presence or absence of a

Table 1 PCR primers

Primer name	*Forward sequence 5′-3′*	*Reverse sequence 5′-3′*	*Annealing temperature (°C)*
D17S71	CACTAAATAGCATCTATCCCCTTCC	CAAAGCCAAATCCTGCCCATG	52
D17S258	TCTAACCCCAAGCAAAGCCACTGC	GAGACCCTGGGATGATAATGGATGG	55
D17S447	GCTCTGGTTCATATTACCCTGACAC	GCTCTCCAGCCTGGGCAACAAGCGTG	55
D17S620	GGGAAGGTGTCTGAAACCCAAGG	CCCACCACTACCTATTGTTCTA	55
D17S740	CTCTTTGTGCTTGGCAGGGT	TACATTTAATGCAGGATGCC	55
D17S1715	GACTGGCAGGGGATAAGTACC	ATCTCTCAAGTGGGGGTCCT	55
D17S2021	TAAAATCAGAGTCGAGAGAAAGTCA	CACACTCAGCAGGATAGAGGC	55
A003A44	GTGACAAAAACACTAAACTTCTA	GGGACGGATATTTATTTC	55
A004R11	GTAAGAGTCAAAAATCTAACAGA	CACTTTATGCAGATTTCC	55
A006R41	CAGAGCATAAGTCCTTTTAA	GACTGTTATTTAACTGAATCAGT	55
A006Y21	TGGTGACTTCCAATCTATATAT	TTACTTTTGACTCACAGTGTTA	52
IB1187	GTATTTTGTACATGTCCATGGG	GATGAACTGCCTTAGCCCAA	52
NIB1041	ATCAACCTGGACATCAAATG	CTGGAAGACAAACAGGAACA	55
stSG8339	AGTGGGAAGGTGTTGAGGG	GTGCCCTGGCATATGAGTG	60
stSG9692	GAGACCAACAGGAACTATT	CTTTGAGAAACATTGAGG	52
T78887	AGACCTCAATAAGCTGTGGG	CCATGAGAATCAGTTTATGCC	55
WI-11472	TTACACACATGAACAGCATTTATTT	TGTGCTGGGGGACGATACT	55
WI-13499	ATCCTTTATTTCAGAAAGGAGAGC	TCCCCCTGTATTCCCCCT	52
COPS3[a]	GCCAGTGACAAGTGTTCGGAG	CCACGACAGACTTTAAAGTTTGCAA	55
FLII[b]	TGCGGTCCAAGGAACATGAGC	GAGTTGGATTTCCCAGACCCCTG	60
LLGLI[b]	AGCGGCATCGCTTCGTGCGT	TTAGCCTGGCTCAGCTTGGG	52
MYO15 Exon 66[c]	TGATGGTCTCCACCTATCTC	TCCTGAGAGGTTCAGTGATT	52
MYO15 Exon 2[c]	CACCCTCCAGCCTGAGGATCC	CTCCAAGGTTGCACAGGGCTCC	65
PEMT2 Exon 6/7[d]	CCCTCACCTACATAATGGCTCTCC	GCAGCTCAATCAGCTCCTCTTG	55
SHMT1[e]	CCTCTGTGGCGCCAGGTGGG	GCCACCCTGAAAGAGTTCAAGG	60
SREBF1[b]	ACATTGAGCTCCTCTCTTGAAG	TAGCCTAACACAGGGGTGGA	55
TOP3A[b]	CACAGTGGAGATCGACAT	TCCATGAGGGCAATGAGG	55

Primers pairs for ESTs and genomic markers were derived from the Genome Database (http://www.gdb.org). Primers pairs used for mapping known genes are indicated. [a]Elsea *et al*[22]; [b]Generated for this investigation; [c]T Friedman, personal communication; [d]D Vance, personal communication; [e]Elsea *et al*[21].

Table 2 ESTs within the SMS critical interval

EST amplimer	*Clone IMAGE*	*GenBank accession*	*Unigene*	*Tissue expression*	*Sequence motif (as of 9/19/01)*	*Best BLAST or Unigene hit (as of 9/19/01)*
IB1187	IB1187	T15437	–	Smear on Northern blot	526 bp MIR/MER21 repeat	Novel
stSG8339	204716	H57290	–	No transcript on Northern blot	93 bp MIR repeat	Novel
WI-13499	29328	R41366	Hs. 78582	2.2 kb transcript in all adult and foetal tissues[a]	94% aa alignment to GTP-binding protein motif	Human *DRG2* (GenBank X80754)
A003A44	683788	AA236905	Hs. 8125	6.5 kb transcript in all adult tissues	None	Similar to target of myb1 homologue-like 2 (*TOM1L2*) (GenBank AJ230803)
D17S2021	712371	AA281720	Hs. 13734	1.8 kb transcript in all adult and foetal tissues; highest in heart	None	28% similar over 225 aa to yeast ATP12 protein (GenPept P22135)
DKFZp434A139Q2	–	AL133649	Hs. 278684	8.0 kb transcript in all adult tissues	*RAI1* contains CAG repeat	Human *RAI1* (GenBank NM_030665)
stSG9692	23808	T77503	Hs. 12537	6.5 kb transcript in all adult tissues	None	36% similar to human *SIM2*
T78887	1258050	AA730163	Hs. 187422	No transcript in Northern blot, isolated from germ cell library	378 bp LINE1 repeat	Novel
A004R11	118648	T91728	Hs. 16899	2.4 kb transcript in adult and foetal liver	None	Novel
A006R41	2177116	AI500641	Hs. 235195	6.0 kb transcript in adult skeletal muscle and foetal brain	None	Hypothetical protein FLJ10193 (GenBank AK001055)
A006Y21	187071	H50830	Hs. 31652	Smear on Northern blot	322 bp Alu repeat; 85.7% alignment to phospholipase D active site motif	51% similar to *Y. pestis* endonuclease
WI-11472	156347	R72633	Hs. 16614	1.6 kb transcript in adult heart,	14.8% alignment to hydrolase domain	Human *NT5M* (GenBank AJ277557)
NIB1041	237257	T16275	Hs. 106359	5 kb transcript in all adult tissues[c]	*RASD1* contains 98% alignment to Ras subfamily	Human *RASD1* (GenBank AF262018)
stSG26124	2223382	AI570799	Hs. 84883	~6.0 kb transcript in all adult and foetal tissues	KIAA0864 is 91% similar (602 aa alignment) to mouse p116Rip protein (GenBank T30868)	KIAA0864 protein (GenBank AB020671)

