Human Molecular Genetics, 2003, Vol. 12, No. 3 **349–356**
DOI: 10.1093/hmg/ddg030

Mutation of the *SBF2* gene, encoding a novel member of the myotubularin family, in Charcot–Marie–Tooth neuropathy type 4B2/11p15

Jan Senderek[1,*], Carsten Bergmann[1], Susanne Weber[2], Uwe-Peter Ketelsen[3], Hubert Schorle[2], Sabine Rudnik-Schöneborn[1], Reinhard Büttner[2], Eckhard Buchheim[4] and Klaus Zerres[1]

[1]Department of Human Genetics, Aachen University of Technology, Aachen, Germany, [2]Institute of Pathology, Bonn University Medical School, Bonn, Germany, [3]Department of Neuropaediatrics and Muscular Diseases, Children's Hospital, University of Freiburg, Freiburg, Germany and [4]Department of Paediatrics, Esslingen Community Hospital, Esslingen, Germany

Received October 24, 2002; Revised and Accepted December 2, 2002

Autosomal recessive hereditary motor and sensory neuropathy or Charcot–Marie–Tooth disease (CMT) is a severe childhood-onset neuromuscular disorder. Autosomal recessive CMT is genetically heterogeneous with one locus mapped to chromosome 11p15 (CMT4B2). The histopathological hallmarks of CMT4B2 are focal outfoldings of myelin in nerve biopsies. Homozygosity mapping, in a Turkish inbred family with four children affected by CMT characterized by focally folded myelin, provided linkage to the CMT4B2 locus. We identified a large, novel gene, named *SET binding factor 2* (*SBF2*), that lies within this interval and is expressed in various tissues, including spinal cord and peripheral nerve. SBF2 is a member of the pseudo-phosphatase branch of myotubularins and was an obvious candidate for CMT4B2 by virtue of its striking homology to myotubularin-related protein 2 (MTMR2), causing another form of autosomal recessive CMT with outfoldings of the myelin sheaths. Molecular study of the *SBF2* gene in the CMT4B family demonstrated the presence of a homozygous inframe deletion of *SBF2* exons 11 and 12 in all four affected individuals. On the protein level, this mutation is predicted to disrupt an N-terminal domain that is conserved in SBF2 and its orthologues across species. Myotubularin-related proteins have been suggested to work in phosphoinositide-mediated signalling events that may also convey control of myelination. Localization of *SBF2* within the candidate interval, cosegregation with the disease, expression in the peripheral nervous system, and resemblance of the histopathological phenotype to that related to mutations in its paralogue *MTMR2* indicate that this gene is the CMT4B2 gene.

INTRODUCTION

Hereditary motor and sensory neuropathy (HMSN) or Charcot–Marie–Tooth disease (CMT) comprises a group of clinically and genetically heterogeneous disorders of the peripheral nervous system. With an overall prevalence of 1 in 2500, CMT is the most common inherited neuromuscular disorder in man (1). The clinical features of CMT include progressive distal muscle weakness and atrophy starting in the legs and spreading to the upper extremities, foot deformities, steppage gait, distal sensory loss and decreased or absent tendon reflexes (2). CMT falls into two main subtypes, the demyelinating form, CMT1, and the axonal type, CMT2 (3). CMT1 is characterized by reduced nerve conduction velocities (NCVs) with values <38 m/s for the median motor nerve, segmental de- and re-myelination, and onion bulb formation. These changes differentiate CMT1 from CMT2 in which NCVs are near normal and nerve pathology shows axonal loss and regenerative sprouting. Considerable advances have been made in recent years into the molecular genetics of the CMTs, especially those with autosomal dominant and X-linked inheritance (for review see 4 and the CMT mutation database at http://molgen-www.uia.ac.be/CMTMutations/). This is now extending to the much rarer CMTs with autosomal recessive inheritance that are clinically similar to the dominant forms, but usually more severe with an earlier age of onset (5,6). At least 10 loci are responsible for autosomal recessive CMT and five genes have been identified so far (4). Focally folded myelin has been recognized as a distinctive neuropathological feature in families linked to two regions on chromosome 11, 11q22, referred to as CMT4B1 (7), and 11p15, referred to as CMT4B2 (8). However,

*To whom correspondence should be addressed at: Department of Human Genetics, Aachen University of Technology, Pauwelsstrasse 30, D-52074 Aachen, Germany. Tel: +49 2418080178; Fax: +49 2418082580; Email: jsenderek@ukaachen.de

this subgroup is even more heterogeneous with at least one additional unidentified recessive locus and *de novo* mutations in the *MPZ* gene implicated in dominant forms of demyelinating CMT (9,10). Mutations in the gene for myotubularin-related protein 2 (MTMR2) have been identified in CMT4B1 families (11), while the gene for CMT4B2 has not yet been found.

The myotubularin-related genes define a family of eukaryotic proteins, most of them initially characterized by the presence of a 10-amino acid consensus sequence related to the active sites of different classes of phosphatases (for review see 12,13). Moreover, myotubularin-related proteins have been found to share additional conserved domains such as GRAM (glucosyltransferases, Rab-like GTPase activators, myotubularins) and SID [SET (Suvar3-9, Enhancer-of-zeste, Trithorax) interacting domain] motifs. Myotubularin (MTM1), the founder member, is mutant in myotubular myopathy (14), while its close homologue MTMR2 is involved in CMT4B1 (11). It has been shown that myotubularin is primarily a lipid phosphatase, acting on phosphatidylinositol 3-phosphate [PI(3)P] (15). Phosphoinositides act as signal mediators in a spatially and temporally controlled manner. Information about the timing and location of their production is received by phosphoinositide-binding proteins and transmitted to multiple lines of intracellular events such as signal transduction, cytoskeletal rearrangement and membrane trafficking (for review see 16). Thus, myotubularins might be involved in the modulation of such pathways by controlling phosphoinositide levels. However, the link between the probable alteration of phosphoinositide signalling and the phenotype of patients carrying *MTM1* and *MTMR2* mutations remains mysterious. Notably, some members of the myotubularin family lack phosphatase activity due to germline-encoded mutation of the active site. These so-called pseudo-phosphatases contain additional conserved domains such as DENN [differentially expressed in normal and neoplastic cells, synonymous with GEF (guanine nucleotide exchange factor) or AEX-3] and PH (pleckstrin homology) motifs. The prototype of pseudo-phospatases, SET binding factor 1 (SBF1), has initially been suggested to be involved in oncogenic transformation (17,18); however, more recent results indicate that SBF1 may release growth inhibition signals, dependent on its subcellular localization (19).

