Published in final edited form as:

Ann Neurol. 2019 March ; 85(3): 316–330. doi:10.1002/ana.25426.

Variation in *SIPA1L2* Is Correlated with Phenotype Modification in Charcot– Marie– Tooth Disease Type 1A

Feifei Tao, MS[1], **Gary W. Beecham, PhD**[1], **Adriana P. Rebelo, PhD**[1], **Susan H. Blanton, PhD**[1], **John J. Moran**[2], **Camila Lopez-Anido, PhD**[2], **John Svaren, PhD**[2], **Jasper M. Morrow, PhD**[3], **Lisa Abreu, MPH**[1], **Devon Rizzo, MS**[4], **Callyn A. Kirk, MSPH**[4], **Xingyao Wu, MD, PhD**[5], **Shawna Feely, MS, CGC**[5], **Camiel Verhamme, MD, PhD**[6], **Mario A. Saporta, MD, PhD**[7], **David N. Herrmann, MD**[8], **John W. Day, MD, PhD**[9], **Charlotte J. Sumner, MD**[10,11], **Thomas E. Lloyd, MD, PhD**[10,11], **Jun Li, MD, PhD**[12], **Sabrina W. Yum, MD**[13], **Franco Taroni, MD**[14], **Frank Baas, MD, PhD**[15], **Byung-Ok Choi, MD**[16], **Davide Pareyson, MD**[14], **Steven S. Scherer, MD, PhD**[17], **Mary M. Reilly, MD**[3], **Michael E. Shy, MD**[5,*], **Stephan Züchner, MD, PhD**[1,*], **Inherited Neuropathy Consortium**

[1]Department for Human Genetics and Hussman Institute for Human Genomics, University of Miami, Miami, FL; [2]Department of Comparative Biosciences and Waisman Center, University of Wisconsin, Madison, WI; [3]Medical Research Council Centre for Neuromuscular Diseases, University College London Institute of Neurology, London, United Kingdom; [4]Data Management and Coordinating Center, Rare Diseases Clinical Research Network, Pediatrics Epidemiology Center, University of South Florida, Tampa, FL; [5]Department of Neurology, University of Iowa, Iowa City, IA; [6]Department of Neurology, Academic Medical Center, Amsterdam, the Netherlands; [7]Department of Neurology, University of Miami, Miami, FL; [8]Department of Neurology, University of Rochester, Rochester, NY; [9]Department of Neurology, Stanford University, Palo Alto, CA; [10]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD; [11]Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD; [12]Department of Neurology, Wayne State University School of Medicine, Detroit, MI; [13]Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA; [14]IRCCS Foundation Carlo Besta Neurological Institute, Milan, Italy; [15]Department of Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands; [16]Department of Neurology, Samsung Medical Center,

Address correspondence to Dr Züchner, University of Miami Miller School of Medicine, Biomedical Research Building (BRB), Room 616, LC: M-860, 1501 NW 10th Avenue, Miami, FL 33136. szuchner@med.miami.edu.

*M.E.S. and S.Z. contributed equally to this work.

Author Contributions

F.Tao, M.E.S., and S.Z. contributed to the conception and design of the study. All authors contributed to the acquisition of data. F.Tao, G.W.B., A.P.R., S.H.B., J.J.M., and C.L.-A. contributed to the analysis of data. F.Tao, S.Z., and J.S. contributed to drafting the original text. All authors contributed to drafting the final manuscript. F.Tao, A.P.R., and C.L.-A. contributed to drafting the figures.

Inherited Neuropathy Consortium authors: Richard Lewis, MD; Gyula Acsadi, MD, PhD; Richard Finkel, MD; Vera Fridman, MD; Sindhu Ramchandren, MD, MS; David Walk, MD; Eric Logigian, MD; Michael Stanton, MD; Katy Eichinger, PT, DPT, NCS; Debra Guntrum, MS, FNP; Cindy Gibson, MS, NP; Joshua Burns, PhD; Isabella Moroni, MD; Chiara Pisciotta, MD, PhD; Matilde Laurá, MD; Francesco Muntoni, MD; Janet E. Sowden; Joan Mountain, RN; Yunhong Bai; Chelsea Bacon; Laurie Gutmann, MD; Tiffany Grider, MS, CGC; Janel Phetteplace, MS, CGC; Reza Seyedsadjadi, MD; Henry Houlden, MD; Andrea Cortese, MD, PhD; Amelie Pandraud, MS; Chiara Pisciotta, MD, PhD; Daniela Calabrese; Paola Saveri; Jessica Richardson; Lois Dankwa; Diana Lee; Carly Siskind, MS, CGC; Renata Maciel, PhD; Dana Bis.

