Am. J. Hum. Genet. 74:1064–1073, 2004

Report

Identification of a Novel Gene (*HSN2*) Causing Hereditary Sensory and Autonomic Neuropathy Type II through the Study of Canadian Genetic Isolates

Ronald G. Lafrenière,[1,*,†] Marcia L. E. MacDonald,[4,*] Marie-Pierre Dubé,[1] Julie MacFarlane,[4] Mary O'Driscoll,[5] Bernard Brais,[2] Sébastien Meilleur,[1] Ryan R. Brinkman,[4] Owen Dadivas,[4] Terry Pape,[4] Christèle Platon,[1] Chris Radomski,[4] Jenni Risler,[4] Jay Thompson,[4] Ana-Maria Guerra-Escobio,[1] Gudarz Davar,[7] Xandra O. Breakefield,[8] Simon N. Pimstone,[4] Roger Green,[5,6] William Pryse-Phillips,[6] Y. Paul Goldberg,[4,9] H. Banfield Younghusband,[5,6] Michael R. Hayden,[4,9] Robin Sherrington,[4] Guy A. Rouleau,[1,3,†] and Mark E. Samuels[4]

[1]Xenon Genetics Research, [2]Département de Médecine de l'Université de Montréal et Centre de Recherche du Centre Hospitalier de l'Université de Montréal, and [3]Centre for Research in Neuroscience, McGill University, Montreal; [4]Xenon Genetics, Burnaby, British Columbia; [5]Discipline of Genetics and [6]Faculty of Medicine, Memorial University, St. John's, Newfoundland; [7]Departments of Neurology and Anesthesiology, Brigham and Women's Hospital, and [8]Departments of Neurology and Radiology, Massachusetts General Hospital and Neuroscience Program, Harvard Medical School, Boston; and [9]Department of Medical Genetics, University of British Columbia, and Center for Molecular Medicine and Therapeutics, Children and Women's Hospital, Vancouver

Hereditary sensory and autonomic neuropathy (HSAN) type II is an autosomal recessive disorder characterized by impairment of pain, temperature, and touch sensation owing to reduction or absence of peripheral sensory neurons. We identified two large pedigrees segregating the disorder in an isolated population living in Newfoundland and performed a 5-cM genome scan. Linkage analysis identified a locus mapping to 12p13.33 with a maximum LOD score of 8.4. Haplotype sharing defined a candidate interval of 1.06 Mb containing all or part of seven annotated genes, sequencing of which failed to detect causative mutations. Comparative genomics revealed a conserved ORF corresponding to a novel gene in which we found three different truncating mutations among five families including patients from rural Quebec and Nova Scotia. This gene, termed "HSN2," consists of a single exon located within intron 8 of the *PRKWNK1* gene and is transcribed from the same strand. The HSN2 protein may play a role in the development and/or maintenance of peripheral sensory neurons or their supporting Schwann cells.

The hereditary sensory and autonomic neuropathies (HSANs) comprise at minimum a group of five clinically heterogeneous disorders, characterized by variable sensory and autonomic dysfunction due to peripheral nerve degeneration (Dyck 1993; Axelrod 2002; Hilz 2002). HSANs are distinct from the more prevalent motor and sensory neuropathies (Charcot-Marie-Tooth disorders) because sensory involvement is primary. Moreover, the abnormalities of autonomic function and varying degrees of analgesia are specific to the HSANs (Axelrod 2002).

Mutations have previously been identified for autosomal dominant, adult-onset HSAN type I (MIM 162400) (Bejaoui et al. 2001; Dawkins et al. 2001); for autosomal recessive HSAN type III, also known as "familial dysautonomia/Riley Day syndrome" (MIM 223900) (Anderson et al. 2001; Slaugenhaupt et al. 2001); and for autosomal recessive HSAN type IV, also called "congenital insensitivity to pain with anhidrosis" (MIM 256800) (Mardy et al. 1999). HSAN type V is the least reported of the five disorders and remains somewhat uncertain in phenotypic description and molecular basis (Houlden et al. 2001; Toscano et al. 2002; Axelrod 2002; Hilz 2002; Einarsdottir et al. 2004). Recently, a

Received December 19, 2003; accepted for publication February 25, 2004; electronically published April 1, 2004.

Address for correspondence and reprints: Dr. Mark E. Samuels, Xenon Genetics Inc., Burnaby, British Columbia V5G 4W8, Canada. E-mail: msamuels@xenongenetics.com

* This work represents equal contributions by these authors.

† Present affiliation: Emerillon Therapeutics Inc., Montreal.

Portions of this research have been financially supported by Xenon Genetics and Xenon Genetics Research.

0002-9297/2004/7405-0026$15.00

rare subtype of HSAN type I (type IB) with cough and gastroesophageal reflux has been described and mapped to a novel locus (MIM 608088) (Kok et al. 2003).

Locus identification has not yet been described for the autosomal recessive HSAN type II (MIM 201300), also called "neurogenic acroosteolysis," "hereditary sensory radicular neuropathy," or "congenital sensory neuropathy." Clinical description of HSAN type II in the literature is not completely consistent; however, diagnosis can be made on the basis of natural history, a group of clinical findings, and pathology (Axelrod 2002). HSAN type II is generally defined by progressive degeneration of peripheral neurons, typically with onset in the first 2 decades, manifesting sensory reduction in the upper and lower limbs (Ogryzlo 1946; Murray 1973; Ohta et al. 1973; Kondo and Horikawa 1974). There have been reports of a nonprogressive, stable congenital form of HSAN type II with neonatal hypotonia and sensory impairment affecting the whole body rather than just the limbs, suggesting two different clinical subtypes of the type II disease (Ferriere et al. 1992). The worldwide prevalence of HSAN type II is low, but clustering of cases in eastern Canada has been reported (Murray 1973; Roddier et al. 2003).

HSAN type II in Newfoundland was first noted in the early 1900s. The original family members came from Dorset, United Kingdom, ~100 years earlier, as part of a mass migration of Protestant settlers from southwestern England and Roman Catholic settlers from southern Ireland (Rahman et al. 2003). The affected individuals ascertained in this study had the following clinical features: Beginning in early childhood, they experienced numbness in their hands and feet, aggravated by cold, together with reduced sensation to pain. They experienced loss of touch, pain, and temperature, with touch most severely affected. The loss was predominantly distal, extending from the elbows to the finger tips and from just above the knees down to the toes. This is often described as the "glove and sock distribution." Typically, the lower limbs were affected more severely than the upper limbs. Progression of the disorder varied within the family, however, involving the trunk in some patients. In some cases, secondary features of muscle atrophy, diminished tendon reflexes, ulcerations, and infections caused spontaneous amputation of digits and surgical amputation of lower limbs (fig. 1*a*). Patients did not express any overt autonomic dysfunction. Sweating and tearing were within normal range, and postural hypotension was not present. Mental development was normal. As in the other HSANs, there was absence of axon flare after intradermal histamine, indicative of defective nociceptive fibers. Biopsy revealed a severe loss of myelinated axons, some loss of nonmyelinated fibres in the sural nerve, and the absence of cutaneous sensory receptors and nerve fibers.