The genetic heterogeneity of autosomal recessive peripheral neuropathies and the limited number and size of families affected by any single disorder are a major obstacle to molecular research and gene identification. We combined classic linkage studies with the tools of bioinformatics based on the hypothesis that paralogues of CMT genes, which reside in CMT critical regions should have a good probability as candidate genes. Using this strategy, we characterized a candidate gene for CMT4B2 and disclosed a *bona fide* disease-causing mutation in a family with autosomal recessive CMT with focally folded myelin.

RESULTS

Linkage analysis

Family CMT281 is a consanguineous Turkish family with four patients affected with severe sensorimotor neuropathy characterized by focally folded myelin sheaths in nerve biopsies (Fig. 1, for details, see Patients and Methods section). After exclusion of mutations in the *PMP22, P0* and *EGR2* genes, linkage analysis for known autosomal recessive CMT loci was performed. Linkage with CMT4A, CMT4C, CMT4B1, HMSNL, CMT4F, HMSNR, ARCMT2A and ARCMT2B was excluded. Affected patients were found to be homozygous for markers from within the CMT4B2 region on chromosome 11p15 (Fig. 1). A maximum two-point LOD score of 2.45 at $\theta = 0$ was calculated for marker D11S4465. Haplotypes were estimated using the SimWalk2 program that generated a peak LOD score of 4.39 between D11S4149 and D11S4465.

Identification and characterization of the CMT4B2 gene

Nine candidate sequences from the CMT4B2 critical interval matched the criteria as stated in the Patients and Methods section, i.e. (1) providing good likelihood of being true genes, (2) showing homology to known CMT genes, and (3) being expressed in spinal cord or sciatic nerve. A cluster of overlapping ESTs (GenBank Accession no. BM015750, BE778113, BI094130, BI255224, BC011143) was selected as the most promising candidate. The predicted protein displayed 36% sequence identity with aminoacids 345–581 of MTMR2 and shared the phosphatase signature and the SID domain. For the remaining candidates, homology was confined to a single motif, partly associated with a high *a priori* likelihood of occurrence, while overall domain organization and composition were different from the CMT paralogue in question.

The selected EST cluster, that we termed *Par540A21* gene, appeared to have a mouse orthologue in mRNA BC015069 (Fig. 2B). The murine mRNA extended 5′ and shared its first two exons with a large human transcript annotated about 23 kb upstream from the EST cluster. This sequence was known as *KIAA1766* (www.kazusa.or.jp/huge; GenBank Accession no. AB051553) and was believed to have a verified 3′ end. However, using mRNA extracted from adult sciatic nerve as template, primers positioned in *KIAA1766* and *Par540A21* yielded a product of the expected size. Northern blots probed with *KIAA1766* exons 18/19 confirmed that the transcript was considerably longer than was suggested by the *KIAA1766* mRNA (Fig. 3A). We went on to determine the structure of the complete transcript and its genomic organization. We used a PCR-based approach with primers strategically positioned within exons of *KIAA1766* and *Par540A21*. Human total RNA prepared from sciatic nerve, spinal cord, cerebral cortex and blood cells was used as template. The sequences of all exons and splicing junctions were determined by sequencing of PCR products. These studies confirmed that *KIAA1766* and *Par540A21* were part of the same gene. The putative first exon, by alignment with ESTs BE256927 and BE255129, was supported by FirstEF prediction. This exon contains an ATG start codon with sequences conforming to a Kozak consensus [(A/G)CC ATG G]. Sequencing confirmed the presence of an inframe stop codon within the last exon of *Par540A21*. A putative polyadenylation signal was discovered 1.7 kb downstream of the stop codon and ascertained by alignment with additional ESTs. Notably, gene prediction tools Fgenesh and

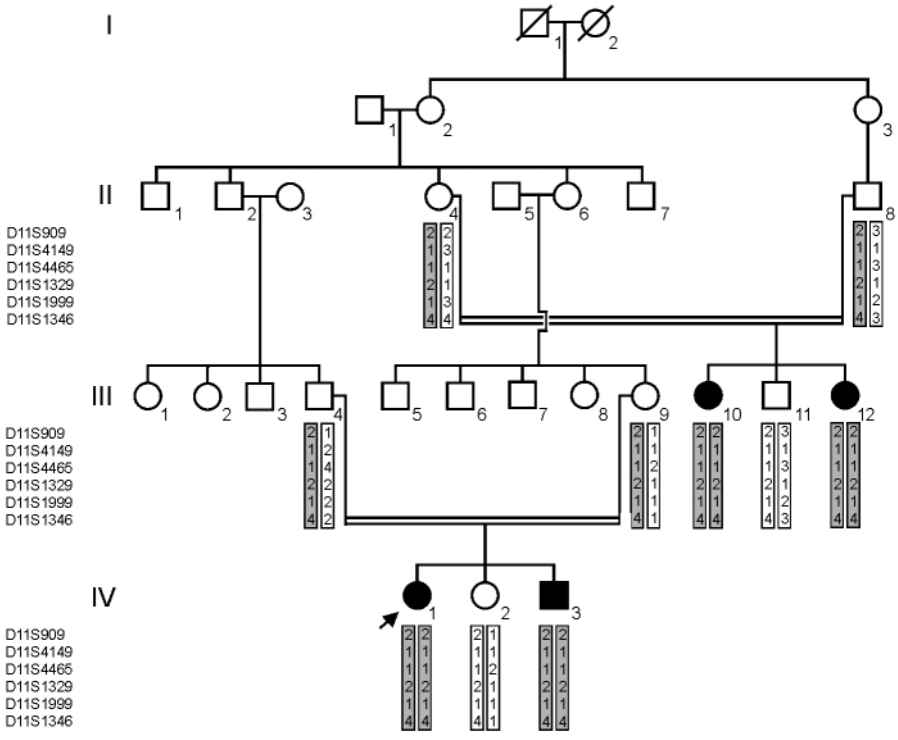

Figure 1. Pedigree of family CMT281. The arrow indicates the proband who underwent nerve biopsy. Haplotypes for markers from the CMT4B2 region are given. Square = male, circle = female, solid symbol = affected patient, cross-hatched symbol = deceased individual.