Potential Conflicts of Interest

Nothing to report.

Sungkyunkwan University School of Medicine, Seoul, South Korea; [17]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

Abstract

Objective: Genetic modifiers in rare disease have long been suspected to contribute to the considerable variance in disease expression, including Charcot-Marie-Tooth disease type 1A (CMT1A). To address this question, the Inherited Neuropathy Consortium collected a large standardized sample of such rare CMT1A patients over a period of 8 years. CMT1A is caused in most patients by a uniformly sized 1.5 Mb duplication event involving the gene *PMP22*.

Methods: We genotyped DNA samples from 971 CMT1A patients on Illumina BeadChips. Genome-wide analysis was performed in a subset of 330 of these patients, who expressed the extremes of a hallmark symptom: mild and severe foot dorsiflexion strength impairment. *SIPA1L2* (signal-induced proliferation-associated 1 like 2), the top identified candidate modifier gene, was expressed in the peripheral nerve, and our functional studies identified and confirmed interacting proteins using coimmunoprecipitation analysis, mass spectrometry, and immunocytochemistry. Chromatin immunoprecipitation and in vitro siRNA experiments were used to analyze gene regulation.

Results: We identified significant association of 4 single nucleotide polymorphisms (rs10910527, rs7536385, rs4649265, rs1547740) in *SIPA1L2* with foot dorsiflexion strength ($p < 1 \times 10^{-7}$). Coimmunoprecipitation and mass spectroscopy studies identified **β**-actin and MYH9 as SIPA1L2 binding partners. Furthermore, we show that *SIPA1L2* is part of a myelination-associated coexpressed network regulated by the master transcription factor SOX10. Importantly, in vitro knockdown of *SIPA1L2* in Schwannoma cells led to a significant reduction of *PMP22* expression, hinting at a potential strategy for drug development.

Interpretation: *SIPA1L2* is a potential genetic modifier of CMT1A phenotypic expressions and offers a new pathway to therapeutic interventions.

Charcot-Marie-Tooth disease (CMT) is a group of inherited neuropathies that affect peripheral motor and sensory axons. It affects 1 in 2,500 individuals worldwide.[1] Typical clinical features include distal muscle weakness and atrophy, sensory loss, steppage gait, and foot deformities. CMT type 1A (CMT1A) is the most common CMT subtype and explains one-third of all CMT cases.[2] It is caused by a uniform 1.5 Mb tandem duplication on chromosome 17p12,[3] which includes the causative gene peripheral myelin protein 22 (*PMP22*). Most CMT1A patients present a typical CMT phenotype, but clinical presentations vary considerably in terms of disease onset, severity, and progression.[2,4] Genetic factors may play important roles in modulating the phenotype in CMT1A patients and could lead to more individualized management of patients as well as offer additional therapeutic targets. There are cases of CMT1A patients with 2 extra copies of *PMP22* (either with homozygous variants from consanguineous marriage,[5,6] or with *PMP22* triplication on 1 chromosome[7,8]) pre-senting a more severe phenotype. These cases represent a rare cause of the severe phenotype in which the fourth copy of *PMP22* further exaggerates the *PMP22* dosage effect and contributes to disease severity. In addition, mutations in

lipopolysaccharide-induced TNF factor (*LITAF*), a gene known to cause demyelinating CMT, have been reported to contribute to a more severe phenotype and earlier onset in CMT1A.[9,10] Recently, micro-RNA 149 (*MIR149*) has been reported as a genetic modifier for disease onset and severity in a group of CMT1A patients with Asian ancestry.[11] Thus far, no studies have systematically screened for genetic modifiers in CMT1A, due to the great challenges of collecting a large uniformly phenotyped sample for this rare disease.

Here, we report our efforts to identify genetic modifiers in CMT1A on a genome-wide scale using a large cohort of 971 CMT1A patients. We identified 4 strongly associated intronic variants in signal-induced proliferation-associated 1 like 2 (*SIPA1L2*), also known as spine-associated RapGAP2 (*SPAR2*), a member of the Rap GTPase-activating proteins (RapGAPs) enriched at synaptic sites.[12–14] Our functional studies imply a role in peripheral nerve myelination and regulating *PMP22* expression. Combined with the genomic evidence, these results support *SIPA1L2* as a potential modifier gene for CMT1A.