EST name, clone ID, GenBank accession number, Unigene number, clone insert size, tissue expression pattern, and best BLAST hit (as of 9/19/01) are indicated in separate columns. All ESTs map to the SMS critical interval. [a]Results reported in Vlangos *et al*[30]; [b]Northern results similar to those reported in Rampazzo *et al*[31]; [c]Northern results similar to those reported in Tu and Wu[53].

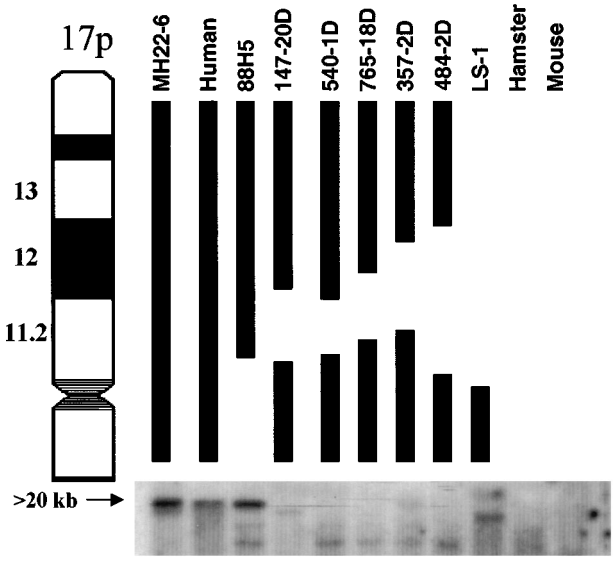

Figure 2 Localisation of *RAI1* to 17p11.2. In order to fine-map *RAI1* to a specific region within 17p11.2, the ~2.0 kb insert from EST DKFZp434A139Q2 was hybridised to a panel of *Eco*RI-digested somatic cell hybrids, carrying deleted chromosomes 17 from SMS patients (147-20D, 540-1D and 484-2D) and non-SMS patients with overlapping deletions along 17p11.2 (765-18D and 357-2D), as well as chromsome 17p positive controls (MH22-6, human genomic, and 88H5), and iso-17q (LS-1), hamster (a23), and mouse (Cl-1D) cell line negative controls. A $<$20 kb *RAI1* band is evident only in the MH22-6, human, 88H5, and 357-2D lanes (present but faint on this exposure), demonstrating that *RAI1* is deleted in all SMS patients and maps to the central portion of the SMS critical interval. Further locatisation to PACs within the transcription map (Figure 1) indicates that this gene is located between D17S258 and D17S620.

certain gene or EST in *Eco*RI-digested BAC or PAC DNA. There remain several portions of the contig where gene order could not be determined absolutely, as indicated (Figure 1), due to the unfinished draft sequence alignments. Transcription orientation was obtained from the Golden Path database, cosmid sequencing (data not shown), and known overlapping segments, such as the 3′ ends of *FLII* and *LLGL1*.[33] Cosmid genomic clones for several of the genes were identified through hybridisation of EST inserts to gridded filters of the Los Alamos flow-sorted chromosome 17-specific cosmid library.[34] The region from *SHMT1* to *FLII* is well-represented by overlapping cosmids, which confirmed gene order (data not shown).[19,21,22] We reconciled HTGS data with naturally-occurring deletions by hybridising genes and ESTs to *Eco*RI-digested rodent:human somatic cell hybrid panel DNA (Figure 2).[8] The SMS critical interval is represented by the SMS patient HOU142-540,[7,8,10] who carries the smallest deletion described to date.

Within the critical interval, we have identified 13 known genes and 14 ESTs. A summary is presented below. Six genomic markers were utilised in the generation of the physical map of the critical region: D17S258 (STS, GDB:178583), D17S740 (CA dinucleotide repeat, GenBank L29354, dbSTS: 3906), D17S1794 (CA dinucleotide repeat, GenBank Z52295, dbSTS: 46953), D17S620 (CA dinucleotide repeat, GenBank L29360, dbSTS: 3170), D17S447 (STS; also called FG2, GDB:186807), and D17S71 (STS, GDB:180951).[7,8]

EST characterisation and analysis

Previously, numerous markers were mapped along chromosome 17p11-p12, including short tandem repeats (STRs), sequence-tagged sites (STSs), expressed sequence tags (ESTs), and 12 identified genes.[8] Binning/fine-mapping of ESTs within 17p11.2 was accomplished through PCR, FISH, or Southern hybridisation to DNA from a panel of rodent : human somatic cell hybrids.[8] In this study, plasmid clones for each of the 14 ESTs that mapped to the critical interval were obtained commercially, inserts were sequenced, and extensive database searches of the nucleotide sequence and 6-frame amino acid translation were conducted in order to identify homologous genes and sequence motifs (Table 2). We determined the tissue expression pattern and transcript size for each EST by Northern analysis of adult human and foetal mRNAs. Figure 3 shows the multiple tissue northern (MTN) blot for EST DKFZp434A139Q2, which represents the 3′ end of the human *RAI1* gene. This clone was obtained from the German Genome Project.[35] An 8.0 kb transcript is seen in all adult tissues.