To identify the HSAN II locus, we expanded a consanguineous multigenerational family (F1) with eight affected members, living or deceased, from the Newfoundland sibship reported previously (Ogryzlo 1946), and we ascertained an additional family (F2) with two affected members (fig. 1*b*). Most of the patients live within a 160-km radius in a geographically isolated region of northeastern Newfoundland, suggesting the possibility of a founder effect. With institutional review board approval and informed consent, we collected DNA samples from 57 individuals, including one deceased but diagnosed affected individual and seven living persons diagnosed with HSAN type II on the basis of quantitative sensory and autonomic testing and, when necessary, sensory and motor nerve conduction studies. In families F1 and F2, the mode of inheritance appeared autosomal recessive, as reported for other families with HSAN type II (Ohta et al. 1973; Kondo and Horikawa 1974). Subsequently, we obtained samples from a group of French Canadians consisting of two affected sisters (F3-301 and F3-302) and one affected individual (F4-301) with no close family ties to the affected F3 sisters from a larger cohort of HSAN type II families in French Canada (Roddier et al. 2003). Finally, DNA was obtained from two affected sisters (F5-301 and F5-302) who were diagnosed with HSAN type II in Nova Scotia and whose parents were consanguineous but of uncertain ethnic origin (Murray 1973; Davar et al. 1996).

We performed a complete 5-cM genome scan of Newfoundland families F1 and F2 using 763 autosomal and 48 X-chromosomal markers. Two-point linkage analysis identified 10 markers in eight chromosomal regions with the sum of two-point LOD scores >2.0 (data not shown). One such marker (D12S352) mapping to the telomere of 12p showed extensive though incomplete allele sharing in the affected individuals (fig. 2*a*), with an additive 2-point LOD score of 2.64 at zero recombination, suggesting linkage to the 12p13 region. Additional microsatellite markers in this region revealed that all seven living affected individuals in both families were homozygous for a common haplotype spanning 1.2 Mb with a highly significant LOD score (table 1, fig. 2*a*). The additional markers defined recombination events consistent with the incomplete allele sharing observed in the genome scan. Subsequently, we genotyped affected individuals from the two French Canadian families (F3 and F4). Haplotype data were consistent with linkage to this chromosomal region and suggested that there may be two distinct mutations in the French Canadian population (fig. 2*a*). A recombination event in one of the sisters in F3 reduced the size of the candidate region to between 1.040 and 1.057 Mb (fig. 2).

A total of seven annotated transcripts were identified in the candidate region from the completed human genomic assembly (fig. 2*b*). These included *PRKWNK1*,

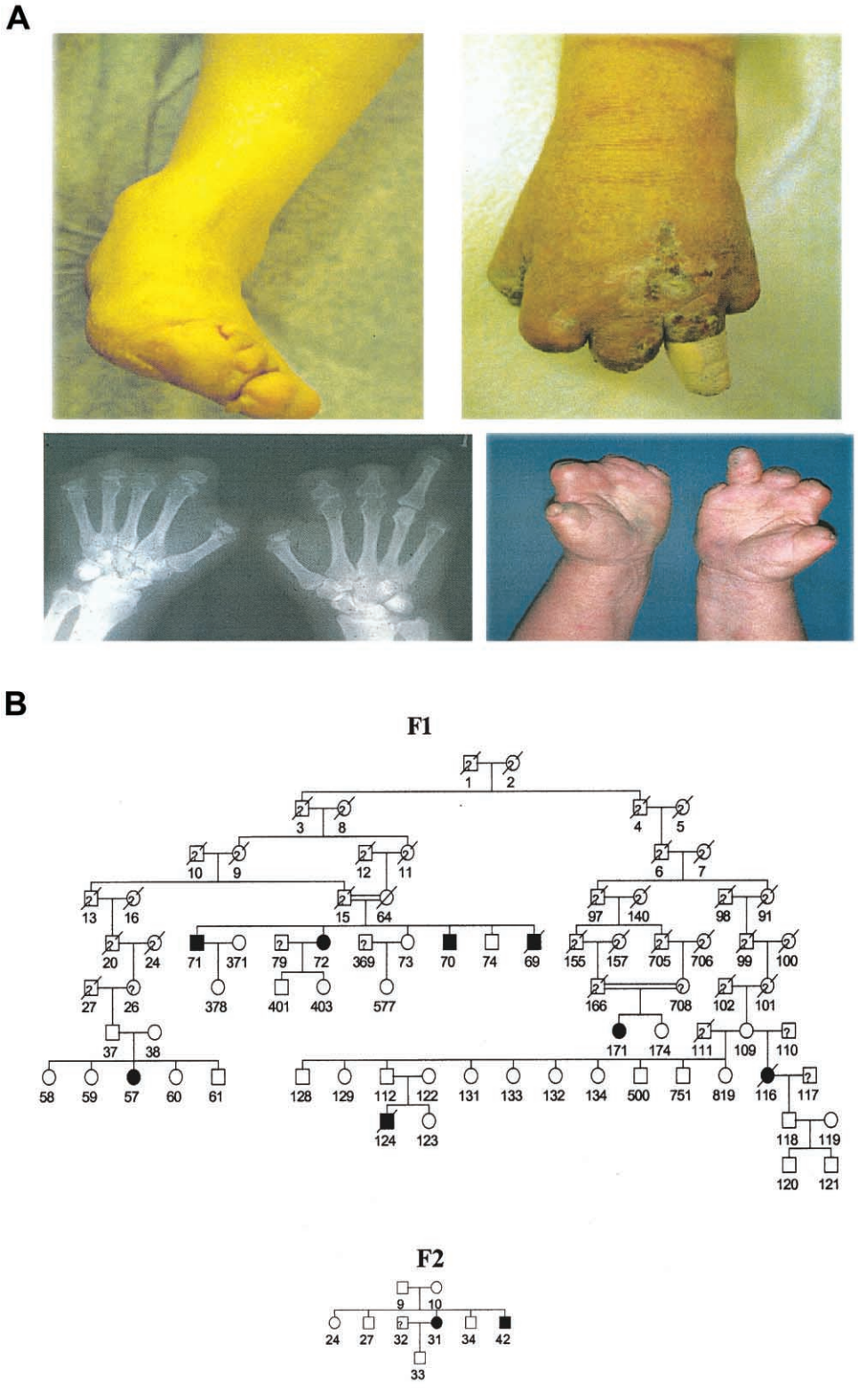

Figure 1 Phenotype of HSAN type II and pedigrees for large Newfoundland families described in this study. *a*, Examples of clinical presentations. Patients in these families suffer from loss of touch and pain sensation leading to ulceration, infection, muscle atrophy, and amputation of digits. *b*, Partial pedigrees of families F1 and F2. Squares and circles represent males and females, respectively. Blackened symbols indicate individuals with HSAN type II. Unblackened symbols indicate individuals normal by clinical examination, and symbols with a question mark indicate individuals who have not been examined. Diagonal slashes indicate deceased individuals. All examined individuals were included in the genome scan or the subsequent fine mapping.

