Localisation of a gene implicated in a severe speech and language disorder

Simon E. Fisher[1], Faraneh Vargha-Khadem[2], Kate E. Watkins[2], Anthony P. Monaco[1] & Marcus E. Pembrey[3]

Between 2 and 5% of children who are otherwise unimpaired have significant difficulties in acquiring expressive and/or receptive language, despite adequate intelligence and opportunity[1,2]. While twin studies indicate a significant role for genetic factors in developmental disorders of speech and language[1], the majority of families segregating such disorders show complex patterns of inheritance, and are thus not amenable for conventional linkage analysis[2]. A rare exception is the KE family, a large three-generation pedigree in which approximately half of the members are affected with a severe speech and language disorder which appears to be transmitted as an autosomal dominant monogenic trait[3]. This family has been widely publicised as suffering primarily from a defect in the use of grammatical suffixation rules[4-7], thus supposedly supporting the existence of genes specific to grammar. The phenotype, however, is broader in nature, with virtually every aspect of grammar and of language affected[8-10]. In addition, affected members have a severe orofacial dyspraxia, and their speech is largely incomprehensible to the naive listener[10]. We initiated a genome-wide search for linkage in the KE family and have identified a region on chromosome 7 which co-segregates with the speech and language disorder (maximum lod score = 6.62 at θ = 0.0), confirming autosomal dominant inheritance with full penetrance. Further analysis of microsatellites from within the region enabled us to fine map the locus responsible (designated *SPCH1*) to a 5.6-cM interval in 7q31, thus providing an important step towards its identification. Isolation of *SPCH1* may offer the first insight into the molecular genetics of the developmental process that culminates in speech and language.

The precise nature of the disorder segregating in the KE family has been the subject of considerable debate[4,5,8-10], and has played a key role in discussions on the innate aspects of language[6,7,11]. Some authors have suggested that the affected family members suffer primarily from an impairment in generating word endings in accordance with grammatical rules[4,5]. This is often cited as evidence in favour of genes related to specific aspects of grammar[6,7,11]. Other studies have, however, shown that the disorder does not predominantly affect any specific facet of grammatical ability, but instead pervades virtually all aspects of grammar and expressive language, and involves grossly defective articulation characterized as a severe speech dyspraxia[9,10]. In addition, affected individuals have an underlying impairment in co-ordination of the orofacial musculature affecting non-speech movements (an orofacial dyspraxia; ref. 10). Both the orofacial and speech dyspraxias are present in all affected family members, appearing early in their development and persisting throughout their adult life[10]. Finally, there is evidence of cognitive impairment among affected family members, with more profound deficits in the verbal domain[10]. As a result, the full-scale IQs of thirteen affected individuals ranged from 63 to 101, with seven
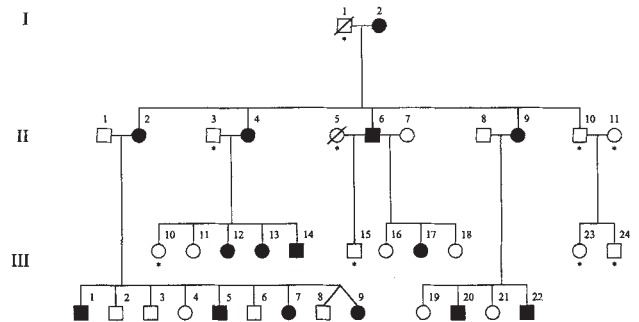

Fig. 1 Pedigree of the speech and language disorder family KE. Affected individuals are shaded. Asterisks indicate those individuals who were unavailable for linkage analysis.

falling below the low average range (80–89) and two scoring 81, in contrast to the eight unaffected individuals tested, whose IQs ranged from 82–118.