Table 2 shows the results of our analysis of ESTs that map to the SMS critical interval. Four ESTs, WI-13499, DKFZp434A139Q2, WI-11472, and NIB1041, represent known genes and are described below. Several of the ESTs that we sequenced, including IB1187, stSG8339, D17S2021, T78887, A003A44, A004R11, and A006R41 were novel, or only matched other unidentified clones in the database (Table 2). A006R41 is of interest, as this gene is expressed in foetal brain and may play a role in development. A004R11 may possibly represent a metabolic gene, as it is expressed in only adult and foetal liver.

We recently mapped marker stSG26124 through sequence analysis and carried out an initial tissue expression analysis (Table 2). The Unigene database lists stSG26124 as representative of the cDNA for the KIAA0864 protein (GenBank AB020671; Unigene Hs.84883). We obtained a clone for KIAA0864 from the Kazusa Research Institute (Database of Human Unidentified Gene-Encoded Large Protein (HUGE) database; http://www.kazusa.or.jp/huge/gfpage/KIAA0864). The cellular role of this protein is unknown, though preliminary sequence analysis suggests that it may be involved in control of the actin cytoskeleton (Table 2).

Three ESTs, IB1187, stSG8339 (represented by a single EST), and T78887 did not show transcripts on Northern blot, possibly due to low transcript abundance or expression in tissues not represented on the Clontech MTN and foetal blots (Table 2).

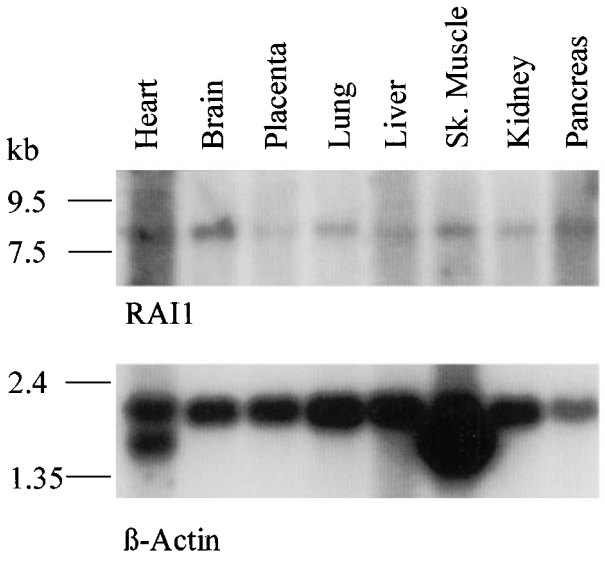

Figure 3 *RAI1* Northern analysis. The ~2.0 kb insert from EST DKFZp434A139Q2, representing the 3′ end of human *RAI1*, was hybridised to a Clontech MTN Northern blot. An ~8.0 kb transcript is evident in all adult tissues. β-actin is included as a mRNA loading control.

HTGS sequence analysis of several of the BACs and PACs within the contig revealed the presence of four other putative genes or expressed transcripts that we have not yet fully analysed, FLJ23022 (GenBank NM_025051, Unigene Hs.2877170), LOC82336 (GenBank XM_012549), DKFZp586M1120 (GenBank NM_031294, Unigene Hs.159068), and MGC 3048 (GenBank NM_024052; Unigene Hs.115437) (Figure 1). Most of these are novel or hypothetical proteins and further investigation is needed to determine if any of these sequences are actively expressed and their potential contribution to the SMS phenotype.

Summary and analysis of candidate genes

Several of the known genes that map to the critical interval including, *MYO15A*,[27,36] *TOP3A*,[19,24,25] *SREBF1*,[8] *PEMT2*,[8,20] and *COPS3*,[28,29] have been well-characterised and we do not consider them a high priority for future work. In the case of *MYO15A*, this gene is involved in hereditary recessive deafness (*DFNB3*), and may not be involved in the deafness seen in several SMS patients.[27,36] *PEMT2*, *SREBF1*, and *TOP3A* have been analysed in knockout mice; heterozygous mice appear normal and healthy[37,38] (D Vance and J Wang, personal communications). The COPS3 protein complex has been analysed in SMS patient and control parent lymphoblastoid cell lines, and no noticeable abnormalities were detected in SMS patient cell lines.[28] SHMT1 has also been analysed in patient lymphoblastoid cell lines, and while haploinsufficiency of SHMT1 was detected, serum folate, glycine, and serine levels were normal.[21] Studies of SHMT1 and related metabolites in other types of patient cell lines or cerebrospinal fluid may be more conclusive. *EEF1A3* is also not considered to be a high priority, as several genes/pseudogenes for this elongation factor are known and haploinsufficiency of *EEF1A3* may have little physiological effect.[8]