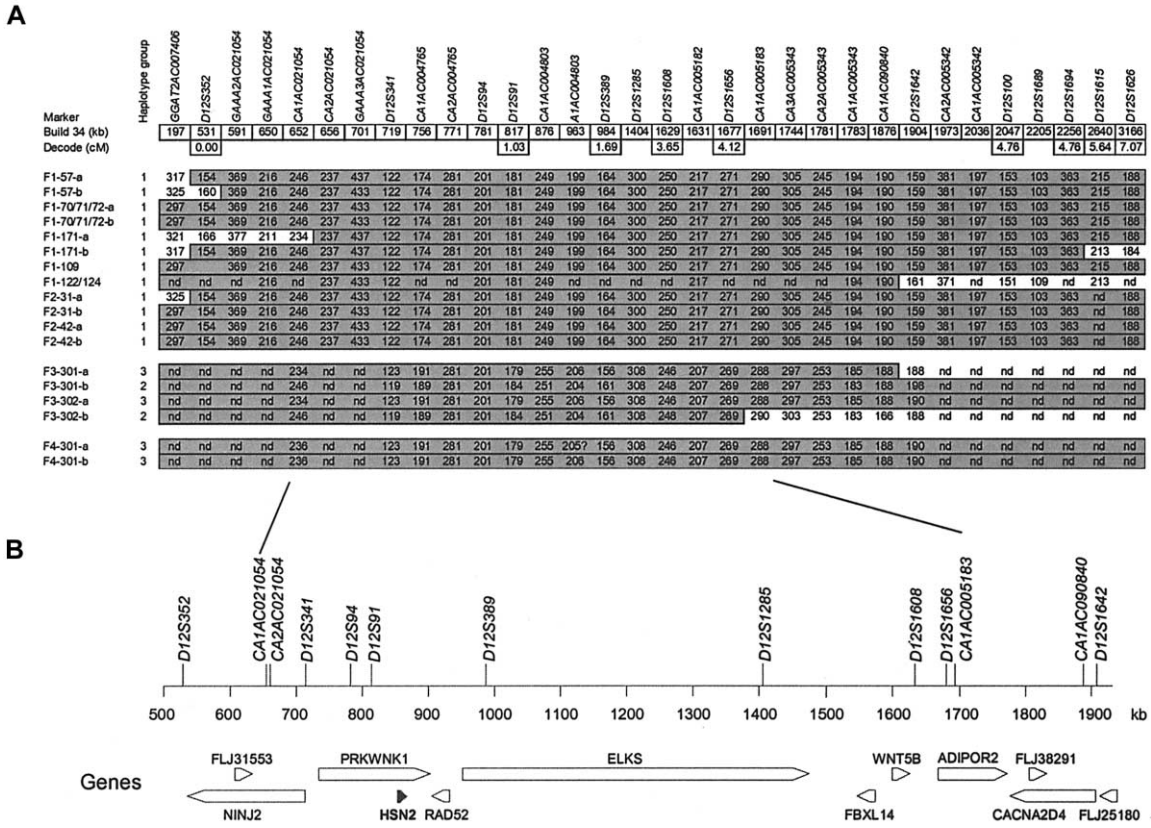

Figure 2 Haplotype analysis and candidate region. *a,* Haplotypes were constructed for all genotyped individuals in families F1 and F2, but here only affected haplotypes are shown. GENEHUNTER v2.1_r2 beta (Kruglyak et al. 1996; Kruglyak and Lander 1998) was used to construct haplotypes on pedigree sections, which were then manually combined. Haplotypes are designated with pedigree (F) number, individual number, and chromosome (a or b) number. The three different haplotype groups bearing causal mutations are numbered 1–3. Allele sizes are given in bp. Markers are indicated at the top, with their physical distances from the telomere given in kb on the basis of the July 2003 build 34 human genome assembly and their genetic positions, according to the deCODE map. Portions of haplotypes encompassing the *HSN2* locus are boxed. Haplotypes for F1 and F2 are phased, except for F1–122/124, which was inferred from genotypes in parent F1-122 and deceased offspring F1-124. Phasing of F3-301 and F3-302 was inferred on the basis of allele sharing with a region of homozygosity in F4-301. Recombination breakpoints in the shared haplotype in individuals F1-171 and F1-124 placed the HSAN II locus between markers CA1AC0021054 and D12S1642. Affected French Canadian samples were subsequently genotyped for markers in this region. A recombination event in one of the sisters in F3 positioned the centromeric boundary at marker CA1AC005183. *b,* Physical map of the HSAN II candidate region, showing markers used in the genetic analysis and candidate genes represented as unblackened arrows (derived from UCSC genome annotation). The arrows indicate the direction of transcription for each gene. The distances are given in kb from the telomere on the basis of the July 2003 build 34 human genome assembly. The *HSN2* gene is indicated by a blackened arrow. To identify unannotated conserved ORFs, the human genomic sequence representing the entire HSAN II candidate interval (~1.0 Mb) was downloaded with case toggled to highlight the mouse translated BLAT track (which represents regions showing significant protein homology between human and mouse). These sequences were assembled into a contig, to which all exons of previously identified candidate genes were added. All sequences were manipulated using DNASTAR software. The contig was scanned for novel conserved fragments (>80% conservation over >100 bp). This identified 64 novel fragments, which were then tested for functional homology using (1) BLASTn against the nr database, (2) BLASTn against the dbEST, and (3) BLASTx against the nr database available from the NCBI Web site. The *HSN2* gene was, by a substantial margin, the most highly conserved and the longest ORF among these fragments. It should be noted that various gene-prediction algorithms—such as ECgene, Geneid, SGP, and Twinscan—also predict the *HSN2* ORF, in whole or in part, either as a separate exon or as a potential alternative exon of the *PRKWNK1* gene (for which there is no evidence among mammalian ESTs).