Patients and Methods

Sample Collection

A total of 971 CMT1A patients from 705 families were recruited by the Inherited Neuropathies Consortium (INC) at multiple sites since 2011. All patients were recruited by board-certified neurologists following a standard protocol. Physicians were trained in blinded cross-examination exercises with CMT patients to reduce interrater variability of phenotypic assessments. Institutional review board approval was obtained from each site. All participants and/or their legal guardians signed informed consents. Patients with a phenotype consistent with CMT1A had to meet at least 1 of 2 criteria to be included in the study: (1) the participant has a documented *PMP22* duplication and/or (2) the participant has a first-or second-degree relative with a documented *PMP22* duplication.

Clinical Evaluations

Patients' demographic information (including age at examination, sex, race, and ethnicity) and clinical data were obtained from each clinical site and collected by the INC. Foot muscle strength was measured using Medical Research Council (MRC) standards, which grade muscle strength from 0 to 5 on an ordinal scale: grade 0 = no contraction; grade 1 = slight contraction without movement; grade 2 = movement with gravity eliminated; grade 3 = movement against gravity, but not against resistance; grade 4 = movement against gravity and some resistance (4−, 4 and 4+ were used to indicate slight, moderate, and submaximal movement, respectively); grade 5 = normal contraction. For our analysis, the clinical outcome is dichotomized into severe cases (minimal MRC = 0–3, both sides) and mild cases (minimal MRC = 5, both sides).

Genotyping and Quality Control

Genotyping was performed at John P. Hussman Institute for Human Genomics, University of Miami using Illumina (San Diego, CA) OmniExpress BeadChip or Illumina OmniExpressExome BeadChip. Genotype calls were generated with Illumina GenomeStudio software. Overlapping markers from 2 genotyping platforms were combined

for data quality control (QC); single nucleotide polymorphisms (SNPs) were excluded if they had call rate < 95%, were monomorphic, or deviated from Hardy-Weinberg equilibrium ($p < 0.00001$). Samples were excluded if they showed call rate < 95%, had sex mismatch, had unconfirmed relatedness, or did not carry the classic 1.5 Mb duplication on chromosome 17p. Copy number variation (CNV) was checked using Illumina cnvPartition v3.2.1. Population stratification was assessed using EigenStrat.[15] All other QC procedures were performed with PLINK v1.07.[16] The combined dataset after QC included 699,650 markers and 903 samples.

Genome-Wide Analysis

Genome-wide analysis was performed with R package GWAF,[17] which uses generalized estimating equations (GEE) to adjust for family relations. The analysis assumes an additive model. Patients' sex and age at examination were included as covariates in association tests. Manhattan and quantile-quantile (QQ) plots were created using the R package qqman. Regional association plots were created using LocusZoom.[18] Haplotype analysis was performed using Haploview 4.2.

Gene-Based Genome-Wide Association Study and Pathway Analysis

Gene-based genome-wide analysis was performed with VEGAS2[19] using GWAF results of the European population. The definition of a gene also includes SNPs within ±50 kb of the gene boundaries. The EUR population from 1000 Genomes Project was used to infer linkage disequilibrium (LD) structure in the analysis. Pathway enrichment was performed with VEGAS2Pathway,[20] which accounts for gene sizes, LD between markers, and pathway sizes.

Resampling-Based Analyses

Bootstrap analysis was performed with 1,000 times resampling and with replacement of the original dataset. GWAF was performed to obtain the beta coefficients of the SNPs in regression models in each resampled dataset. The 95% confidence interval of beta coefficients was obtained from bootstrap samples. Permutation test was performed by randomly rearranging the phenotype labels (severe or mild cases) in the dataset 10,000 times. GWAF was performed to obtain the chi-squared test statistics in regression models. The original GWAF chi-squared statistic was then compared with the permutation chi-squared test statistics to obtain an empirical p value.

Immunoprecipitation

Rat Schwannoma cells (RT4-D6P2T; ATCC, Manassas, VA) and mouse Neuro-2a cells (ATCC) were transfected with Myc-DDK-tagged *SIPA1L2* construct (OriGene Technologies, Rockville, MD) using Lipofectamine 3000 (Thermo Fisher Scientific, Waltham, MA). After 24 hours, protein lysates were collected for immunoprecipitation with Pierce Magnetic IP/Co-IP Kit (Thermo Fisher Scientific). A monoclonal anti-Myc antibody (#2276; Cell Signaling Technology, Danvers, MA) was crosslinked to magnetic beads and incubated with protein lysates for 1 hour. Beads were washed and eluted in a low-pH elution buffer. Coimmunoprecipitated proteins were analyzed by 4 to 20% sodium dodecyl sulfate-

polyacrylamide gel electrophoresis gel followed by silver staining using Pierce Silver Stain Kit (Thermo Fisher Scientific).