Genscan did not predict the full-length mRNA, however, both algorithms recognized that *KIAA1766* was part of a larger gene.

Using the BLAT tool implemented on the UCSC website we aligned the composite mRNA with the genomic sequence. The gene is flanked by STR markers tel-D11S4149 and D11S1904-cen, runs cen-tel, and spans at least 515 kb of genomic sequence (Fig. 2A). The total number of exons that we identified is 40. This may be a conservative estimate, since splice variants might be detected if mRNA from a suitable tissue was used as template. The size of exons ranged from 69 to 263 bp. All splice junctions followed the canonical AG/GT rule. Lengths of introns were between 446 bp and 150 kb (supplement 1).

Commercially available human adult multiple-tissue northern blots were hybridized with a probe corresponding to exons 19/20 (former *KIAA1766* exons 18/19). A single, discrete message of about 8.5 kb was detected in a variety of tissues with highest expression in brain (Fig. 3A). The signal in spinal cord was also well detectable. Using RT–PCR we could show that levels of expression are comparable in spinal cord and sciatic nerve (Fig. 3B). The composite cDNA experimentally amplified, from sciatic nerve cDNA, as four overlapping fragments, is 7.4 kb in length (supplement 2). Considering likely occurrence of 5′ truncation and poly(A) tailing, this mRNA corresponds to the size of the about 8.5 kb band seen in the northern blot reasonably well. Northern blotting did not yield evidence of splice variants. However, in RT–PCR experiments performed on spinal cord and sciatic nerve mRNA, we could confirm the existence of the *KIAA1766* message as initially described in the HUGE database. This variant probably results from leakiness of intron 26 donor splice signal. A stop codon is reached after intronic read-through of 4 bp and the 5′ portion of intron 26 forms the 3′ UTR (Fig. 2C).

By conceptual translation, the composite cDNA is predicted to encode a protein of 1849 amino acids (supplement 3). As found initially, this novel protein shares homology with MTMR2. BLAST2 alignment revealed 30% sequence identity with MTMR2 amino acids 223–581. In addition to the phosphatase signature and the SID motif, SMART analysis predicted a GRAM domain in the novel protein. However, similarity was too weak to the MTMR2 GRAM motif to produce a BLAST alignment (at parameters set to default). Sequence comparison of the phosphatase domain with the consensus sequence revealed that the catalytic cysteine was substituted by a leucine residue predicted to abolish phosphatase activity. The protein was predicted, by SMART, to contain additional domains, DENN and PH (Fig. 2C). Overall, domain organization suggested that the novel protein represents a pseudo-phosphatase within the myotubularin family (Fig. 4). BLASTP analysis revealed that, in human, it has highest homology to the prototype of such pseudo-phosphatases, SET binding factor 1, SBF1, with the region of homology stretching over the full length of both proteins, sharing 59% identical amino acids. Hereafter, we named the novel protein SET binding factor 2, SBF2. Even stronger similarity was noted between SBF2 and the protein predicted from murine mRNA BC015069, initially found to link *Par540A21* and *KIAA1766.* This suggests that SBF2, its putative mouse orthologue, and by extension, their homologues in other species, constitute a

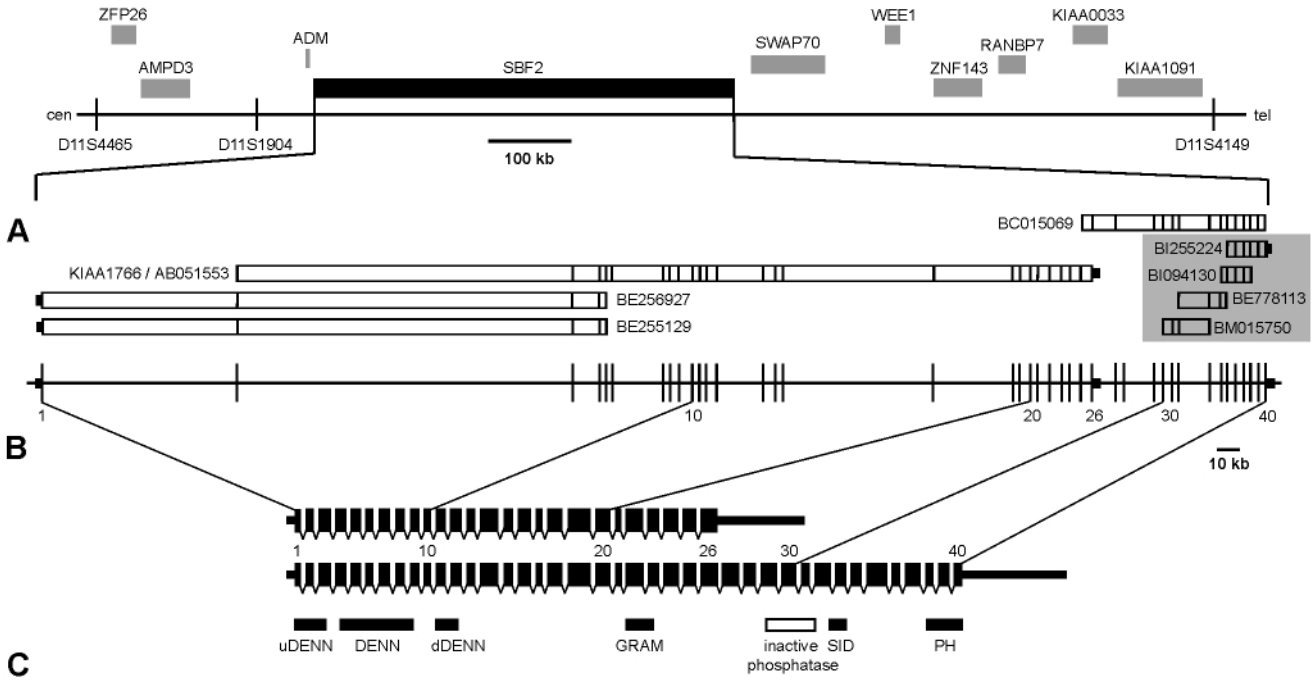

Figure 2. Genomic organization of *SBF2* and structure of its transcript. (**A**) Partial physical map of the CMT4B2 candidate region containing the *SBF2* gene. This map is based on the December 2001 release of the UCSC genome assembly. (**B**) Schematic representation of the intron/exon structure. The set of expressed sequences initially considered as an *MTMR2* paralogue is located at the gene's 3′ end (grey box). Mouse mRNA BC015069 provided a link between this EST cluster and *KIAA1766* from the HUGE database. (**C**) *SBF2* mRNA as assembled from overlapping RT–PCR fragments. A set of 40, non-overlapping exons spans the entire length *SBF2* mRNA. The shorter, 26 exon transcript corresponding to *KIAA1766*, was not observed by Northern blotting but detected by RT–PCR.

novel branch in the phylogenetic tree of myotubularin-related pseudo-phosphatases.