RAD52, *ELKS*, *FBXL14*, *WNT5B*, and portions of *ADIPOR2* and *NINJ2*. The syntenic region in the mouse genome is ~800 kb in size, located on chromosome 6, and conserves this gene order. All seven annotated transcripts in the candidate region were sequenced in samples from affected patients, but no causative mutations were observed. A more exhaustive search for unannotated genes was therefore conducted on the basis of conserved homologies between the human and mouse genome assemblies, together with a requirement for an extended ORF of >100 amino acids. This search identified a novel unannotated yet well-conserved 434–amino acid ORF

Table 1

Linkage Analysis in Region 12p

	Position		LOD for Family and Recombination Fraction (θ)														
			F1					F2					F1 and F2				
Marker	NCBI Build 34 (kb)	deCODE Map (cM)	.00	.01	.05	.10	Max	.00	.01	.05	.10	Max	.00	.01	.05	.10	Max
GGAT2AC007406	197		.00	.33	.88	1.06	1.06	−4.14	−1.04	−.41	−.18	.01	−4.14	−.71	.48	.88	.90
D12S352	531	.00	3.40	3.49	3.45	3.14	3.49	.84	.82	.72	.59	.84	4.24	4.31	4.16	3.73	4.31
GAAA2AC021054	591		2.84	2.92	2.85	2.52	2.92	.37	.37	.33	.28	.37	3.22	3.29	3.18	2.80	3.29
GAAA1AC021054	650		4.69	4.73	4.57	4.14	4.73	.00	.00	.00	.00	.00	4.69	4.73	4.57	4.14	4.73
CA1AC021054	652		2.94	3.02	2.95	2.63	3.02	.37	.37	.33	.28	.37	3.32	3.39	3.28	2.91	3.39
CA2AC021054	656		5.10	5.07	4.81	4.36	5.10	.97	.94	.82	.67	.97	6.07	6.01	5.63	5.03	6.07
GAAA3AC021054	701		2.54	2.56	2.54	2.37	2.56	−.64	−.41	−.07	.05	.09	1.91	2.15	2.46	2.42	2.46
D12S341	719		5.10	4.98	4.48	3.86	5.10	.08	.08	.07	.06	.08	5.18	5.06	4.55	3.92	5.18
CA1AC004765	756		5.50	5.37	4.83	4.16	5.50	.11	.13	.17	.18	.18	5.61	5.50	5.00	4.34	5.61
CA2AC004765	771		.95	.93	.82	.70	.95	.00	.00	.00	.00	.00	.95	.92	.82	.70	.95
D12S94	781		7.26	7.12	6.54	5.80	7.26	.00	.00	.00	.00	.00	7.26	7.12	6.54	5.80	7.26
D12S91	817	1.03	5.08	4.96	4.45	3.81	5.08	.37	.37	.33	.28	.37	5.46	5.32	4.78	4.09	5.46
CA1AC004803	876		4.26	4.14	3.69	3.12	4.26	.97	.94	.82	.67	.97	5.22	5.08	4.50	3.79	5.22
D12S389	984	1.69	3.88	3.80	3.46	2.99	3.88	.37	.37	.33	.28	.37	4.25	4.17	3.78	3.27	4.25
A1AC004803	963		.66	.63	.53	.41	.66	.37	.37	.33	.28	.37	1.03	.99	.86	.69	1.03
D12S1285	1404		6.54	6.41	5.88	5.20	6.54	.97	.94	.82	.67	.97	7.51	7.35	6.70	5.86	7.51
D12S1608	1629	3.65	4.80	4.68	4.19	3.58	4.80	.97	.94	.82	.67	.97	5.77	5.61	5.00	4.24	5.77
CA1AC005182	1631		6.56	6.43	5.88	5.17	6.56	.97	.94	.82	.67	.97	7.53	7.36	6.69	5.84	7.53
D12S1656	1677	4.12	3.12	3.02	2.66	2.21	3.12	.97	.94	.82	.67	.97	4.08	3.96	3.47	2.88	4.08
CA1AC005183	1691		2.14	2.07	1.82	1.51	2.14	.97	.94	.82	.67	.97	3.10	3.01	2.64	2.18	3.10
CA3AC005343	1744		3.73	3.62	3.20	2.68	3.73	.97	.94	.82	.67	.97	4.69	4.56	4.02	3.35	4.69
CA2AC005343	1781		5.98	5.97	5.72	5.21	5.98	.97	.94	.82	.67	.97	6.95	6.91	6.53	5.87	6.95
CA1AC005343	1783		7.48	7.34	6.77	6.03	7.48	.97	.94	.82	.67	.97	8.44	8.27	7.58	6.70	8.44
CA1AC090840	1876		6.82	6.68	6.12	5.41	6.82	.97	.94	.82	.67	.97	7.78	7.62	6.94	6.07	7.78
D12S1642	1904		1.52	1.48	1.31	1.10	1.52	.08	.08	.07	.06	.08	1.60	1.56	1.38	1.15	1.60
CA2AC005342	1973		6.53	6.40	5.86	5.17	6.53	.97	.94	.82	.67	.97	7.50	7.33	6.67	5.83	7.50
CA1AC005342	2036		5.49	5.44	5.13	4.62	5.49	.97	.94	.82	.67	.97	6.46	6.38	5.95	5.29	6.46
D12S100	2047	4.76	6.50	6.37	5.83	5.14	6.50	.97	.94	.82	.67	.97	7.47	7.30	6.65	5.81	7.47
D12S1689	2205		2.89	2.80	2.46	2.05	2.89	.08	.08	.07	.06	.08	2.97	2.89	2.54	2.11	2.97
D12S1694	2256	4.76	7.46	7.32	6.73	5.98	7.46	.97	.94	.82	.67	.97	8.43	8.25	7.55	6.65	8.43
D12S1615	2640	5.64	3.67	3.87	4.03	3.81	4.03	.00	.00	.00	.00	.00	3.67	3.87	4.03	3.81	4.03
D12S1626	3166	7.07	2.88	2.97	2.93	2.62	2.97	.97	.94	.82	.67	.97	3.84	3.91	3.74	3.29	3.91
D12S1652	3583	9.04	3.51	3.58	3.46	3.07	3.58	.97	.94	.82	.67	.97	4.48	4.52	4.28	3.73	4.52
D12S1725	4316	13.14	−1.14	−.76	−.11	.16	.23	−4.35	−1.16	−.51	−.27	.00	−5.49	−1.92	−.62	−.11	.14
D12S1624	4581		−1.56	−1.22	−.50	−.11	.13	−4.50	−1.06	−.46	−.26	.00	−6.06	−2.28	−.95	−.37	.05
D12S314	4839	13.96	−1.24	−.82	−.08	.25	.37	−4.50	−1.06	−.46	−.26	.00	−5.74	−1.88	−.54	−.01	.25
D12S93	5201	15.00	−1.22	−.91	−.39	−.18	.00	−.21	−.20	−.16	−.10	.00	−1.43	−1.11	−.55	−.28	.00
D12S99	5435	15.20	−.83	−.51	.05	.25	.25	−5.54	−2.85	−1.49	−.93	.00	−6.37	−3.36	−1.43	−.67	.00