Analysis of the KE pedigree (Fig. 1) supports a simple genetic basis for their disorder. About half of the family members are affected, and these include males and females in similar number. In addition, the affected male II-6 has passed on the trait to only one of his three daughters. This inheritance pattern weighs against X-linkage and suggests that the disorder is due to a fully penetrant single locus with an autosomal dominant form of transmission. Cytogenetic analysis of affected subjects II-6 and II-9 failed to reveal any visible chromosomal abnormality.

In light of this, we initiated a genome-wide search for linkage in 27 members of this pedigree, using fluorescence-based geno-
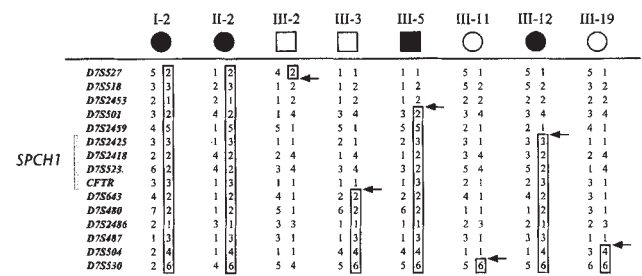

Fig. 2 Haplotypes of a representative selection of informative markers from the *D7S527–D7S530* region of chromosome 7 for key members of the KE pedigree. Numbers used to identify individuals correspond to those in Fig. 1. Paternal haplotypes are on the left, maternal on the right. Boxed areas represent the haplotype which co-segregates with the disorder. Arrows show the positions of recombination events involving the disease chromosome. The region implicated by this analysis is indicated by a shaded line to the left of the loci. All affected members who are not shown in this figure have inherited the non-recombinant disorder-associated haplotype for this region.

[1]*Wellcome Trust Centre for Human Genetics, University of Oxford, Windmill Road, Oxford, OX3 7BN, UK.* [2]*Cognitive Neuroscience Unit, Institute of Child Health, The Wolfson Centre, Mecklenburgh Square, London, WC1N 2AP, UK.* [3]*Mothercare Unit of Clinical Genetics and Fetal Medicine, Institute of Child Health, 30 Guilford St., London, WC1N 1EH, UK. Correspondence should be addressed to A.P.M. e-mail: anthony.monaco@well.ox.ac.uk*

typing of microsatellite markers spaced evenly throughout the genome[12]. Strong evidence in favour of linkage was found for markers on the long arm of chromosome 7, with two closely linked markers, *D7S486* and *CFTR*, giving maximal pairwise lod scores of 6.22 and 5.46, respectively, both at zero recombination (Table 1). We also observed positive lod scores for *D7S527* ($Zmax=1.82$ at $\theta=0.158$) and *D7S530* ($Zmax=2.76$ at $\theta=0.131$) which map on either side of *D7S486* and *CFTR*.

Analysis of the marker haplotypes identified six critical recombinants (Fig. 2), three placing the *SPCH1* locus telomeric to *D7S527* and three placing it centromeric to *D7S530*, thereby localising the gene to the 27.4-cM region between these markers. We further investigated family KE with a large number of additional markers from this interval, selected from the most recent version of the Généthon human linkage map[13]. Of the 20 markers which were found to be informative for this pedigree, all showed linkage to the trait, and 18 gave pairwise maximal lod scores exceeding 3 (Table 1).

Multipoint mapping indicated that the 3 lod unit confidence interval for the location of *SPCH1* spans a 5.6-cM region in chromosomal band 7q31, flanked by *D7S2459* and *D7S643* (Fig. 3). Peak multipoint lod scores of 6.62 were found in the 3.8-cM interval between *D7S2425* and *CFTR*. These results were verified by direct inspection of haplotypes. The recombinations in affected female III-12 and unaffected male III-3 define *D7S2459* and *D7S643* as proximal and distal limits for the locus associated with the disorder (Fig. 2).