There are several known genes within the critical interval that have not been extensively analysed that we consider to be promising SMS candidate genes. These six genes are summarised below:

FLII The *Drosophila melanogaster flightless-I* (*fliI*) protein is a member of the gelsolin family of proteins, which play a role in actin binding and severing.[39] FLII, the human homologue of *fliI* is 58% identical to the *Drosophila* protein and contains a leucine-rich repeat (LRR) N-terminal domain, which is believed to play a role in cell adhesion and protein–protein interaction.[40] In *Drosophila*, less severe mutations in the *FLII* locus lead to abnormal myofibrillar arrangements in the indirect flight muscles which result in flightlessness; more debilitating mutations are embryonic lethal.[40] FLII (GenBank NM_002018, Unigene Hs. 83849, OMIM 600362) has the ability to bind to actin *in vitro* and is highly expressed in human muscle, though little else is known about its cellular role.[33] An interacting protein, termed FLAP (FLI-LRR associated protein) has been identified, which may link FLII to the actin cytoskeleton.[33,41] Genomic clones for human *FLII* demonstrate that the 3′ end of this gene overlaps with the 3′ end of the *LLGL1* gene.[33] The murine homologue of *FLII*, *Fliih* has recently been identified and mapped to mouse chromosome 11B.[42] The 3′ overlap region between *Fliih* and *Llglh* is conserved in the mouse.[42] Human *FLII* was mapped to the SMS critical interval by FISH and somatic cell hybrids and remains a potential candidate gene.[22]

LLGL1 *LLGL1* (GenBank NM_004140, Unigene Hs. 95659, OMIM 600966) is the human homologue of the *Drosophila* tumour suppressor gene, *D-lethal (2) giant larvae*. When disrupted, *D-lgl* can produce tumours in imaginal disks and abnormal transformation of the adult optic centers in larval brains.[43] Human *LLGL1* is expressed in brain, kidney, and muscle, and shows an association with the cytoskeleton.[43] Antibodies against *LLGL1* were able to co-immunoprecipitate *LLGL1* and non-muscle mysosin heavy chain; this interaction is conserved between human and *Drosophila*.[43] Both *Drosophila* and human *LLGL1* are also associated with a serine kinase, which is able to specifically recognise and phosphorylate serine residues within the protein.[43] *LLGL1* was mapped to chromosome 17p11.2 by FISH and Southern blotting[43] and was demonstrated to overlap at the 3′ end with the *FLII* gene.[33] The contribution of this gene to the SMS phenotype is unknown.

DRG2 Directly adjacent to *MYO15A* is developmentally-regulated GTP-binding protein 2 (*DRG2*) (GenBank

NM_001388, Unigene Hs.78582, OMIM 602986). The *DRG2* protein is a member of a novel family of small GTP-binding proteins that are highly conserved and may play a role in oncogenesis.[44] *DRG2* is highly homologous to *DRG1*, a developmentally regulated gene that is expressed in embryonic and neoplastic tissue and down-regulated in adult tissues.[45,46] Northern analysis of *DRG2* shows low levels of expression in all adult and foetal tissues examined.[30,44] This preliminary evidence suggests that *DRG2* is not developmentally regulated, but further studies are required. The role of *DRG2* in SMS, if any, is unknown.

RAI1 In this study, we have mapped *RAI1* (GenBank NM_030665, Unigene Hs.278684) by HTGS analysis to the central portion of the SMS critical interval. We confirmed our sequence analysis by Southern hybridisation of the 3′ end of the human *RAI1* gene to the somatic cell hybrid mapping panel (Figure 2). This result demonstrated that *RAI1* is deleted in all SMS patients. *RAI1* (*retinoic acid-induced 1*) is the human homolog of the mouse GT1 gene, which is up-regulated in mouse embryonal carcinoma cells following retinoic acid treatment to induce neuronal differentiation.[47] *In situ* hybridisation demonstrated that GT1 protein expression in mouse brain is localised to neurons.[47] A recent study revealed that the human *RAI1* protein, which contains a CAG repeat, may modify the age of onset of spinocerebellar ataxia type 2 (SCA2) possibly through a protein-protein interaction involving the polyglutamine regions of the ataxin-2 and RAI1 proteins.[48] As *GT1* shows ubiquitous expression in brain and may influence neuronal differentiation, [47] additional studies are required to assess the role of *RAI1* in the SMS phenotype. Northern analysis indicates that human *RAI1* is expressed as an 8.0 kb transcript in all adult tissues (Table 2 and Figure 3). As this manuscript was under review, Seranski *et al*.[49] published the genomic structure of *RAI1* and determined the expression pattern of *RAI1* by Northern analysis; these results are in agreement with Figure 3.