Mass Spectrometry Analyses

Protein bands from silver staining were excised and submitted for mass spectrometry (MS) analysis at Scripps Center for Metabolomics (La Jolla, CA). Peptides were analyzed by reverse-phase chromatography prior to MS analysis using the following method. Nanoelectrospray capillary column tips were made in-house by using a P-100 laser puller (Sutter Instruments, Novato, CA). The columns were packed with Zorbax SB-C18 stationary phase (Agilent Technologies, Santa Clara, CA) purchased in bulk (5 mm particles, with a 15 cm length and a 75 mm inner diameter). The reverse-phase gradient separation was performed by using water and acetonitrile (0.1% formic acid) as the mobile phases. The gradient consisted of 5% acetonitrile for 10 minutes followed by a gradient to 8% acetonitrile for 5 minutes, 35% acetonitrile for 113 minutes, 55% acetonitrile for 12 minutes, 95% acetonitrile for 15 minutes, and re-equilibration with 5% acetonitrile for 15 minutes.

Data-dependent MS/MS data were obtained with an LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific) using a home-built nanoelectrospray source at 2 kV at the tip. One MS spectrum was followed by 4 MS/MS scans on the most abundant ions after the application of the dynamic exclusion list. All MS/MS data were searched against the Mammalia (mammals) National Center for Bio-technology Information database using Mascot (v2.3.02; Matrix Science, London, United Kingdom). Mascot searches were conducted using a peptide mass tolerance of 2.0 Da, a fragment ion mass tolerance of 0.80 Da, fixed modifications of carbamidomethylation, variable modifications of oxidation, an enzyme of trypsin, a maximum of 1 missed cleavage, and a decoy database. Proteins with $p < 0.05$ (corresponding to a Mascot ion score > 61) were identified with ≥2 peptides and with a peptide false discovery rate of 0.02 considered at the 95% confidence level.

Coimmunoprecipitation and Western Blot

Schwannoma cells were cotransfected with SIPA1L2-Myc and myosin heavy chain 9 (MYH9)-HA. After 24 hours, protein lysates were collected for coimmunoprecipitation with immunoprecipitation Kit Dynabeads Protein A (Thermo Fisher Scientific) following the manufacturer's protocol. Coimmunoprecipitation results were analyzed by Western blot probed against primary antibodies anti-Myc (Cell Signaling Technology #2276) and anti-HA (Cell Signaling Technology #3724) and secondary antibodies EasyBlot antimouse IgG (GTX221667–01 Lot No. 42527; GeneTex, Irvine, CA) and EasyBlot antirabbit IgG (GeneTex GTX221666–01 Lot No. 41082).

Immunocytochemistry

Schwannoma cells were cotransfected with Myc-DDK-tagged *SIPA1L2* construct and *MYH9*-HA tagged construct using Lipofectamine 3000 (Invitrogen, Carlsbad, CA). After 24 hours, cells were fixed with paraformaldehyde for 20 minutes, permeabilized with 0.1% Triton-X, and incubated with a monoclonal anti-Myc antibody (Cell Signaling Technology #2276) and a polyclonal anti-HA antibody (Cell Signaling Technology #3724) for 2 hours at room temperature. Fluorescent secondary antibodies (Alexa Fluor 488 goat antimouse,

Thermo Fisher Scientific A11001; Alexa Fluor 647 goat antirabbit, Thermo Fisher Scientific A21244) were incubated along with Alexa Fluor 555 phalloidin (Thermo Fisher Scientific A34055) for 1 hour. Cells were then mounted onto slides and imaged with a confocal microscope (LSM710; Carl Zeiss, Oberkochen, Germany).

Gene Expression and Chromatin Immunoprecipitation Analysis

Histone and transcription factor (TF) chromatin immunoprecipitation (ChIP)-Seq datasets were previously published.[21–25] To convert the coordinates of regulatory elements from rn5 to hg19, we implemented the liftOver utility from the University of California, Santa Cruz genome browser.[26] The S16 rat Schwann cell line (provided by Richard Quarles) was grown as described. S16 cells were transfected with Lonza (Basel, Switzerland) Amaxa SE Cell Line 4D-Nucleofector X Kit L (Cat# V4XC-1024) using 40 nmol of siRNA directed toward *Sox10* (s131239; Ambion, Austin, TX), Integrated DNA Technologies (IDT; Coralville, IA) #rn.Ri.Sipa1l2.13.1 (rGrCrUrCrUrArUrUrU rUrUrU rUrGrG rArUrG rUrAr C rArGT A and rUrArC rUrGrU rArCrA rUrCrC rArArA rArArA rArUrA rGrArG rCrArC), or an IDT Negative Control DsiRNA (NC-1, IDT #51–01-14–04). At 48 hours after transfection, RNA was purified and analyzed by quantitative real-time polymerase chain reaction (qRT-PCR) with the following primers: *Egr2*, forward GCTGGAG ATGGCATGATCAAC, reverse TCTGGTTTCTAGGCG CAGAGA; *Mpz*, forward CTGCAGTCAAATCCCCCAG TA, reverse CCTGGAGGTGACGGTCACTT; *Sipa1l2*, forward ACATTCTCTTCAGCACGACTC, reverse TCT GCACACTTGGACTTGTC; *Pmp22*, forward GTCAT CTTCAGCGTCCTGTC, reverse TGCACTCATCACAC ACAGAC; *Actb*, forward CCTAGCACCATGAAGATC AAGA, reverse CTCATCGTACTCCTGCTTGC.