The region of 7q implicated in the KE family contains several genes encoding proteins of known functions which may be good candidates to test for mutations in affected individuals[14,15]. These include genes encoding a G-protein–activated phosphoinositide-3 kinase, the interferon-related protein PC4 (which may play a role in differentiative pathways induced by nerve growth factor), Bravo/Nr-CAM (a neuronal cell adhesion molecule), and WNT-2 (a putative signalling molecule involved in

development). In addition, a number of as yet uncharacterized expressed sequence tags are known to map to this interval[14,15].

A search of the Human Cytogenetics Database[16] revealed a single report of an interstitial deletion involving 7q31 (ref. 17). A seven-year old boy presenting with a dysmorphic face and absence of speech, despite language comprehension and psychomotor development equivalent to those of a five-year-old, was found to have a deletion of 7q31.2–32.3. In the absence of further cytogenetic clues, identification of additional families with an articulatory/expressive language disorder co-segregating with 7q31 markers is likely to be important for isolation of *SPCH1*.

A recent genome screen for susceptibility to autism has suggested linkage to an interval of approximately 40cM between *D7S527* and *D7S483* in 7q, with a maximum multipoint lod score of 3.55 in 56 UK affected sib-pair families[18]. Given the high prevalence of language delay amongst autistic probands, it is intriguing to note that the region implicated herein overlaps with that identified in the autism study.

Our data indicate that the disorder segregating in the KE family is likely to have resulted from the disruption of a single gene. At this stage, however, we cannot discount the alternative possibility that the different components of the phenotypic profile are the consequence of a contiguous microdeletion involving several genes in 7q31. This issue is being resolved by the analysis of a somatic hybrid containing the disease chromosome using a series of closely spaced markers from 7q31 (ref. 15). If these investigations do uncover a microdeletion, they will greatly facilitate the future identification of the gene or genes involved.

Table 1• Pairwise lod scores for linkage of *SPCH1* to 7q markers

Locus	Dist (cM)	Recombination fractions (θ)				θ_{max}	Z_{max}
		0.00	0.05	0.10	0.20		
*D7S527**	3.9	$-\infty$	1.16	1.69	1.77	0.158	1.82
D7S518	3.8	$-\infty$	2.74	2.94	2.67	0.105	2.94
D7S2453	3.1	$-\infty$	1.27	1.46	1.19	0.103	1.46
D7S501	0.9	$-\infty$	4.55	4.36	3.59	0.045	4.55
D7S2420	0.0	$-\infty$	4.41	4.19	3.37	0.043	4.41
D7S496	0.0	$-\infty$	3.34	3.44	2.97	0.087	3.45
D7S2459	0.8	$-\infty$	4.83	4.62	3.79	0.043	4.84
D7S2425	0.0	5.67	5.25	4.81	3.84	0.000	5.67
D7S692	1.9	6.62	6.11	5.57	4.40	0.000	6.62
D7S687	0.0	6.02	5.55	5.06	3.99	0.000	6.02
D7S2418	0.5	5.72	5.30	4.85	3.88	0.000	5.72
D7S523	0.0	6.62	6.11	5.57	4.40	0.000	6.62
D7S2554	1.2	5.72	5.27	4.80	3.78	0.000	5.72
D7S522	0.0	5.41	5.01	4.53	3.66	0.000	5.41
D7S2460	0.1	5.79	5.33	4.85	3.80	0.000	5.79
D7S633	0.1	4.53	4.17	3.79	2.97	0.000	4.53
*D7S486**	0.0	6.22	5.71	5.17	4.00	0.000	6.22
*CFTR**	1.0	5.46	5.05	4.63	3.70	0.000	5.46
D7S643	0.9	$-\infty$	4.83	4.62	3.79	0.043	4.83
D7S480	2.2	$-\infty$	4.62	4.41	3.61	0.044	4.63
D7S2486	0.6	$-\infty$	4.55	4.36	3.59	0.045	4.55
D7S487	2.6	3.58	3.25	2.90	2.15	0.000	3.58
D7S504	3.8	$-\infty$	3.45	3.34	2.75	0.054	3.45
*D7S530**		$-\infty$	2.27	2.71	2.59	0.131	2.76