RASD1 (formerly AGS1/DEXRAS1) We previously localised the EST marker NIB1041 to the SMS critical interval by PCR to somatic cell hybrids.[8] Recent sequence analysis in our laboratory has determined that the ~1.5 kb NIB1041 insert is a 99% match to 3′ end of cDNA sequence for activator of G-signalling 1 (*AGS1*) (GenBank AF222979, Unigene Hs.25829) as well as the cDNA sequence for dexamethasone-induced ras-related protein (*DEXRAS1*) (GenBank AF262018, Unigene Hs.25829). Unigene has grouped *AGS1* and *DEXRAS1* together in the same entry, indicating that these sequences represent the same gene, now formally designated *RASD1*. *AGS1* (activator of G-protein signalling), was isolated through a *Saccharomyces cerevisiae* screen to identify mammalian nonreceptor activators of G-protein signaling pathways.[50] *AGS1*, a ras-related protein, is able to process guanosine triphosphate exchange of the heterotrimeric G protein G_{α} subunit.[50,51] *DexRas1* was discovered independently in mice through differential display analysis to identify genes that were induced by glucocorticoid hormone (dexamethasone) treatment in AtT-20 cells.[52] Murine *Dex-Ras1* is expressed in heart, brain, and liver, and *DEXRAS1* mRNA levels in these tissues significantly increased after glucocorticoid treatment.[52] Human *DEXRAS1* cDNA was isolated through a yeast two-hybrid screen and subsequently cloned.[53] The mouse and human DEXRAS1 protein are 98% identical at the amino acid level and both are members of the Ras superfamily of GTPases.[53] Expression of human *DEXRAS1* was also shown to be strongly stimulated by dexamethasone, in a variety of tissues (Table 2, in agreement with our NIB1041 data).[53] Recent investigation into the role of rat Dexras1 in signalling demonstrates a ternary complex formation between neuronal nitric acid synthase (nNOS), the nNOS adaptor protein CAPON and Dexras1 to enhance Dexras1 GTP/GDP exchange.[54] The downstream signalling targets of Dexras1 are unknown.[54] Human *DEXRAS1* may also play a role in cell adhesion and extracellular matrix interactions,[53] though the contribution to the SMS phenotype is unknown.

NT5M (dNT-2) Our BLAST searches of the ~2.1 kb insert of EST WI-11472 revealed that WI-11472 is a 99% match to the cDNA for deoxyribonucleotidase-2 (*dNT-2/NT5M*) (GenBank NM_020201, Unigene Hs.16614), which codes for a mitochondrial deoxyribonucleotidase.[31] Independently, Rampazzo *et al*.[31] mapped NT5M to the SMS region based on sequence analysis of a genomic clone (GenBank AC006071) and the proximity of *NT5M* to *COPS3*. Deoxyribonucleotidases hydrolyze the monophosphate ester linkages of the 5′ or 3′ carbon of deoxyribonucleotides, producing an inorganic phosphate and the corresponding deoxyribonucleoside.[55] Northern analysis of *NT5M* revealed high expression in the heart, brain, and skeletal muscle, a 'typical' pattern for a mitochondrial enzyme.[31] These data are in agreement with our Northern analysis of EST WI-11472 (Table 2). The authors demonstrated that NT5M is targetted to the mitochondria by GFP-fusion localisation and by *in vitro* incubation of the recombinant protein with mitochondria.[31] NT5M activity is limited to the dephosphorylation of thymine and uracil deoxyribonucleotides and may play an important role in modulation of dTTP substrate pools and removal of excess dTTP from the mitochondria.[31] A potential role in SMS is yet to be identified.

Discussion

In order to identify possible candidates for SMS and to characterise the genomic content of distal 17p11.2, we sought to create the first contiguous physical and transcription map of the SMS critical region. Given the increased stability in the case of manipulation of BACs and PACs compared to yeast artificial chromosomes (YACs), we chose

BACs and PACs to complete the contig of the SMS critical interval. As well, given that the human genome sequence was being determined using BAC and PAC library clones, our mapping information will be invaluable in validating the draft sequence. We also relied on previously-published maps of the region,[14,17] including a partial YAC contig, to confirm the placement of markers, although many discrepancies existed between various maps for the localisation of certain genes and STSs. An example of conflicting data is the mapping of marker D17S258 in the YAC contig and in our current map.[14] We localised this marker, by sequence analysis and by hybridisation, adjacent to *SREBF1*, within the SMS critical interval. Chen *et al*.[14] localised this marker in a more centromeric position, adjacent to D17S29. Furthermore, an earlier, preliminary transcription map of chromosome 17p11.2, generated in order to identify regions of chromosomal instability contains many gaps that we have subsequently filled.[17] Our current contig also differs somewhat from the published human draft sequence by the Human Genome Sequencing Consortium.[56] The latter map separates the overlapping genes *FLII* and *LLGL1*[33] and places several genes between them, including *ALDH3*, *ALHD10*, and *MFAP4*, which have been previously localised to a more proximal region of 17p11.2.[8] In our transcription map, we rearranged several markers and confirmed their order from the telomeric to centromeric regions of 17p11.2 (Figure 1).

In order to identify possible SMS candidate genes, we are especially interested in genes that are developmentally-regulated and those that may play a role in mental and behavioural development. In addition, genes that are expressed in the neural crest could affect craniofacial and heart development.[57–59] We are also very interested in transcription factors, as they may have a more global effect on development and are often dosage-sensitive.[60,61] The T-box family transcription factor *TBX1* was recently identified as the gene responsible for the cardiac defects in DiGeorge syndrome, a disorder resulting from a microdeletion involving 22q11.[60–62]

As discussed in depth above, genes that are of high priority for further study are *RAI1*, *FLII*, *LLGL1*, *DRG2*, *RASD1*, and *NT5M*. *NT5M* is an especially intriguing SMS candidate gene because of the possible role of the NT5M protein on the modulation of dTTP substrate pools in the mitochondria.[31] We speculate that some of the characteristics of SMS, including hypotonia, mental retardation, and behavioural abnormalities, may be an effect of excess dTTP and defective mitochondria. Furthermore, modulation of nucleotide metabolism has been shown to be important in disorders with similar physical and behavioural features to SMS, such as Lesch-Nyhan syndrome.[63,64] We also consider the nine novel ESTs to be possible candidate genes, especially EST A006R41, as it is expressed in foetal brain (Table 2).