Note.—Samples from 15 family members from F1 and 7 family members from F2, including a total of seven living affected individuals from Newfoundland, were genome scanned. Genomic DNA was extracted from blood samples using a standard salt-extraction method. DNA was also obtained from tissue samples of one deceased individual in F1, previously diagnosed with HSAN type II. We carried out a 5-cM whole-genome screen using the LMS2 HD-5 microsatellite screening set (Applied Biosystems) and Prism 3100 and 3700 Genetic Analyzers running GeneMapper software (Applied Biosystems). For fine mapping, we analyzed an additional 16 Genethon markers and 1 CHLC marker (derived from NCBI, GDB, and Marshfield databases), and designed 18 novel markers containing short tandem repeats on the basis of publicly available genomic sequences. Genome-scan markers are indicated in boldface italics in the table. Primer sequences for all markers are indicated in table A1 (online only); map locations are described using the deCODE genetic maps (Kong et al. 2002) and physical locations are based on the July 2003 build 34 genome assembly. Mendelian inheritance of alleles was verified using the PedCheck program (O'Connell and Weeks 1998). We performed pairwise linkage analysis on the F1 and F2 pedigrees using MLINK from the FASTLINK v4.1p program (Cottingham et al. 1993; Schaffer et al. 1994) and allowing for known pedigree consanguinity loops. We calculated maximum pairwise cumulative LOD scores as the maximum LOD over tested θ of the sum of θ-specific LODs for the F1 and F2 pedigrees. We chose to not carry out multipoint linkage since pairwise results were sufficiently convincing and we wished to retain the consanguinity information of F1 for linkage. We set equal marker allele frequencies for the genomewide linkage scan. For the follow-up linkage analysis at 12p13, allele frequencies were estimated using 16 chromosomes from 8 unrelated individuals from the Newfoundland population, as well as 8 untransmitted chromosomes from F3 and F4 (F3-9, F3-10, F4-37, F4-38, F4-112, F4-122, F4-371). Frequencies were estimated using allele counting. Unobserved alleles were set to 0.02 frequency, and all frequencies were proportionally transformed to sum to 1. Two consanguineous loops of F1 were broken at individuals 15 and 708. We used an autosomal recessive model with a disease allele frequency of 0.007, penetrance of 0.975, and phenocopy rate of 0.000003, as calculated from the disease prevalence in the population. Linkage tests using equal allele frequencies were carried out in parallel for the fine-mapping data presented in the table. We believe that using the estimated allele frequencies from the Newfoundland population does provide more accurate LOD score estimates; however, LOD scores calculated using equal allele frequencies provided significant results as well, with a maximum LOD of 5.88 at D12S1285.

located within intron 8 of the *PRKWNK1* gene (fig. 2*b*). Sequencing of samples containing the three linked haplotypes revealed that this ORF harbored three different truncating mutations in the affected patients (figs. 3, 4*a*). One mutation was found in homozygous state in the affected Newfoundland subjects. This mutation was found only in Newfoundland and cosegregated perfectly with the presumptive affected haplotype in families F1 and F2. The other two mutations found in the French Canadian families were in either the homozygous or the compound heterozygous state, consistent with the haplotypes in F3 and F4. Subsequent work has shown that, in all identified French Canadian patients, mutations cosegregate in either the homozygous or the compound heterozygous state, consistent with haplotype analysis (B.B., unpublished observations). One of these mutations was also homozygous in both Nova Scotia F5 samples.

None of these three mutations were found in 192 normal chromosomes of European descent, in 180 Coriell

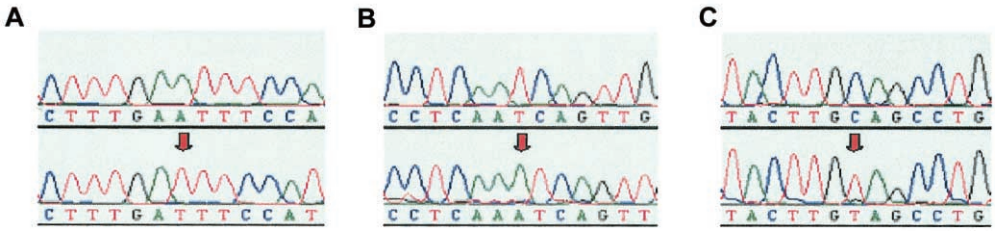

Figure 3 Mutations in the *HSN2* gene in affected individuals. To screen the *HSN2* ORF for mutations, three separate amplicons were designed (for primers, see appendix A [online only]) using PrimerSelect (DNASTAR) and were purchased from BioCorp. Amplified products were sequenced at the Montreal Genome Centre sequencing facility. Sequence traces were aligned using SeqManII (DNASTAR) and were inspected visually for mutations. Each panel displays a sequencing trace from a normal control (*above*) and from an affected individual (*below*) according to the following: *a,* Homozygous c.594delA mutation in patient F1-70 from Newfoundland, causing a frameshift in codon 198 leading to premature truncation to a 206-aa peptide. This mutation cosegregated perfectly with the disease in the F1 and F2 families from Newfoundland as predicted from the haplotype analysis. *b,* Homozygous c.918–919insA mutation in patient F5-301 from Nova Scotia, causing a frameshift in codon 307 and truncation to a 318-aa peptide. Both Nova Scotia siblings shared this mutation, which was also heterozygous in the French Canadian F3 samples. *c,* Homozygous c.943C→T nonsense mutation in French Canadian patient F4-301, changing codon 315 (CAG, encoding glutamine) to a TAG stop codon, and truncating the protein to 314 aa. This mutation was also heterozygous in the F3 samples. All sequence traces are from the forward strand. To genotype each of the three mutations in additional family members and population controls, we used PCR-RFLP or capillary electrophoresis analysis (details of mutation detection sequencing and genotyping in appendix A [online only]).

control chromosomes of mixed ethnicity, or in any public sequence databases. The clustering of three different disrupting mutations within a small conserved ORF in five formally unrelated HSAN type II families strongly supports the conclusion that these mutations are the direct cause of the HSAN type II phenotype. We therefore propose that the "HSN2" gene encodes this novel ORF; this nomenclature has been accepted by HUGO.