Mutation analysis

PCR-amplification of *SBF2* exons 11 and 12 failed in DNA samples from the patients of family CMT281, while it was obtained in healthy sibs and the parents from the nucleus families. Amplification of sequences corresponding to exons 10 and 13 was possible in all family members. We hypothesized that these findings may have resulted from a genomic deletion with putative breakpoints in intron 10/exon 11 and exon 12/intron 12. Direct sequencing of the predicted *SBF2* coding exons 1–40 did not unravel any additional alteration.

Deletion breakpoints were defined by PCR amplification with primers strategically placed in introns 10 and 12. Primers placed 0.6 kb upstream of exon 11 and 0.9 kb downstream of exon 12 resulted in amplification of an estimated band of about 2.3 kb in healthy controls. Individuals from family CMT281 found heterozygous for the CMT4B2 haplotype were characterized by an additional PCR product at about 1 kb. In case of an affected patient, PCR resulted only in the smaller product. Sequencing of this putative junction fragment allowed identification of the deletion breakpoints which were found at position 31 of exon 11 and at position 646 downstream of exon 12 encompassing a deleted sequence of 1.301 bp. PCR with primers as mentioned did not detect the junction fragment corresponding to the deletion allele in altogether 150 DNAs of healthy, unrelated control individuals. Sixty of these controls were of Turkish origin.

We used the repeat masker program (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker) to identify whether the genomic region encompassing the deletion had any homology to known human repetitive elements. Homology with several motifs detected was only weak, and cleavage did not occur there. There was no sequence homology between the 5′ and the 3′ junction fragments, suggesting a non-homologous recombination event.

On the transcript level different consequences might be expected for this deletion, i.e. fusion of exons 10 and 13, use of a cryptic splice donor site for exon 11, or read-through of the remainder of intron 12. Alternatively, the altered mRNA might be unstable and undergo rapid degeneration. To test for these hypotheses, we obtained leukocyte mRNA from patient IV.1 and demonstrated by RT–PCR and sequencing that exons 11 and 12 were absent from the mature transcript with exon 10 fused to exon 13 (c.Ex11_Ex12del, Fig. 5). As the junction of exons 10 and 13 is in frame, this deletion is not expected to result in premature stop of tanslation. However, the deletion disrupts the N-terminal DENN domain by removing the unique dDENN motif which is represented by translation of exons 11 and 12 (L351_E432del).

To determine the frequency of *SBF2* mutations in autosomal recessive or sporadic sensorimotor neuropathies, we screened 15 additional affected individuals. These patients were from non-consanguineous German families too small for linkage analysis. None of these patients had focally folded myelin on nerve biopsy. Mutations in *PMP22, Cx32, MPZ, EGR2, GDAP1* and *PRX* genes had been excluded previously. We tested for the deletion mutation found in family CMT281 and

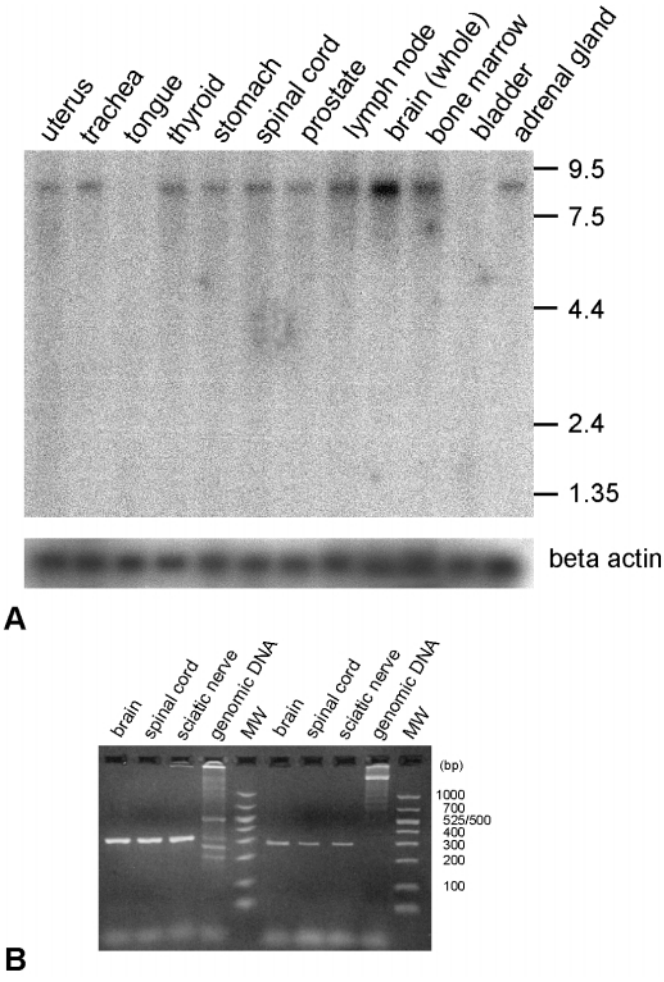

Figure 3. SBF2 expression profile. (**A**) Human adult multiple-tissue northern blot probed with SBF2 exons 19/20. A single message of about 8.5 kb is detected in a variety of tissues with the strongest signal observed in brain. Control hybridisation with a beta actin probe is given below. (**B**) Expression analysis by semiquantitative RT–PCR. Left: beta 2-microglobulin specific primers were used in control reactions. Right: amplification of a 298 bp *SBF2* fragment corresponding to exons 19 and 20 yielded similar results with mRNA derived from spinal cord and sciatic nerve. Genomic DNA was employed to ensure specific amplification from cDNA targets.