In this study, we have generated the first overlapping and contiguous transcription map of the entire SMS critical interval, linking the proximal 17p11.2 region near the SMS-REPM and the distal region near D17S740 in a minimum tiling path of 16 BACs and two PACs. As the presence of low-copy repeats and duplicons that flank this region can make sequence alignment difficult, and some of the information from the human genome project was inconsistent with naturally-occurring deletions described in patients,[7,8] we combined sequence analysis with molecular biological techniques such as PCR and Southern analysis to give our map greater coverage. Concurrently, we were able to identify six promising SMS candidate genes and numerous ESTs that warrant further investigation. We intend to pursue mouse knockout models and in depth *in vitro* characterisation of these genes and ESTs, including analysis in SMS patient and parental control cell lines.

Acknowledgments

The authors thank Joy Gumin, for technical assistance and cell culture work; and Raymond Schoener-Scott, Matt Stratton, Nakia Smith, Megan Loomer, Sarah Stapels, Nicole Urban, Tiffany Newton, Greg Stromberg, Anita Allman, and Ruth Ann Adell for technical assistance. The authors also thank Drs Patrick Venta and Karen Friderici for critical reading of the manuscript. The clone for TACI was a generous gift of Dr Richard J Braun at the Mayo Clinic. The work was funded in part by NIH grants HD08010 and HD38534 to SHE and HD28458 to PIP. Gene order and sequence information are discussed with respect to the status of the NCBI databases on September 19, 2001. Any changes to these databases since that time are not addressed in this manuscript.

References

1 Greenberg F, Guzzetta V, Montes de Oca-Luna R *et al*: Molecular analysis of the Smith-Magenis syndrome: a possible contiguous-gene syndrome associated with del(17)(p11.2). *Am J Hum Genet* 1991; 49: 1207–1218.
2 Smith AC, McGavran L, Robinson J *et al*: Interstitial deletion of (17)(p11.2p11.2) in nine patients. *Am J Med Genet* 1986; **24**: 393–414.
3 Stratton RF, Dobyns WB, Greenberg F *et al*: Interstitial deletion of (17)(p11.2p11.2): report of six additional patients with a new chromosome deletion syndrome. *Am J Med Genet* 1986; **24**: 421–432.
4 Greenberg F, Lewis RA, Potocki L *et al*: Multi-disciplinary clinical study of Smith-Magenis syndrome (deletion 17p11.2). *Am J Med Genet* 1996; **62**: 247–254.
5 Colley AF, Leversha MA, Voullaire LE, Rogers JG: Five cases demonstrating the distinctive behavioural features of chromosome deletion 17(p11.2 p11.2) (Smith-Magenis syndrome). *J Paediatr Child Health* 1990; **26**: 17–21.
6 Trask BJ, Mefford H, van den Engh G *et al*: Quantification by flow cytometry of chromosome-17 deletions in Smith-Magenis syndrome patients. *Hum Genet* 1996; **98**: 710–718.
7 Juyal RC, Figuera LE, Hauge X *et al*: Molecular analyses of 17p11.2 deletions in 62 Smith-Magenis syndrome patients. *Am J Hum Genet* 1996; **58**: 998–1007.
8 Elsea SH, Purandare SM, Adell RA *et al*: Definition of the critical interval for Smith-Magenis syndrome. *Cytogenet Cell Genet* 1997; **79**: 276–281.
9 Guzzetta V, Franco B, Trask BJ *et al*.: Somatic cell hybrids, sequence-tagged sites, simple repeat polymorphisms, and yeast artificial chromosomes for physical and genetic mapping of proximal 17p. *Genomics* 1992; **13**: 551–559.