To confirm the comparative genomic analysis, additional *HSN2* gene orthologs from rat and zebrafish, as well as mouse, were identified by a BLAST search of their respective genomic assemblies (fig. 4*a*). Each *HSN2* ortholog mapped within the conserved intron of the *PRKWNK1* ortholog and was encoded on the same strand as *PRKWNK1*. To verify the presumptive structure of the *HSN2* gene, we closely examined ESTs overlapping or mapping near the ORF. As such, a number of human cDNA clones, derived from a variety of adult and fetal tissues, were identified from genome databases. In one of these cDNA clones containing a portion of the ORF (accession number BF695311, corresponding to IMAGE:4244848 cloned from skeletal muscle, purchased from Research Genetics and resequenced by us), a long polyA tail was detected just downstream of an AATAAA polyA addition signal; however, the clone was incomplete at the 5′ end. A placentally derived pig-cDNA clone with similarity at the protein level (corresponding to the partial sequence in accession number BI399422, originally identified through the Gene Discovery Program for Functional Genomics in Pig Reproduction) was obtained from Open Biosystems, and the insert was sequenced. This cDNA clone encodes the entire predicted HSN2 peptide (fig. 4*a*), plus 475 bp of 5′-UTR, 561 bp of 3′-UTR, and a polyA tail at a site aligning with the human polyA site (fig. 4*b*). Thus, the pig cDNA may represent a nearly full-length *HSN2* transcript. *HSN2* transcripts could not be detected by northern blotting using an adult human tissue panel (Clontech 12-lane human 636818 including brain, heart, skeletal muscle, colon, thymus, spleen, kidney, liver, small intestine, placenta, lung, and peripheral blood leukocyte), suggesting that the gene may be expressed at very low levels. RT-PCR experiments using cDNA from human fibroblast, thyroid, liver, testes, small intestine, adrenal gland, brain, and dorsal root ganglion were similarly inconclusive; these are complicated by the absence of a splice junction and thus potential contamination with either trace genomic DNA or incompletely processed *PRKWNK1* transcripts. Further analysis will be required to define the exact transcript structure and expression pattern of *HSN2*. The low level and distribution of expression suggested by ESTs is consistent with a gene specifically expressed from peripheral sensory neurons or supporting cells, which are distributed throughout many organs and tissues but constitute only a very small percentage of any given tissue's mass.

These observations suggest that the *HSN2* gene spans ~3 kb and consists of a single exon nested within an intron of the *PRKWNK1* gene and transcribed from the same strand. The structure of the *HSN2* gene is thus unusual and may be the first example of a causative mutation in a human gene mapping within the intron of another gene. Single exon genes are somewhat atypical but well-documented; moreover, the *HSN2* sequence is unique in each genome assembly examined, which indicates that it is not a pseudogene. The *PRKWNK1*

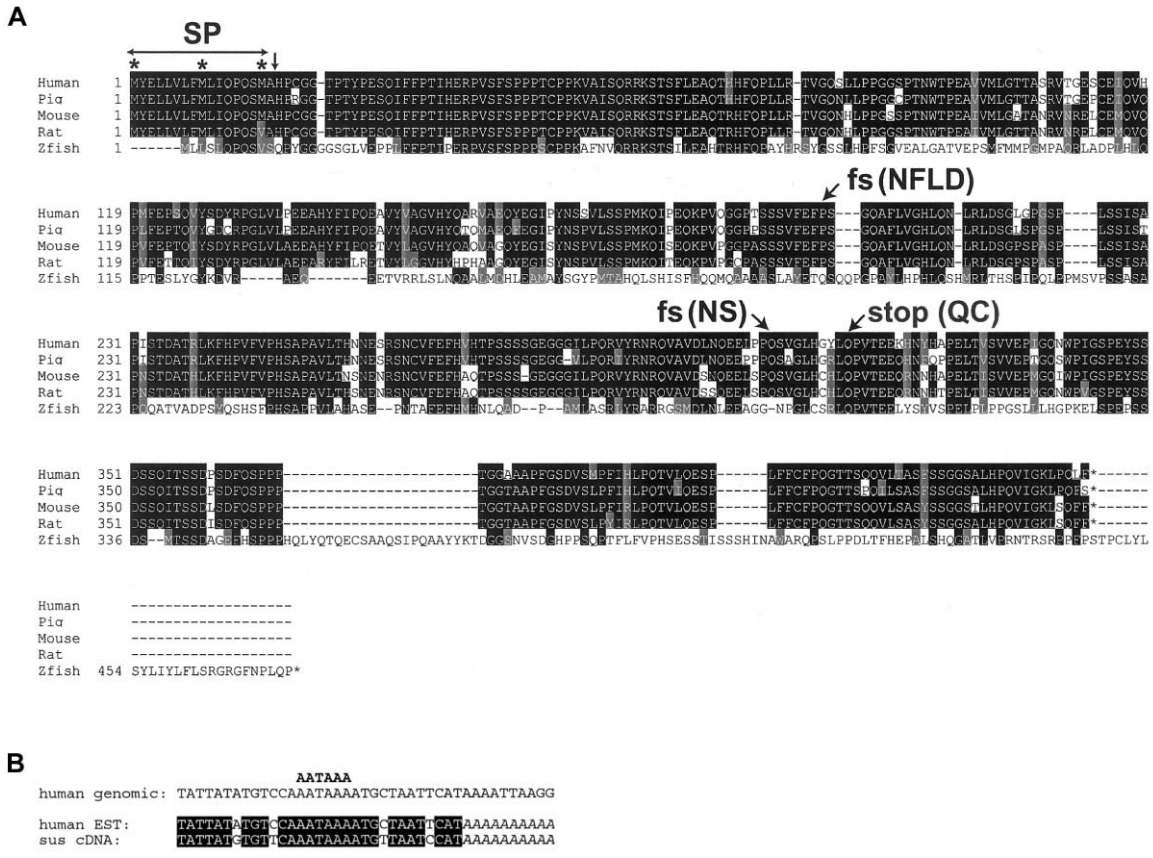

Figure 4 Alignment of predicted HSN2 peptide sequences for vertebrate orthologs. *a,* The conserved *HSN2* ORF was identified from the genomic assemblies of human, mouse, rat, and zebrafish (Zfish) and from the pig cDNA clone. BLAST searches were done through the NCBI Web site. Genomic sequences for *HSN2* orthologs were identified within GenBank sequence files AC004765 (human), AC106932 and AC106348 (rat), AC113092 (mouse), and BX321885 (zebrafish). Prosite searches were done through the EBI server. For the predicted peptide sequence alignment, sequences were aligned using Clustal 1.8, available from the BCM Search Launcher Web site. Aligned sequences then were shaded using BOXSHADE. Functional bioinformatics analysis was performed using SignalP. Conserved residues are shaded. Residues are numbered from the most upstream start ATG. Asterisks (*) above N-terminal residues indicate potential initiating methionines; asterisks at the C-termini indicate the stop codons. The signal peptide and predicted cleavage residue are indicated at the N-terminus. Positions of causative mutations in human populations are indicated. *b,* Alignment of 3′ ends of resequenced human and pig cDNA clones with the human genomic sequence, showing conservation of polyadenylation sites. The putative polyadenylation signal is shown above the aligned sequences.

gene is expressed in polarized epithelia in a number of tissues of patients with hypertension, and dominant mutations are thought to disrupt the regulation of epithelial Cl^- flux, leading to pseudohypoaldosteronism type IIC (MIM 605232) (Wilson et al. 2001; Choate et al. 2003). There is no overlap in observed phenotypes between HSAN type II and pseudohypoaldosteronism, supporting the notion that transcription of these two genes may be differentially regulated.