Markers are listed in the order proximal-distal. Distances between neighbouring markers are from the latest Généthon linkage map[13]. Markers indicated with an asterisk were part of the HGMP set[12] used in the genome-wide linkage search; the remaining markers were used for fine mapping of the implicated region. Shading indicates the 3.8–5.6 cM region that co-segregates perfectly with the disorder.
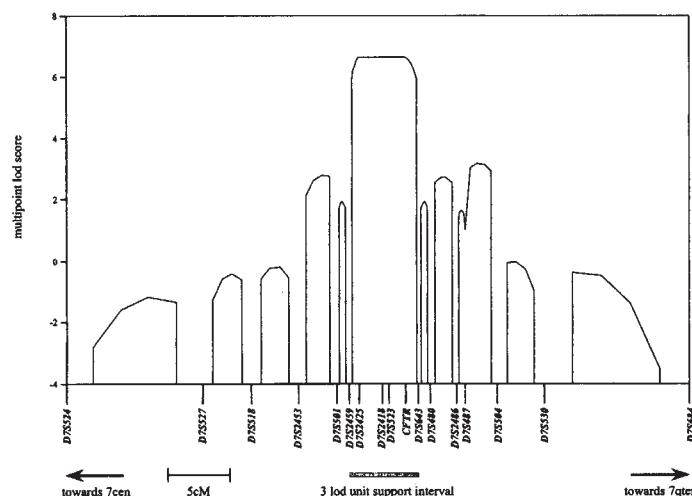

Fig. 3 Multipoint mapping using data from all informative markers in the region identifies a 5.6-cM interval in 7q31 as the 3 lod unit support interval for the location of the *SPCH1* locus. Multipoint lod scores were obtained from a series of sequential 9-point analyses using the VITESSE algorithm[19] and the latest Généthon map[13]. Positions of the selection of markers from Fig. 2 are indicated.

Recently, brain imaging studies of affected and unaffected members of the KE pedigree were carried out to investigate the neural correlates of the disorder found in this family (F.V.-K. *et al.*, unpublished studies). The findings suggest that mutation at the *SPCH1* locus results in functional abnormalities in motor-related areas of the frontal lobe, and that these are due, in turn, to abnormal anatomical development of several brain areas, with a key site of pathology being located in the neostriatum. Analysis of the gene, once identified, could therefore further our understanding of both the structure and function of the fronto-neostriatal system.

In conclusion, this study has provided the first formal evidence for a single autosomal dominant gene involved in a speech and language disorder, and represents a major step towards its identification. Given the phenotypic profile of affected family members, this gene is unlikely to be one specifically involved in grammar. Nevertheless, it is clearly crucial for the normal acquisition of language skills, and isolation of the *SPCH1* locus may thus provide valuable insight into the molecular genetics of the developmental process that culminates in speech and language.

Methods

Semi-automated genotyping. A selection of microsatellite markers from the UK Human Genome Mapping Project (HGMP) set[12] were used in the genome search. PCR primers were labelled with either 6-FAM, HEX or TET phosphoramidite (Applied Biosystems). Pooled products were run on a 373A sequencer (Applied Biosystems), followed by analysis with GENESCAN™ (Ver 2.0) and GENOTYPER™ (Ver 1.1) software to derive allele sizes.