10 Juyal RC, Finucane B, Shaffer LG *et al*: Apparent mosaicism for del(17)(p11.2) ruled out by fluorescence in situ hybridisation in a Smith-Magenis syndrome patient. *Am J Med Genet* 1995; **59**: 406–407.
11 Juyal RC, Greenberg F, Mengden GA *et al*: Smith-Magenis syndrome deletion: a case with equivocal cytogenetic findings resolved by fluorescence in situ hybridization. *Am J Med Genet* 1995; **58**: 286–291.
12 Schmickel RD: Contiguous gene syndromes: a component of recognizable syndromes. *J Pediatr* 1986; **109**: 231–241.
13 Ji Y, Eichler EE, Schwartz S, Nicholls RD: Structure of chromosomal duplicons and their role in mediating human genomic disorders. *Genome Res* 2000; **10**: 597–610.
14 Chen KS, Manian P, Koeuth T *et al*.: Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nat Genet* 1997; **17**: 154–163.
15 Potocki L, Chen KS, Park SS *et al*: Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet* 2000; **24**: 84–87.
16 Scheurlen WG, Seranski P, Mincheva A *et al*: High-resolution deletion mapping of chromosome arm 17p in childhood primitive neuroectodermal tumors reveals a common chromosomal disruption within the Smith-Magenis region, an unstable region in chromosome band 17p11.2. *Genes Chromosomes Cancer* 1997; **18**: 50–58.
17 Seranski P, Heiss NS, Dhorne-Pollet S *et al*: Transcription mapping in a medulloblastoma breakpoint interval and Smith-Magenis syndrome candidate region: identification of 53 transcriptional units and new candidate genes. *Genomics* 1999; **56**: 1–11.
18 Koyama K, Fukushima Y, Inazawa J, Tomotsune D, Takahashi N, Nakamura Y: The human homologue of the murine Llglh gene (LLGL) maps within the Smith-Magenis syndrome region in 17p11.2. *Cytogenet Cell Genet* 1996; **72**: 78–82.
19 Elsea SH, Fritz E, Schoener-Scott R, Meyn MS, Patel PI: Gene for topoisomerase III maps within the Smith-Magenis syndrome critical region: analysis of cell-cycle distribution and radiation sensitivity. *Am J Med Genet* 1998; **75**: 104–108.
20 Walkey CJ, Shields DJ, Vance DE: Identification of three novel cDNAs for human phosphatidylethanolamine N-methyltransferase and localization of the human gene on chromosome 17p11.2. *Biochim Biophys Acta* 1999; **1436**: 405–412.
21 Elsea SH, Juyal RC, Jiralerspong S *et al*: Haploinsufficiency of cytosolic serine hydroxymethyltransferase in the Smith-Magenis syndrome. *Am J Hum Genet* 1995; **57**: 1342–1350.
22 Chen KS, Gunaratne PH, Hoheisel JD *et al*: The human homologue of the Drosophila melanogaster flightless-I gene (flil) maps within the Smith-Magenis microdeletion critical region in 17p11.2. *Am J Hum Genet* 1995; **56**: 175–182.
23 Garrow TA, Brenner AA, Whitehead VM *et al*: Cloning of human cDNAs encoding mitochondrial and cytosolic serine hydroxymethyltransferases and chromosomal localization. *J Biol Chem* 1993; **268**: 11910–11916.
24 Hanai R, Caron PR, Wang JC: Human TOP3: a single-copy gene encoding DNA topoisomerase III. *Proc Natl Acad Sci USA* 1996; **93**: 3653–3657.
25 Fritz E, Elsea SH, Patel PI, Meyn MS: Overexpression of a truncated human topoisomerase III partially corrects multiple aspects of the ataxia-telangiectasia phenotype. *Proc Natl Acad Sci USA* 1997; **94**: 4538–4542.
26 Liang Y, Wang A, Probst FJ *et al*: Genetic mapping refines DFNB3 to 17p11.2, suggests multiple alleles of DFNB3, and supports homology to the mouse model shaker-2. *Am J Hum Genet* 1998; **62**: 904–915.
27 Wang A, Liang Y, Fridell RA *et al*: Association of unconventional myosin MYO15 mutations with human nonsyndromic deafness DFNB3. *Science* 1998; **280**: 1447–1451.
28 Elsea SH, Mykytyn K, Ferrell K *et al*: Hemizygosity for the COP9 signalosome subunit gene, SGN3, in the Smith- Magenis syndrome. *Am J Med Genet* 1999; **87**: 342–348.
29 Potocki L, Chen KS, Lupski JR: Subunit 3 of the COP9 signal transduction complex is conserved from plants to humans and maps within the Smith-Magenis syndrome critical region in 17p11.2. *Genomics* 1999; **57**: 180–182.
30 Vlangos CN, Das P, Patel PI, Elsea SH: Assignment of developmentally regulated GTP-binding protein (DRG2) to human chromosome band 17p11.2 with somatic cell hybrids and localization to the Smith-Magenis syndrome critical interval. *Cytogenet Cell Genet* 2000; **88**: 283–285.
31 Rampazzo C, Gallinaro L, Milanesi E, Frigimelica E, Reichard P, Bianchi V: A eoxyribonucleotidase in mitochondria: involvement in regulation of dNTP pools and possible link to genetic disease. *Proc Natl Acad Sci USA* 2000; **97**: 8239–8244.
32 Fruhwald MC, O'Dorisio MS, Dai Z *et al*: Aberrant hypermethylation of the major breakpoint cluster region in 17p11.2 in medulloblastomas but not supratentorial PNETs. *Genes Chromosomes Cancer* 2001; **30**: 38–47.
33 Campbell HD, Fountain S, Young IG *et al*: Genomic structure, evolution, and expression of human FLII, a gelsolin and leucine-rich-repeat family member: overlap with LLGL. *Genomics* 1997; **42**: 46–54.
34 Kallioniemi OP, Kallioniemi A, Mascio L *et al*: Physical mapping of chromosome 17 cosmids by fluorescence in situ hybridization and digital image analysis. *Genomics* 1994; **20**: 125–128.
35 Vente A, Korn B, Zehetner G, Poustka A, Lehrach H: Distribution and early development of microarray technology in Europe. *Nat Genet* 1999; **22**: 22.
36 Liang Y, Wang A, Belyantseva IA *et al*: Characterization of the human and mouse unconventional myosin XV genes responsible for hereditary deafness DFNB3 and shaker 2. *Genomics* 1999; **61**: 243–258.
37 Li W, Wang JC: Mammalian DNA topoisomerase III(is essential in early embryogenesis. *Proc Natl Acad Sci USA* 1998; **95**: 1010–1013.
38 Shimano H, Shimomura I, Hammer RE *et al*: Elevated levels of SREBP-2 and cholesterol synthesis in livers of mice homozygous for a targeted disruption of the SREBP-1 gene. *J Clin Invest* 1997; **100**: 2115–2124.
39 Sun HQ, Yamamoto M, Mejillano M, Yin HL: Gelsolin, a multifunctional actin regulatory protein. *J Biol Chem* 1999; **274**: 33179–33182.
40 Campbell HD, Schimansky T, Claudianos C *et al*: The Drosophila melanogaster flightless-I gene involved in gastrulation and muscle degeneration encodes gelsolin-like and leucine-rich repeat domains and is conserved in Caenorhabditis elegans and humans. *Proc Natl Acad Sci USA* 1993; **90**: 11386–11390.
41 Liu YT, Yin HL: Identification of the binding partners for flightless I, A novel protein bridging the leucine-rich repeat and the gelsolin superfamilies. *J Biol Chem* 1998; **273**: 7920–7927.
42 Campbell HD, Fountain S, Young IG, Weitz S, Lichter P, Hoheisel JD: Fliih, the murine homologue of the Drosophila melanogaster flightless I gene: nucleotide sequence, chromosomal mapping and overlap with Llglh. *DNA Seq* 2000; **11**: 29–40.
43 Strand D, Unger S, Corvi R *et al*: A human homologue of the Drosophila tumour suppressor gene l(2)gl maps to 17p11.2-12 and codes for a cytoskeletal protein that associates with nonmuscle myosin II heavy chain. *Oncogene* 1995; **11**: 291–301.
44 Schenker T, Lach C, Kessler B, Calderara S, Trueb B: A novel GTP-binding protein which is selectively repressed in SV40 transformed fibroblasts. *J Biol Chem* 1994; **269**: 25447–25453.
45 Sazuka T, Kinoshita M, Tomooka Y, Ikawa Y, Noda M, Kumar S: Expression of DRG during murine embryonic development. *Biochem Biophys Res Commun* 1992; **189**: 371–377.
46 Sazuka T, Tomooka Y, Ikawa Y, Noda M, Kumar S: DRG: a novel developmentally regulated GTP-binding protein. *Biochem Biophys Res Commun* 1992; **189**: 363–370.