Results

Study Cohort

CMT1A patients were enrolled over a timeframe of 7 years with the purpose of creating a uniformly phenotyped cohort for natural history and detailed genomic studies. To ensure clinical assessment reliability, only trained neurologists performed physical examinations who had completed cross-training CMT phenotyping exercises. A total of 971 ascertained CMT1A patients were considered for this study (Fig 1A, Table 1). Motor and sensory functions are affected in CMT1A; however, motor involvement is often the more disabling component, leading to specific gait abnormalities and hand weakness. Weakness of the anterior tibialis muscle leads to ankle dorsiflexion weakness, including foot drop, which in turn causes patients to develop steppage gait, a hallmark of CMT. We thus decided to focus genotype correlations on foot dorsiflexion strength. Using the MRC scale, a standard clinical measure of muscle strength, we distinguished severe weakness (MRC = 0–3) from no weakness (MRC = 5; see Fig 1B). Phenotype analysis of our cohort showed that age contributes to higher likelihood of developing severe weakness (odds ratio [OR] = 2.20, $p = 2 \times 10^{-16}$ for every 10 years of age).[2] In addition, females were more likely to develop severe muscle strength impairment than males independent of age (OR = 1.91, $p = 0.015$).

Cases were genotyped in 2 waves as they became available on Illumina BeadChips. As deviations from the canonical chromosome 17p breakpoints have been reported,[27] we first

performed CNV analysis in each patient to confirm the classic 1.5 Mb duplication. We identified a small subset of patients with different duplication sizes: 8 patients (0.8%) with an extended duplication and 17 patients (1.8%) with a shorter-sized duplication, all of which included the *PMP22* locus. These samples were removed from the dataset to eliminate potential confounding effects caused by genetic heterogeneity at the *PMP22* locus. Next, to avoid genotype bias introduced through genetic ancestries, we performed principal component analysis using EigenStrat.[15] The top 2 principal components identified 2 distinct populations in the cohort: 652 individuals with European ancestry and 224 individuals with Asian ancestry. Twenty-nine individuals with unclear genetic ancestry were excluded from further analyses. After further data quality control, 644 European ancestry samples and 213 Asian ancestry samples entered the genotype analysis stage.

Genome-Wide Association Analyses

We then performed genome-wide association analyses to identify SNPs associated with weakness in foot dorsiflexion. Taking into account the substantial difference in genetic ancestry, the initial analysis was performed in the European-ancestry cohort only (n = 644). Foot dorsiflexion data were reported in 552 individuals (85.7%). After removing 92 samples with missing phenotype data and 222 samples with moderate severity (foot dorsiflexion of 4−, 4, or 4+), we obtained a final dataset of 330 individuals (183 severe cases and 147 mild cases from 265 families). Quality control measures as described in Patients and Methods yielded a final 633,827 SNPs (619,496 autosomal SNPs and 14,331 X chromosome SNPs) for further analyses in this cohort. To take advantage of related family members in this cohort, we performed genome-wide association analysis using the R package GWAF, which uses a GEE model to adjust for relatedness by treating each pedigree as a cluster. Sex and age at examination were included as covariates in the analysis.

For a rare disease, it is challenging to collect a large sample size for a well-controlled genome-wide study. Only relatively large effect sizes would be expected to create a genome-wide signal after correction for multiple testing. A strict Bonferroni threshold for this study would require a $p < 7.9 \times 10^{-8}$ (633,827 tests).[28] Considering LD between SNPs, this Bonferroni threshold may result in overly conservative correction, because tests for each SNP are not completely independent. Therefore, we lowered our required threshold to $p < 1 \times 10^{-7}$. It has been reported that a substantial proportion of borderline-significant genome-wide association study (GWAS) findings (p value between 5×10^{-8} and 1×10^{-7}) were later proven replicable,[29] suggesting that relaxing our GWAS threshold is reasonable.