Linkage analysis. Two point lod scores were calculated using the MLINK and ILINK options of the LINKAGE package (version 5.1; ref. 19). Multipoint analyses were performed with the VITESSE algorithm[20], using subsets of eight markers at a time. The sex-averaged marker map used for this analysis was derived from the latest Généthon data[13] and was as follows: *D7S669–7.1–D7S524–11.0–D7S527–3.9–D7S518–3.8–D7S2453–3.1–D7S501–0.9–D7S2420–0.0–D7S496–0.0–D7S2459–0.8–D7S2425–0.0–D7S692–1.9–D7S687–0.0–D7S2418–0.5–D7S523–0.0–D7S2554–1.2–D7S522–0.0–D7S2460–0.1–D7S633–0.1–D7S486–0.0–CFTR–1.0–D7S643–0.9–D7S480–2.2–D7S2486–0.6–D7S487–2.6–D7S504–3.8–D7S530–13.2–D7S684*, with distances in centiMorgans. Marker haplotypes inferred from our data were independently verified using the SIMWALK2 computer program[21].

1. Bishop, D.V.M., North, T. & Donlan, C. Genetic basis of specific language impairment: evidence from a twin study. *Dev. Med. Child Neurol.* **37**, 56–71 (1995).
2. Smith, S.D., Gilger, J.W. & Pennington, B.F. Dyslexia and other specific learning disorders. in *Principles and Practice of Medical Genetics* (eds. D.L. Rimoin, J.M. Connor & R.E. Pyeritz) 1767–1789 (Churchill Livingston, New York, 1996).
3. Hurst, J.A., Baraitser, M., Auger, E., Graham, F. & Norell, S. An extended family with a dominantly inherited speech disorder. *Dev. Med. Child Neurol.* **32**, 347–355 (1990).
4. Gopnik, M. Feature-blind grammar and dysphasia. *Nature* **344**, 715 (1990).
5. Gopnik, M. & Crago, M.B. Familial aggregation of a developmental language disorder. *Cognition* **39**, 1–50 (1991).
6. Pinker, S. Rules of language. *Science* 253, 530–535 (1991).
7. Pinker, S. *The Language Instinct.* (Allen Lane, London, 1994).
8. Fletcher, P. Speech and language defects. *Nature* **346**, 226 (1990).
9. Vargha-Khadem, F. & Passingham, R.E. Speech and laguage defects. *Nature* **346**, 226 (1990).
10. Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P. & Passingham, R. Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proc. Natl. Acad. Sci. USA* **92**, 930–933 (1995).
11. Maynard Smith, J. & Szathmáry, E. *The Major Transitions in Evolution.* (W.H. Freeman, Oxford, 1995).
12. Reed, P.W. *et al.* Chromosome-specific microsatellite sets for fluorescence-based, semi-automated genome mapping. *Nature Genet.* **7**, 390–395 (1994).
13. Dib, C. *et al.* A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380**, 152–154 (1996).
14. Schuler, G.D. *et al.* A gene map of the human genome. *Science* **274**, 540–546 (1996).
15. Bouffard, G.G. *et al.* A physical map of human chromosome 7: an integrated YAC contig map with average STS spacing of 79 kb. *Genome Res.* **7**, 673–692 (1997).
16. Schinzel, A. *Human Genetics Database.* (Oxford University Press, Oxford, 1997).
17. Sarda, P., Turleau, C., Cabanis, M.-O., Jalaguier, J., de Grouchy, J. & Bonnet, H. Interstitial deletion in the long arm of chromosome 7. *Ann. Genet.* **31**, 258–261 (1988).
18. The International Molecular Genetic Study of Autism Consortium. A full genome screen for autism with evidence for linkage to a region on chromosome 7q. *Hum. Mol. Genet.* (in press).
19. Lathrop, G.M., Lalouel, J.-M., Julier, C. & Ott, J. Strategies for multilocus linkage analysis in humans. *Proc. Natl. Acad. Sci. USA* **81**, 3443–3446 (1984).
20. O'Connell, J.R. & Weeks, D.E. The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recording and fuzzy inheritance. *Nature Genet.* **11**, 402–408 (1995).
21. Sobel, E. & Lange, K. Descent graphs in pedigree analysis: applications to haplotyping, location scores and marker sharing statistics. *Am. J. Hum. Genet.* **58**, 1323–1337 (1996).