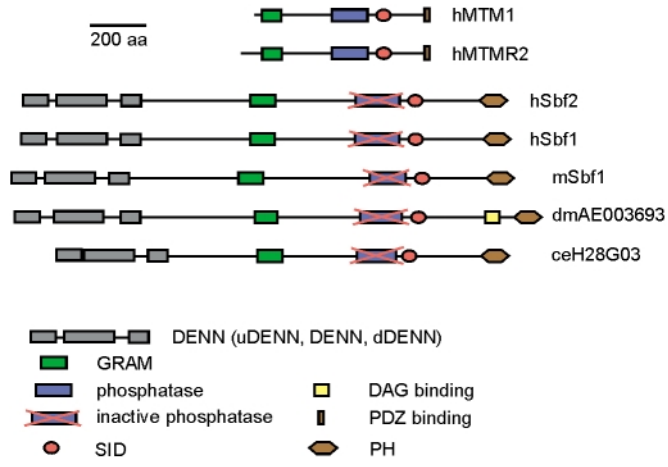

Figure 4. Comparison of the domain architecture of SBF2 and other myotubularin-related proteins. The proteins and the respective motifs are represented to scale. Those members lacking phosphatase activity (pseudophosphatases) constitute a distinct subgroup with additional domains conserved from *Caneorhabditis elegans* (ce) and *Drosophila melanogaster* (dm) to mouse (m) and human (h).

sequenced the entire coding region of *SBF2*. We identified five intronic polymorphisms, but none of these individuals was found to have a causative mutation in *SBF2*.

DISCUSSION

We have reported here the identification of a novel gene, *SET binding factor 2, SBF2*, mutated in a form of autosomal recessive CMT with focally folded myelin sheaths (CMT4B2/11p15). *SBF2* was a *bona fide* candidate for CMT4B2 as the encoded protein displays significant similarity to myotubularin related protein 2, MTMR2, involved in another form of CMT with focally folded myelin (11).

By carrying out homozygosity mapping, we found linkage to the CMT4B2 locus in a Turkish family with autosomal recessive CMT. Notably, this family shared the distinct neuropathologic changes and clinical features with the original CMT4B2 family. In both families, the disease started during the first two decades of life, muscle weakness and atrophy were distally pronounced, nerve conduction was severely slowed, and sural nerve biopsies revealed a demyelinating type of neuropathy with focally folded myelin sheaths (8). Although two-point LOD scores did not reach significance in family CMT281, haplotype based, multipoint data, exclusion of linkage to all other known CMT loci, and phenotypic homogeneity provided high prior probability that the mutation was in the so far unknown CMT4B2 gene. *SBF2* was regarded as the most plausible candidate within the CMT4B2 interval and subsequently subjected to mutation analysis. Detection of a gross genomic inframe deletion which showed segregation with the disease phenotype in an extended pedigree makes it highly likely that this gene is the disease susceptibility locus. Nevertheless, it is theoretically possible that an inframe deletion may represent merely a rare genetic variant, which does not cause the disorder. However, from what is known from compiling human mutation data, deletion of exons encoding non-repetitive parts of a protein are unexpected to remain without clinical consequences (Michael Krawczak, personal communication).

Autosomal recessive CMT is heterogeneous. To date, nine loci for autosomal recessive CMT have been assigned, but none has been shown to make up more than at best 20% of cases. Therefore, it was not unexpected that we did not find any *SBF2* mutation in 15 German patients with autosomal recessive or sporadic CMT. Notably, none of these additional patients presented with focally folded myelin sheaths on nerve biopsy which might be a distinctive feature in the setting of a *SBF2* mutation. Interestingly, mutations in *MTMR2* seem to be almost exclusively restricted to patients with this peculiar type of peripheral myelin pathology (20).

SBF2 adds to the expanding family of myotubularin-related genes which are highly conserved among eukaryote genomes

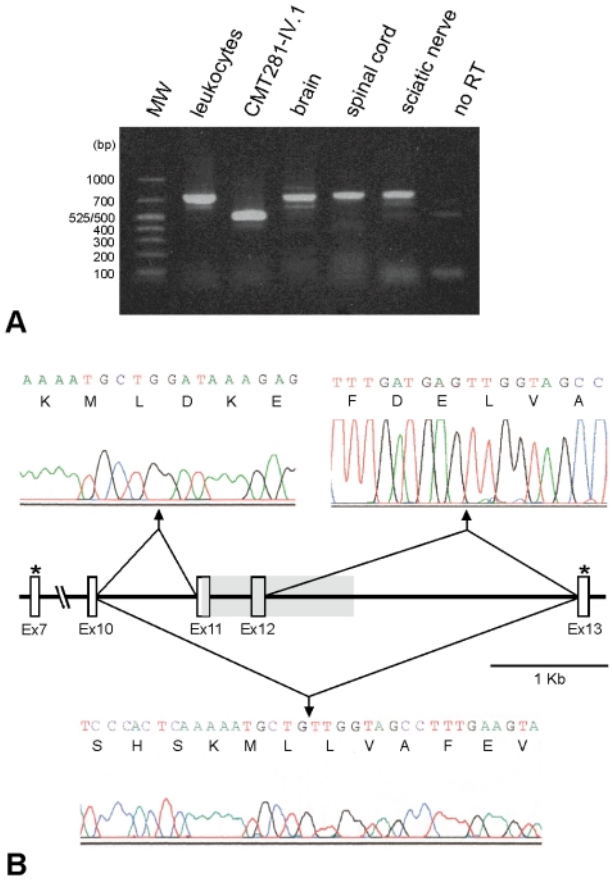

Figure 5. The *SBF2* c.Ex11_Ex12del (L351_E432del) mutation in family CMT281. (**A**) Detection of the deletion mutation by RT–PCR. Primers placed in exons 7 and 13 yielded a fragment of 740 bp in case of wild-type mRNA from different sources. A shorter 497 bp fragment was obtained when using patient's leukocyte mRNA as a template (no RT: no reverse transcriptase added for cDNA synthesis). (**B**) Sequencing of the exon10/13 junction. Top: exon 10/11 and exon 11/12 splice junctions found on wild-type mRNA. Middle: genomic organization of *SBF2* around exons 10–13. The extent of the deletion is indicated as a grey box. The positions of primers for RT–PCR as described in (A) are indicated by asterisks. Bottom: exon 10/13 junction as observed in case of the deletion mutation.