We identified 4 SNPs on chromosome 1 reaching or approaching a genome-wide significance threshold p value of 1×10^{-7}: rs10910527 ($p = 6.30 \times 10^{-8}$, OR = 12.794), rs7536385 ($p = 6.30 \times 10^{-8}$, OR = 12.794), rs4649265 ($p = 1.06 \times 10^{-7}$, OR = 10.881), and rs1547740 ($p = 1.13 \times 10^{-7}$, OR = 10.859; see Fig 1C, Tables 2 and 3). The minor alleles of the SNPs were associated with increased disease severity in muscle strength loss. In the severe extremes (foot dorsiflexion = 0–3), the minor allele frequency (MAF) of the top SNP rs10910527 was 0.0492, whereas in the severe extremes (foot dorsiflexion = 5) the MAF was 0.0068. Individuals with moderate severity (foot dorsiflexion = 4) had MAF = 0.0294. The QQ plot (see Fig 1D) did not show signs of systemic bias in the association results (λ =

1.04). The 4 SNPs are located in the intronic regions of the gene *SIPA1L2* (Fig 2), which codes for a ubiquitously expressed putative GTPase with unknown functions. In addition, 2 SNPs were identified at different loci with suggestive significance ($p < 1 \times 10^{-5}$): rs303143 ($p = 7.60 \times 10^{-7}$, OR = 0.033) and rs2213767 ($p = 4.37 \times 10^{-6}$, OR = 4.039; see Fig 1C, Table 2). The SNP rs303143 is an intronic polymorphism in spermatogenesis-associated 5 (*SPATA5*) on chromosome 4. This gene encodes a member of the ATPase family and is proposed to be involved in the maintenance of mitochondrial functions. The SNP rs2213767 is located in an intergenic region on chromosome 22, close to a noncoding gene *LINC01422* and an uncharacterized gene *LOC284898*.

Gene-based association studies offer a gene-centric correlation map with a reduced number of individual tests. We performed gene-based genome-wide analysis in the European population using VEGAS2.[19] The *SIPA1L2* gene showed a *p* value of 0.0065 with 87 SNPs from the original genotyping dataset included in the gene-based test.

At an even higher level of abstraction, pathway analysis is able to identify groups of connected genes that function together. Pathway analysis using VEGAS2Pathway[20] did not identify any significantly enriched pathways after Bonferroni correction for 6,211 pathways ($p < 8.05 \times 10^{-6}$).

Locus-Specific Analyses

From the genome-wide analysis, *SIPA1L2* is the most promising candidate modifier gene as ranked by *p* values and ORs of the lead SNPs. *SIPA1L2* is a large GTPase gene that spans 21 exons over 100 kb of genomic space. The 4 associated SNPs lie in introns 6, 7, 8, and 12 of transcript NM_020808.4. There is strong LD between the markers (pairwise R^2 ranges from 0.78 to 1 in our dataset, and from 0.69 to 1 in 1000 Genomes EUR). To confirm the genetic association of the 4 SNPs, we performed resampling-based association analyses, including bootstrap analysis and permutation test. In the bootstrap analysis, we performed resampling 1,000 times and calculated the beta coefficients of the SNPs from GWAF regression analysis. The 95% confidence interval of beta coefficient obtained from bootstrapping was 0.697–18.344 for rs10910527, 0.697–18.344 for rs7536385, 0.481–18.071 for rs4649265, and 0.469–18.075 for rs1547740. None of the intervals include 0, suggesting the effects of the SNPs are significant. In the permutation test, the phenotype labels (severe or mild cases) were randomly rearranged 10,000 times, and the chi-squared test statistic from GWAF regression analysis was obtained from each permutation. In the top SNP rs10910527, the original chi-squared statistic of 29.3 was more extreme than the permutation chi-squared statistics, which ranged from 0 to 15.3, with a median of 0.408. Similar results were also obtained from the other 3 SNPs, suggesting that the *p* values of the SNPs were more significant than $p = 0.0001$ (for 10,000 permutations). Both analyses further confirmed the robustness of the significant association of the 4 SNPs in the European cohort.