Bioinformatic analysis of the predicted HSN2 amino acid sequence identifies a potential signal peptide motif at the amino terminus of the mammalian proteins (fig. 4*a*). However, there are three methionines closely spaced at the N-terminus of the ORF, and the occurrence of this signal peptide thus depends on which of these is the true initiating methionine. If that is the case, then HSN2 may conceivably be a secreted protein and possibly a novel neurotrophic factor. The *NTRK1* gene underlying HSAN type IV is a known receptor of a nerve growth factor; it is conceivable that this receptor may also bind a secreted HSN2 peptide, although there is no obvious sequence similarity between HSN2 and any known protein.

Identification of the hereditary basis of HSAN type II represents a key advance in the diagnosis and management of this and other peripheral neuropathies. *HSN2* represents the fifth identified gene that when mutated results in a sensory and autonomic neuropathy phenotype. There is no obvious unifying theme among all these genes, which include a biosynthetic enzyme in a lipid pathway (HSAN type I/SPTLC1) (Bejaoui et al. 2001; Dawkins et al. 2001), a presumptive transcriptional reg-

ulator (HSAN type III/IKBKAP) (Anderson et al. 2001; Slaugenhaupt et al. 2001), a nerve growth factor receptor (HSAN type IV/NTRK1) (Mardy et al. 1999), a nerve growth factor (NGFB) (Einarsdottir et al. 2004), and a novel protein (HSAN type II/HSN2). If HSN2 is a novel growth factor, then, by virtue of its association with peripheral neuropathy, it may conceivably act via the *NTRK1* gene, and IKBKAP could potentially play a role in a subsequent signal transduction pathway. There are numerous clinical differences among these disorders, however, as well as potential differences in neuropathology. Further work will be required to resolve this question and to elucidate the role of these genes in neuronal development and function.

Given that the HSAN type II phenotype is characterized by absence of pain sensations owing to loss of fibers in the peripheral nervous system (Murray 1973; Dyck 1993), the HSN2 protein may play a role in the development or maintenance of peripheral sensory neurons or their supporting cells. The pathology of HSAN type II resembles in many respects that of diabetic neuropathy. Affecting as many as 50% of diabetic patients, initial symptoms of peripheral diabetic neuropathy (tingling, burning, and numbness in feet) mimic the presenting features of HSAN type II. In addition, sural nerve biopsy shows, in both diseases, a reduction in myelinated and unmyelinated nerve fibers (Llewelyn et al. 1991; Zorrilla Hernandez et al. 1994; Yasuda et al. 2003). It is possible that a therapeutic targeted to upregulate *HSN2* or a protein therapeutic derived from HSN2 could be used to prevent the nerve-degeneration features of diabetic neuropathy, which currently has no specific treatment modality. In addition to diabetic neuropathy, other neuropathies that could potentially benefit from an *HSN2* targeted therapeutic include HIV- and Hepatitis C-related neuropathies (Polydefkis et al. 2002; Luciano et al. 2003) and other less prevalent but significant metabolic and mitochondrial disorders.

In the literature, there are additional reports of phenotypes resembling HSAN type II (Ferriere et al. 1992; Basu et al. 2002; Krishna Kumar et al. 2002; Sanvito et al. 2003), and it is now feasible to determine whether those conditions are because of mutations in the *HSN2* gene or the result of additional genetic determinants. Furthermore, with the reported higher incidence of HSAN type II in French Canada, Nova Scotia, and Newfoundland, there is potentially a high carrier rate in parts of eastern Canada. Because of the severe disability of the disease and the early age at onset, carrier testing of *HSN2* and genetic counseling is now a feasible option for at-risk relatives of affected individuals. It may also be possible to examine whether carriers are at increased risk of neuropathic complications secondary to other diseases, such as diabetes.

Acknowledgments

The authors thank the patients with HSAN type II and their families for participating in the study, D. Jewer (Faculty of Medicine, Memorial University) for collection and clinical diagnosis, Carole Doré and Pierre Lepage for sequencing at the Montreal Genome Centre Sequencing facility, and Katel Roddier for technical assistance. M.R.H. holds a Canada Research Chair in Human Genetics and is a University Killam Professor. H.B.Y., R.G., and W.P.P. received support from the Medical Research Foundation, Memorial University. G.A.R is supported by the Canadian Institutes of Health Research and the Fonds de Recherche en Sante du Quebec. B.B. is supported by Démogénique and Génétique communautaire axes of the RMGA, the ECOGENE-21 project, and the l'Association de la neuropathie sensorielle et autonomique héréditaire de type II, and he is a chercheur-boursier of the FRSQ. X.O.B. is supported by NINDS NS24279.

Electronic-Database Information

Accession numbers and URLs for data presented herein are as follows:

BCM, http://searchlauncher.bcm.tmc.edu/multi-align/multi-align.html
BOXSHADE, http://www.ch.embnet.org/software/BOX_form.html
EBI, http://www.ebi.ac.uk/ppsearch/
Gene Discovery Program for Functional Genomics in Pig Reproduction, http://pigest.genome.iastate.edu/
HUGO Gene Nomenclature Committee, http://www.gene.ucl.ac.uk/nomenclature/
GDB, http://www.gdb.org/
Marshfield, http://research.marshfieldclinic.org/genetics/
NCBI and Genbank, http://www.ncbi.nlm.nih.gov/ (Accession numbers for *HSN2* sequences in human, mouse, rat, and pig genes are BK004108, BK004107, BK004106, and AY517853, respectively.)
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=OMIM
SignalP, http://www.cbs.dtu.dk/services/SignalP/
UCSC, http://genome.ucsc.edu/

References

Anderson SL, Coli R, Daly IW, Kichula EA, Rork MJ, Volpi SA, Ekstein J, Rubin BY (2001) Familial dysautonomia is caused by mutations of the IKAP gene. Am J Hum Genet 68:753–758

Axelrod FB (2002) Hereditary sensory and autonomic neuropathies: familial dysautonomia and other HSANs. Clin Auton Res Suppl 12:2–14

Basu S, Paul DK, Basu S (2002) Four siblings with type II hereditary sensory and autonomic neuropathy. Indian Pediatr 39:870–874

Bejaoui K, Wu C, Scheffler MD, Haan G, Ashby P, Wu L, de Jong P, Brown RH Jr. (2001) SPTLC1 is mutated in hereditary sensory neuropathy, type I. Nat Genet 27: 261–262

Choate KA, Kahle KT, Wilson FH, Nelson-Williams C, Lifton RP (2003) WNK1, a kinase mutated in inherited hypertension with hyperkalemia, localizes to diverse Cl- -transporting epithelia. Proc Natl Acad Sci USA 100:663–668

Cottingham RW Jr, Idury RM, Schaffer AA (1993) Faster sequential genetic linkage computations. Am J Hum Genet 53: 252–263

Davar G, Shalish C, Blumenfeld A, Breakfield X (1996) Exclusion of p75NGFR and other candidate genes in a family with hereditary sensory neuropathy type II. Pain 67:135–139