suggesting an important role through evolution. Notably, SBF2 shows 90% identity with a partial sequence of its putative murine orthologue (GenBank Accession no. AAH15069). By homology, SBF2 belongs to the pseudo-phosphatase branch of the myotubularin family, characterized by germline-encoded alterations of their phosphatase motif that renders them catalytically inactive. Although the current concepts of how mutations of the *SBF2* paralogue *MTMR2* cause CMT4B1 are far from being complete, loss of phosphatase activity seems to be a common feature of naturally occurring MTMR2 mutants (21). Since SBF2 lacks intrinsic phosphatase activity it presumably acts differently from MTMR2 in the phosphoinositide signalling network. However, possibly both signalling events converge on a common pathway conveying control of myelin formation in the peripheral nervous system. It has been suggested that active and inactive members of the myotubularin family might work in specific pairs (22), however, this intriguing hypothesis remains to be confirmed.

The *SBF2* deletion observed in the CMT4B2 family disrupts the dDENN motif which, with uDENN and DENN, constitutes the DENN/GEF/AEX-3 signature evolutionary conserved in a multitude of signalling molecules (23). Although the exact function of the DENN motif remains elusive there is experimental evidence that its mutation alters the intrinsic properties of the respective protein. MAPK/ERK2 regulatory protein p70 selectively lacking dDENN has a transforming effect on NIH3T3 cells when overexpressed (24). Moreover, disruption of the DENN domain of SBF1 results in oncogenic transformation of transfected NIH3T3 cells (19). Interestingly, while wild-type SBF1 exclusively localized to the cytoplasm, deletion of the DENN domain resulted in partial nuclear localization. It is therefore conceivable that DENN domains sequester proteins to the cytoplasmic compartment. Disruption of DENN might then allow shifting of the molecule to another cellular localization, e.g. the nucleus, where it may interfere with signalling pathways that result in oncogenic transformation (as to SBF1) or initiate a program of irregular myelination (as to SBF2). Given excessive, displaced myelin outgrowth in CMT4B it would not be surprising if this form of CMT resulted from mutation of a cellular signalling pathway rather than from mutation of a structural myelin protein.

In summary, our observations do not only define a cause of hereditary neuropathies but they also extend the spectrum of myotubularin-related proteins involved in human disease. We have shown that alteration of SBF2, unlike what is suspected for its homologue SBF1, does not predispose to neoplasia or dysfunctions in rapidly replicating tissues, but is instead implicated in the pathogenesis of progressive demyelinating neuropathy. Further studies should clarify the mechanism by which SBF2 acts to maintain structural and functional integrity of the peripheral nervous system.

PATIENTS AND METHODS

Family CMT281

This inbred family (Fig. 1) with two loops of consanguinity was of Turkish origin. The parents in the nucleus families were normal by history and two of them (III.4, III.9) were confirmed unaffected by clinical and electrophysiological investigation. The four affected individuals presented a severe sensorimotor neuropathy phenotype. Clinical variability was mainly related to age. Patients III.10 and III.12, aged 32 and 27 years, were ascertained by the Department of Neurology of Istanbul University. They had disease onset at around age 5 and are only able to ambulate with the help of walking aids due to distal muscle wasting and foot deformity. Neurophysiologic examination revealed severely reduced NCVs. In patient IV.1, difficulty with gait became evident at the age of 4. She progressively developed hammertoes and pes cavus. On examination at age 13 she walked with steppage gait. Atrophy and weakness were distally pronounced in the upper and lower limbs and were most prominent in anterior tibial, peroneal and intrinsic foot muscles (grade 2 according to the MRC scale). She had difficulty with grip due to atrophy of thenar muscles and contractures of the fingers. Proximal muscular properties were widely preserved. Upper limb tendon reflexes were diminished and lower limbs

were areflexic. Sensory qualities were also impaired in the lower limbs. Motor NCV was reduced to 16.5 m/s in the median nerve and to 18.8 m/s in the ulnar nerve. At age 4, her brother (IV.3) presented with slight gait ataxia but muscular properties were preserved. No foot deformity was noted. Lower limb deep tendon reflexes could not be elicited and NCV measurement revealed slowing of peroneal nerve motor conduction (18.5 m/s).

Sural nerve biopsy performed on patient IV.1 at age 8 revealed a severe loss of myelinated fibres (fibre density 2720/mm^2). Surviving myelinated axons were generally hypomyelinated but focally the myelin protruded from the outer aspect of the sheath. Electron microscope sections through such regions demonstrated that these protrusions consisted of myelin thickenings superficially resembling tomacula but which were smaller in diameter and more irregular in contour. Myelin protrusions projecting into the axons were also observed.

Genotyping

Blood was collected from participating family members and DNA was extracted according to standard procedures. Mutations in the *PMP22, MPZ, Cx32* and *EGR2* genes for demyelinating CMT were excluded by sequencing analysis of all coding exons. Genotyping for the known ARCMT loci on chromosomes 8q21 (CMT4A), 5q32 (CMT4C), 11q22 (CMT4B1), 11p15 (CMT4B2), 8q24 (HMSNL), 19q13 (CMT4F), 10q23 (HMSNR), 1q21 (ARCMT2A) and 19q13 (ARCMT2B) was performed to test for homozygosity. The following microsatellite markers were used: D8S286, D8S551, D8S548 and D8S1829 (CMT4A); D5S658, D5S436, D5S2090 and D5S636 (CMT4C), D11S4118, D11S1757, D11S1333 and D11S1366 (CMT4B1); D11S909, D11S4149, D11S4465 and D11S1346 (CMT4B2); D8S378, D8S259, D8S256 and D8S1796 (HMSNL); D19S881, D19S47, D19S223 and D19S400 (CMT4F); D10S1670, D10S210, D10S1678 and D10S1742 (HMSNR); D1S1153, D1S2777, D1S2721 and D1S2624 (ARCMT2A); D19S902, D19S879, D19S867 and D19S907 (ARCMT2B). After observing homozygosity for CMT4B2 markers, five additional microsatellites were typed: D11S932, D11S1329, D11S1999, D11S875 and D11S4189. Primers for PCR amplification were as published by the Genome Data Base (www.gdb.org). Sense primers were labelled with FAM fluorophores (Pharmacia, Uppsala, Sweden) for electrophoresis and analysis on an ABI PRISM 377 DNA sequencer (Applied Biosystems, Weiterstadt, Germany). Two-point LOD scores were calculated using the program MLINK of LINKAGE (25). The disorder was coded as fully penetrant and recessive. Disease-gene frequency was defined as 0.0001. SimWalk2 program was utilized to perform multipoint analyses (26).