We then tested whether these markers were also correlated to foot dorsiflexion outcomes in the 213 individuals of Asian ancestry. After removing individuals with moderate foot dorsiflexion strength and/or missing clinical data, the Asian cohort included only 55 severe cases (foot dorsiflexion = 0–3) and 25 mild cases (foot dorsiflexion = 5). The 4 SNPs identified in the European cohort were not correlated with foot dorsiflexion in this small

sample. We compared haplotype structure and found profound differences in the distribution of LD and number and size of haploblocks; there are 15 haplotype blocks in the European and 12 haplotype blocks in the Asian population (Fig 3). Entire haploblocks were not associated with the genomic locus in either population. The European population has stronger LD in the genomic region, whereas the Asian population shows a more fragmented genomic substructure. Thus, the lead SNPs in the European population may not account for association in the Asian population, where other variation could indicate association. We expanded the association analyses in the Asian cohort from 4 SNPs to the entire gene locus of *SIPA1L2*, including 120 kb up- and downstream. We identified 4 variants located 70 kb upstream of *SIPA1L2* that reached $p < 0.05$ (rs11810061, rs1334401, rs2025724, rs2210669). The small sample size of the Asian cohort prevented us from more detailed analyses.

We also explored potential causal SNPs in the *SIPA1L2* locus using public databases. First, we explored the regulatory potential of the significant *SIPA1L2* SNPs identified from GWAS. We annotated the SNPs using regulomeDB score,[30] which grades the regulatory potential of SNPs on a scale from 1 (high regulatory potential, likely to affect binding and linked to expression of a gene target) to 6 (low regulatory potential, with minimal binding evidence). Low scores indicate increasing evidence for a variant to be located in a functional region. The regulomeDB scores for the top 4 *SIPA1L2* SNPs were 5 for rs10910527, 6 for rs7536385, 6 for rs4649265, and 2b for rs1547740 (Table 4). SNP rs1547740 is likely to affect binding (TF binding + any motif + DNase Footprint + DNase peak), suggesting it may have regulatory functions.

To explore other potential causal SNPs with regulatory functions in the *SIPA1L2* locus, we searched for proxy SNPs in LD ($R^2 > 0.6$) with the top GWAS SNP rs10910527 in 1000 Genomes EUR (https://ldlink.nci.nih.gov/). A total of 28 additional SNPs were identified in LD with rs10910527. Among those SNPs, 2 SNPs show regulatory potential as rated by regulomeDB score (see Table 4). SNP rs10495327 (chr1:232,608,616, MAF = 0.0338, R^2 = 0.74 with top SNP) shows a regulomeDB score of 2b (likely to affect binding). SNP rs6662895 (chr1:232,590,092, MAF = 0.0348, R^2 = 0.82 with top SNP) shows a regulomeDB score of 3a (TF binding + any motif + DNase peak). Together, the data show that SNPs rs1547740, rs10495327, and rs6662895 in the *SIPA1L2* locus may be candidate causal variants with regulatory functions for further studies in the future.

Identification of Interaction between SIPA1L2, Actin, and MYH9

To begin to understand the molecular functions of *SIPA1L2*, we performed coimmunoprecipitation analysis in Neuro-2a and Schwannoma cells transfected with *SIPA1L2-Myc* to identify protein binding partners. Silver staining analysis of immunoprecipitates using an anti-Myc antibody revealed prevalent bands of ~60 kDa and ~200 kDa in both Neuro-2a and Schwannoma SIPA1L2 samples that were absent in the control IgG samples (Fig 4A). MS analysis identified these bands as β-actin and MYH9. To validate these protein interactions, Schwannoma cells were cotransfected with SIPA1L2-Myc and MYH9-HA and cell lysates were immunoprecipitated with an anti-Myc antibody. Western blot detected specific pulldown of both SIPA1L2 and MYH9, confirming the

interactions between the two proteins (see Fig 4B). To investigate further the interaction between these proteins, colocalization analysis by fluorescence microscopy was performed. Schwannoma and Neuro-2a cells were cotransfected with SIPA1L2-Myc and MYH9 followed by phalloidin staining, which binds with high affinity to filamentous actin. SIPA1L2 was localized in the cytoplasm and cellular projections with accumulations at the tips of projections. SIPA1L2 strongly colocalized with both MYH9 and actin predominantly in growth cones (see Fig 4C).

Regulation of Sipa1l2 Expression

To investigate the biology of *SIPA1L2* in Schwann cells, we mined our previously published gene expression and ChIP-Seq analyses. In rodents, *Sipa1l2* is expressed throughout Schwann cell development, from embryonic to adult stages of nerve.[31,32] In addition, *Sipa1l2* was previously identified by some of us as an Sox10-regulated gene by siRNA analysis performed in a Schwann cell line.[21] We performed a confirmatory qRT-PCR and found that downregulation of *Sox10* mRNA leads to lower levels of *SIPA1L2* mRNA (Fig 5A). To identify Schwann cell-specific regulation, we examined previous histone and TF ChIP-Seq analyses.[21,23] Histone 3 lysine 27 acetylation (H3K27ac) ChIP-Seq in P15 sciatic nerve revealed several putative enhancers around the rat *Sipa1l2* gene locus, and there are also binding sites present for Sox10 that coincide with these enhancers. TF ChIP-Seq revealed that Sox10 appears to bind at least 2 of these intragenic enhancers (see Fig 5B).