47 Imai Y, Suzuki Y, Matsui T, Tohyama M, Wanaka A, Takagi T: Cloning of a retinoic acid-induced gene, GT1, in the embryonal carcinoma cell line P19: neuron-specific expression in the mouse brain. *Brain Res Mol Brain Res* 1995; **31**: 1 – 9.
48 Hayes S, Turecki G, Brisebois K *et al*: CAG repeat length in RAI1 is associated with age at onset variability in spinocerebellar ataxia type 2 (SCA2). *Hum Mol Genet* 2000; **9**: 1753 – 1758.
39 Seranski P, Hoff C, Radelof U *et al*: RAI1 is a novel polyglutamine encoding gene that is deleted in Smith-Magenis syndrome patients. *Gene* 2001; **270**: 69 – 76.
50 Cismowski MJ, Takesono A, Ma C *et al*: Genetic screens in yeast to identify mammalian nonreceptor modulators of G-protein signaling. *Nat Biotechnol* 1999; **17**: 878 – 883.
51 Cismowski MJ, Ma C, Ribas C *et al*: Activation of heterotrimeric G-protein signaling by a ras-related protein. Implications for signal integration. *J Biol Chem* 2000; **275**: 23421 – 23424.
52 Kemppainen RJ, Behrend EN: Dexamethasone rapidly induces a novel ras superfamily member-related gene in AtT-20 cells. *J Biol Chem* 1998; **273**: 3129 – 3131.
53 Tu Y, Wu C: Cloning, expression and characterization of a novel human Ras-related protein that is regulated by glucocorticoid hormone. *Biochim Biophys Acta* 1999; **1489**: 452 – 456.
54 Fang M, Jaffrey SR, Sawa A, Ye K, Luo X, Snyder SH: Dexras1: a G protein specifically coupled to neuronal nitric oxide synthase via CAPON. *Neuron* 2000; **28**: 183 – 193.
55 Paglia DE, Valentine WN, Brockway RA: Identification of thymidine nucleotidase and deoxyribonucleotidase activities among normal isozymes of 5′-nucleotidase in human erythrocytes. *Proc Natl Acad Sci USA* 1984; **81**: 588 – 592.
56 Lander ES, Linton LM, Birren B *et al*: Initial sequencing and analysis of the human genome. *Nature* 2001; **409**: 860 – 921.
57 Bronner-Fraser M: Mechanisms of neural crest cell migration. *Bioessays* 1993; **5**: 221 – 230.
58 Noden DM: Cell movements and control of patterned tissue assembly during craniofacial development. *J Craniofac Genet Dev Biol* 1991; **11**: 192 – 213.
59 Tan SS, Morriss-Kay GM: Analysis of cranial neural crest cell migration and early fates in postimplantation rat chimaeras. *J Embryol Exp Morphol* 1986; **98**: 21 – 58.
60 Merscher S, Funke B, Epstein JA *et al*: TBX1 is responsible for cardiovascular defects in velo-cardio- facial/DiGeorge syndrome. *Cell* 2001; **104**: 619 – 629.
61 Lindsay EA, Vitelli F, Su H *et al*: Tbx1 haploinsufficieny in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* 2001; **410**: 97 – 101.
62 Jerome LA, Papaioannou VE: DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nat Genet* 2001; **27**: 286 – 291.
63 Sculley DG, Dawson PA, Emmerson BT, Gordon RB: A review of the molecular basis of hypoxanthine-guanine phosphoribosyltransferase (HPRT) deficiency. *Hum Genet* 1992; **90**: 195 – 207.
64 Nyhan WL: The recognition of Lesch-Nyhan syndrome as an inborn error of purine metabolism. *J Inherit Metab Dis* 1997; **20**: 171-178.