In silico identification of candidate genes in the CMT4B2 interval

The CMT4B2 region was flanked by markers D11S1331 and D11S4189 encompassing an interval of about 5.3 cM on the Marshfield sex averaged map (http://research.marshfieldclinic.org/genetics/). This interval equalled a physical distance of about 4.5 Mb in the December 2001 release of the public genome assembly (http://genome.ucsc.edu/) and was reduced to 4.15 Mb after correction of several errors of clone orientation arising from assembly artefacts. Twenty-two known genes as curated by the RefSeq project at NCBI were annotated within this interval. The UCSC map provides aligned mRNAs and ESTs from GenBank, gene builds by Ensembl and Acembly, gene predictions from Genscan and Fgenesh, as well as alignment of other genomes by similarity. We categorized these sequences in terms of the levels of supporting evidence and hence likelihood of being a true gene: (1) Ensembl or Acembly annotation or evidence from mRNA or at least two spliced ESTs; (2) evidence from all three—(a) a single EST and (b) gene prediction and (c) comparative genomics; (3) evidence from two out of (a), (b), (c); (4) evidence from one out of (a), (b), (c). A set of 20 clusters corresponding to putative genes was generated by this approach.

We hypothesized that paralogues of known CMT genes mapping to critical intervals should have a good potential as candidate genes. Therefore, by using BLAST2 alignments (www.ncbi.nlm.nih.gov/BLAST), the 42 proteins predicted from the CMT4B2 interval were investigated for homology with the 16 known proteins involved in the advent of CMT and related disorders (http://molgen-www.uia.ac.be/CMTMutations/). Sequences which showed homology were tested for 3′ and 5′ truncation by carefully checking the genome assembly at these sites, and by BLAST searches against GenBank and dbEST. Putative full-length sequences were again aligned with CMT proteins by BLAST2 and analysed for domain organization and architecture using the SMART program (http://smart.embl-heidelberg.de/).

We went on to question whether these candidates were expressed in spinal cord and peripheral nerve by querying public data repositories (www.ncbi.nlm.nih.gov/dbEST/index.html, www.ncbi.nlm.nih.gov/SAGE/) for EST matches and SAGE tags. This was done as the last step as, due to scarcity of libraries pulled from these tissues, a high dropout was expected. Therefore, in parallel, expression of putative candidates was tested by RT–PCR.

RT–PCR

Products were generated using 1 µg of sciatic nerve, spinal cord, cerebral cortex or leukocyte mRNA as template. The total RNA was extracted by using the RNeasy system (Qiagen, Hilden, Germany) and was reverse transcribed by oligo-dT primers and AMV reverse transcriptase (Promega, Mannheim, Germany). Sequencing of the PCR fragments was done using ABI Prism BigDye chemistry on an ABI310 genetic analyser (Applied Biosystems).

Northern blot

A 298 bp probe was amplified by use of a cloned gene fragment as template. PCR products were gel purified, [^{32}P]-labelled by the multiprime method, and hybridized to human adult multiple tissue northern blots (Clontech, Palo Alto, CA, USA). Hybridization was performed for 1 h at 68°C, with Express Hyb solution (Clontech), and blots were washed under the following conditions: 3 × 15 min in 2 × SSC, 0.05% SDS at room temperature followed by 3 × 15 min in 0.1 × SSC, 0.1% SDS at 50°C. Filters were exposed to autoradiography films for 72 h. Control hybridisations were performed employing beta-actin probes.

Gene prediction

Complete exon–intron structure of the identified gene was deduced from genomic alignment of overlapping RT–PCR fragments amplified from peripheral nerve and spinal cord mRNA. Putative transcription start sites were checked for homology to Kozak sequences (27) and the likelihoods of exons being the first exon of the gene were tested with the FirstEF algorithm (http://rulai.cshl.org/cgi-bin/tools/FirstEF/fef.cgi). Putative polyadenylation signals were considered using the Polyadq algorithm (http://largon.cshl.org/tabaska/polyadq4.cgi).

Mutation detection

PCR primers flanking individual exons and offset, by 20–60 bp, from intron–exon junctions were designed by the program Primer3 (www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi/) and were used to amplify 20 ng of genomic DNA from patients and controls. Mutation detection was performed by direct sequencing by the use of ABI Prism BigDye technology. Samples were run and analysed on an ABI PRISM 310 genetic analyser (Applied Biosystems). A PCR-based strategy was adapted to search for a putative deletion of exons 11 and 12. Primers positioned in introns 10 and 12 were variably combined for amplification of the breakpoint region. The junction fragment obtained was directly sequenced with the PCR primers. These primers were subsequently employed for detection of the mutation in family members and controls.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

The authors are indebted to the CMT patients and their relatives for taking part in this study. We thank Professor Tiemo Grimm and Dr Bertram Müller-Mysok for performing multipoint-linkage analysis. We are grateful to Professor Michael Krawczak for sharing information on human mutation data.