Dawkins JL, Hulme DJ, Brahmbhatt SB, Auer-Grumbach M, Nicholson GA (2001) Mutations in SPTLC1, encoding serine palmitoyltransferase, long chain base subunit-1, cause hereditary sensory neuropathy type I. Nat Genet 27:309–312

Dyck P (1993) Neuronal atrophy and degeneration predominantly affecting peripheral sensory and autonomic neuron. In: Dyck P, Thomas P, Griffin J, Low P, Poduslo J (eds) Peripheral neuropathies. W.B. Saunders, Philadelphia, Pennsylvania, pp 1065–1093

Einarsdottir E, Carlsson A, Minde J, Toolanen G, Svensson O, Solders G, Holmgren G, Holmberg D, Holmberg M (2004) A mutation in the nerve growth factor beta gene (NGFB) causes loss of pain perception. Hum Mol Genet 13: 799–805

Ferriere G, Guzzetta F, Kulakowski S, Evrard P (1992) Nonprogressive type II hereditary sensory autonomic neuropathy: a homogeneous clinicopathologic entity. J Child Neurol 7:364–370

Hilz MJ (2002) Assessment and evaluation of hereditary sensory and autonomic neuropathies with autonomic and neurophysiological examinations. Clin Auton Res Suppl 12:33–43

Houlden H, King RH, Hashemi-Nejad A, Wood NW, Mathias CJ, Reilly M, Thomas PK (2001) A novel TRK A (NTRK1) mutation associated with hereditary sensory and autonomic neuropathy type V. Ann Neurol 49:521–525

Kok C, Kennerson ML, Spring PJ, Ing AJ, Pollard JD, Nicholson GA (2003) A locus for hereditary sensory neuropathy with cough and gastroesophageal reflux on chromosome 3p22-p24. Am J Hum Genet 73:632–637

Kondo K, Horikawa Y (1974) Genetic heterogeneity of hereditary sensory neuropathy. Arch Neurol 30:336–337

Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigrun Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K (2002) A high resolution recombination map of the human genome. Nat Genet 31:241–247

Krishna Kumar R, Shashi Kiran N, Subba Reddy V (2002) Congenital insensitivity to pain (hereditary sensory and autonomic neuropathy) HSAN: a report of two cases. J Indian Soc Pedod Prev Dent 20:51–53

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1347–1363

Kruglyak L, Lander ES (1998) Faster multipoint linkage analysis using Fourier transforms. J Comput Biol 5:1–7

Llewelyn JG, Gilbey SG, Thomas PK, King RH, Muddle JR, Watkins PJ (1991) Sural nerve morphometry in diabetic autonomic and painful sensory neuropathy: a clinicopathological study. Brain 114:867–892

Luciano CA, Pardo CA, McArthur JC (2003) Recent developments in the HIV neuropathies. Curr Opin Neurol 16: 403–409

Mardy S, Miura Y, Endo F, Matsuda I, Sztriha L, Frossard P, Moosa A, Ismail EA, Macaya A, Andria G, Toscano E, Gibson W, Graham GE, Indo Y (1999) Congenital insensitivity to pain with anhidrosis: novel mutations in the TRKA (NTRK1) gene encoding a high-affinity receptor for nerve growth factor. Am J Hum Genet 64:1570–1579

Murray TJ (1973) Congenital sensory neuropathy. Brain 96: 387–394

O'Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63:259–266

Ogryzlo M (1946) A familial peripheral neuropathy of unknown etiology, resembling Morvan's disease. Can Med Assoc J 54:547–553

Ohta M, Ellefson RD, Lambert EH, Dyck PJ (1973) Hereditary sensory neuropathy, type II. Clinical, electrophysiologic, histologic, and biochemical studies of a Quebec kinship. Arch Neurol 29:23–37

Polydefkis M, Yiannoutsos CT, Cohen BA, Hollander H, Schifitto G, Clifford DB, Simpson DM, Katzenstein D, Shriver S, Hauer P, Brown A, Haidich AB, Moo L, McArthur JC (2002) Reduced intraepidermal nerve fiber density in HIV-associated sensory neuropathy. Neurology 58:115–119

Rahman P, Jones A, Curtis J, Bartlett S, Peddle L, Fernandez BA, Freimer NB (2003) The Newfoundland population: a unique resource for genetic investigation of complex diseases. Hum Mol Genet Spec No 2 12:R167–R172

Roddier K, Thomas T, Marleau G, Gagnon A, Dicaire M, Saint-Denis A, Sarrazin A, Larbrisseau A, Lambert M, Vanasse M, Samuels M, Rouleau G, Brais B (2003) Two distinct mutations are responsible for the higher prevalence of Hereditary Sensory and Autonomic Neuropathy Type II (HSAN II) in French Canadians (Abstract). Paper presented at 53rd Annual Meeting of The American Society of Human Genetics, Los Angeles, California, November 4–8

Sanvito WL, Cataldo BO, Costa AR (2003) [Hereditary sensory and autonomic neuropathy type 2: two cases report]. Arq Neuropsiquiatr 61:654–658

Schaffer AA, Gupta SK, Shriram K, Cottingham RW, Jr. (1994) Avoiding recomputation in linkage analysis. Hum Hered 44: 225–37

Slaugenhaupt SA, Blumenfeld A, Gill SP, Leyne M, Mull J, Cuajungco MP, Liebert CB, Chadwick B, Idelson M, Reznik L, Robbins C, Makalowska I, Brownstein M, Krappmann D, Scheidereit C, Maayan C, Axelrod FB, Gusella JF (2001) Tissue-specific expression of a splicing mutation in the IKBKAP gene causes familial dysautonomia. Am J Hum Genet 68:598–605

Toscano E, Simonati A, Indo Y, Andria G (2002) No mutation in the TRKA (NTRK1) gene encoding a receptor tyrosine kinase for nerve growth factor in a patient with hereditary sensory and autonomic neuropathy type V. Ann Neurol 52: 224–227

Wilson FH, Disse-Nicodeme S, Choate KA, Ishikawa K, Nelson-Williams C, Desitter I, Gunel M, Milford DV, Lipkin GW, Achard JM, Feely MP, Dussol B, Berland Y, Unwin RJ,

Mayan H, Simon DB, Farfel Z, Jeunemaitre X, Lifton RP (2001) Human hypertension caused by mutations in WNK kinases. Science 293:1107–1112

Yasuda H, Terada M, Maeda K, Kogawa S, Sanada M, Haneda M, Kashiwagi A, Kikkawa R (2003) Diabetic neuropathy and nerve regeneration. Prog Neurobiol 69:229–285

Zorrilla Hernandez E, Frati Munari A, Lozano Castaneda O, Villalpando Hernandez S, Boulton AJ (1994) [Diabetic neuropathy: current concepts on etiopathogenesis, diagnosis, and treatment]. Gac Med Mex 130:18–25