REFERENCES

1. Skre, H. (1974) Genetic and clinical aspects of Charcot–Marie–Tooth's disease. *Clin. Genet.*, **6**, 98–118.
2. Harding, A.E. and Thomas, P.K. (1980) The clinical features of hereditary motor and sensory neuropathy types I and II. *Brain*, **103**, 259–280.
3. Dyck, P.J., Chance, P., Lebo, R. and Carney, J.A. (1993) Hereditary motor and sensory neuropathies. In Dyck, P.J., Thomas, P.K., Griffin, J.W., Low, P.A. and Podulso, J.F. (eds), *Peripheral Neuropathy.* WB Saunders, Philadelphia, PA, pp. 1094–1136.
4. Vance, J.M. (2000) The many faces of Charcot–Marie–Tooth disease. *Arch. Neurol.*, **57**, 638–640.
5. Harding, A.E. and Thomas, P.K. (1980) Autosomal recessive forms of hereditary motor and sensory neuropathy. *J. Neurol. Neurosurg. Psychiatry*, **43**, 669–678.
6. Thomas, P.K. (2000) Autosomal recessive hereditary motor and sensory neuropathy. *Curr. Opin. Neurol.*, **13**, 565–568.
7. Bolino, A., Brancolini, V., Bono, F., Bruni, A., Gambardella, A., Romeo, G., Quattrone, A. and Devoto, M. (1996) Localization of a gene responsible for autosomal recessive demyelinating neuropathy with focally folded myelin sheaths to chromosome 11q23 by homozygosity mapping and haplotype sharing. *Hum. Mol. Genet.*, **5**, 1051–1054.
8. Othmane, K.B., Johnson, E., Menold, M., Graham, F.L., Hamida, M.B., Hasegawa, O., Rogala, A.D., Ohnishi, A., Pericak-Vance, M., Hentati, F. and Vance, J.M. (1999) Identification of a new locus for autosomal recessive Charcot–Marie–Tooth disease with focally folded myelin on chromosome 11p15. *Genomics*, **62**, 344–349.
9. Fabrizi, G.M., Taioli, F., Cavallaro, T., Rigatelli, F., Simonati, A., Mariani, G., Perrone, P. and Rizzuto, N. (2000) Focally folded myelin in Charcot–Marie–Tooth neuropathy type 1B with Ser49Leu in the myelin protein zero. *Acta Neuropathol.*, **100**, 299–304.
10. Nakagawa, M., Suehara, M., Saito, A., Takashima, H., Umehara, F., Saito, M., Kanzato, N., Matsuzaki, T., Takenaga, S., Sakoda, S. *et al.* (1999) A novel MPZ gene mutation in dominantly inherited neuropathy with focally folded myelin sheaths. *Neurology*, **52**, 1271–1275.
11. Bolino, A., Muglia, M., Conforti, F.L., LeGuern, E., Salih, M.A., Georgiou, D.M., Christodoulou, K., Hausmanowa-Petrusewicz, I., Mandich, P., Schenone, A. *et al.* (2000) Charcot–Marie–Tooth type 4B is caused by mutations in the gene encoding myotubularin-related protein-2. *Nat. Genet.*, **25**, 17–19.
12. Laporte, J., Blondeau, F., Buj-Bello, A. and Mandel, J.L. (2001) The myotubularin family: from genetic disease to phosphoinositide metabolism. *Trends Genet.*, **17**, 221–228.
13. Wishart, M.J., Taylor, G.S., Slama, J.T. and Dixon, J.E. (2001) PTEN and myotubularin phosphoinositide phosphatases: bringing bioinformatics to the lab bench. *Curr. Opin. Cell Biol.*, **13**, 172–181.
14. Laporte, J., Hu, L.J., Kretz, C., Mandel, J.L., Kioschis, P., Coy, J.F., Klauck, S.M., Poustka, A. and Dahl, N. (1996) A gene mutated in X-linked myotubular myopathy defines a new putative tyrosine phosphatase family conserved in yeast. *Nat. Genet.*, **13**, 175–182.
15. Taylor, G.S., Maehama, T. and Dixon J.E. (2000) Inaugural article: myotubularin, a protein tyrosine phosphatase mutated in myotubular myopathy, dephosphorylates the lipid second messenger, phosphatidylinositol 3-phosphate. *Proc. Natl Acad. Sci. USA*, **97**, 8910–8915.
16. Itoh, T. and Takenawa, T. (2002) Phosphoinositide-binding domains. Functional units for temporal and spatial regulation of intracellular signalling. *Cell. Signal.*, **14**, 733–743.
17. Cui, X., De Vivo, I., Slany, R., Miyamoto, A., Firestein, R. and Cleary, M.L. (1998) Association of SET domain and myotubularin-related proteins modulates growth control. *Nat. Genet.*, **18**, 331–337.
18. De Vivo, I., Cui, X., Domen, J. and Cleary, M.L. (1998) Growth stimulation of primary B cell precursors by the anti-phosphatase SBF1. *Proc. Natl Acad. Sci. USA*, **95**, 9471–9476.
19. Firestein, R. and Cleary, M.L. (2001) Pseudo-phosphatase SBF1 contains an N-terminal GEF homology domain that modulates its growth regulatory properties. *J. Cell Sci.*, **114**, 2921–2927.
20. Bolino, A., Lonie, L.J., Zimmer, M., Boerkoel, C.F., Takashima, H., Monaco, A.P. and Lupski, J.R. (2001) Denaturing high-performance liquid chromatography of the myotubularin-related 2 gene (MTMR2) in unrelated patients with Charcot–Marie–Tooth disease suggests a low frequency of mutation in inherited neuropathy. *Neurogenetics*, **3**, 107–109.
21. Berger, P., Bonneick, S., Willi, S., Wymann M. and Suter, U. (2002) Loss of phosphatase activity in myotubularin-related protein 2 is associated with Charcot–Marie–Tooth disease type 4B1. *Hum. Mol. Genet.*, **11**, 1569–1579.
22. Hunter, T. (1998) Anti-phosphatases take the stage. *Nat. Genet.*, **18**, 303–305.
23. Levivier, E., Goud, B., Souchet, M., Calmels, T.P., Mornon, J.P. and Callebaut, I. (2001) uDENN, DENN, and dDENN: indissociable domains in Rab and MAP kinases signaling pathways. *Biochem. Biophys. Res. Commun.*, **287**, 688–695.
24. Majidi, M., Gutkind, J.S. and Lichy, J.H. (2000) Deletion of the COOH terminus converts the ST5 p70 protein from an inhibitor of RAS signaling to an activator with transforming activity in NIH-3T3 cells. *J. Biol. Chem.*, **275**, 6560–6565.
25. Lathrop, G.M., Lalouel, J.M., Julier, C. and Ott, J. (1984) Strategies for multilocus linkage analysis in humans. *Proc. Natl Acad. Sci. USA*, **81**, 3443–3446.
26. Sobel, E. and Lange, K. (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing statistics. *Am. J. Hum. Genet.*, **58**, 1323–1337.
27. Kozak, M. (1989) Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol. Cell. Biol.*, **9**, 5073–5080.