Although *Sipa1l2* gene expression is relatively ubiquitous, its expression is reduced after sciatic nerve injury, ~1.6-fold 3 days postcrush[33] and ~2.7-fold in the distal stump at 7 days after transection.[34] To identify potentially critical enhancers required for *Sipa1l2* expression, we assessed H3K27ac ChIP-Seq profiles around the gene in sham-injured sciatic nerves compared with 3 days after transection.[22] One intragenic enhancer was lost after injury (see Fig 5C), thus suggesting that it may play a role in driving expression of *Sipa1l2*. We also found that the promoter region of Sipa1l2, marked by H3K4me3 (see Fig 5D), contains a strong Egr2 binding site, the TF Egr2 that plays a major role in driving myelin gene expression.[35] Because the expression of Egr2 decreases after nerve injury,[36] Egr2 may be involved in the injury regulation of Sipa1l2 (see Fig 5C).

We next tested whether the downregulation of *SIPA1L2* leads to any transcriptional changes. We transfected the S16 Schwann cell line as described above with *Sipa1l2* siRNA and found that its mRNA level was efficiently reduced (see Fig 5E). Interestingly, we found that the mRNA levels of other myelin-related genes, including *Pmp22, Mpz*, and *Egr2*, were also significantly reduced (see Fig 5E). These results suggest that *Sipa1l2* participates in a myelination-associated gene regulatory network in Schwann cells. In addition to its role in actin association, even a subtle change in myelin gene expression caused by *Sipa1l2* dysregulation could exacerbate the consequences of the CMT1A-associated *PMP22* gene duplication.

We also analyzed publicly available datasets in the Genotype-Tissue Expression project (https://gtexportal.org/home/) to examine whether *SIPA1L2* SNPs are involved as expression quantitative trait loci (eQTLs). In tibial nerves (n = 361), the *p* values of the associated SNPs did not support correlation with *SIPA1L2* expression. Interestingly, the eQTL *p* values for

these SNPs and correlation with *PMP22* expression were 0.056 for rs10910527, rs7536385, and rs4649265, as well as 0.074 for rs1547740.

Discussion

The identification of genetic modifiers reveals additional aspects of genetic architecture, expands concepts of Mendelian disease, and offers progress for diagnosis and treatment. However, the challenge in rare diseases is the limited sample sizes and locus and allelic heterogeneity. CMT1A is a valuable target for modifier gene studies, as it is by far the most common form of CMT and is caused by a uniform 1.5 Mb duplication on chromosome 17p11.2 (97.4% of participants in this study, the largest series of patients to date). Despite this genetic homogeneity of CMT1A, patients present with considerable clinical phenotypic variability.[2,4] Existing studies had limited success in identifying genetic CMT1A modifiers.[9–11]

Over a timeframe of 7 years, the INC has collected nearly 1,000 CMT1A samples under a standardized clinical protocol. This led to this hitherto largest genomic study in CMT1A and one of the largest uniform cohort studies in any rare disease. In a case-only GWAS design and after consideration of genetic ancestry, we identified a genome-wide significant signal in the gene *SIPA1L2*. Four intronic SNPs (rs10910527, rs7536385, rs4649265, rs1547740) showed association with foot dorsiflexion in CMT1A patients ($p < 1 \times 10^{-7}$). Patients carrying the minor alleles of these variants are at increased risk of developing more severe weakness of the anterior tibialis muscle, leading to foot drop and steppage gait, a hallmark clinical symptom of CMT neuropathy.

Foot dorsiflexion strength, measured by the MRC scale, is a direct and immediate indicator of disease severity of motor function impairment in CMT. More generally, CMT disease severity has been quantified using the CMT neuropathy score (CMTNS), a composite scoring system that combines symptoms, signs, and neurophysiology.[37] Of the 9 items on the CMTNS, MRC contributes directly to 2 signs (item 6 [strength in legs] and item 7 [strength in arms]), and is also related to 2 symptoms (item 2 [motor symptoms in legs] and item 3 [motor symptoms in arms]). Foot dorsiflexion strength is significantly correlated with CMTNS, with a Spearman correlation ρ of −0.657 and p value of $<2.2 \times 10^{-16}$ (n = 403 in the European-ancestry cohort). In addition, MRC is significantly correlated with quantitative magnetic resonance imaging (MRI) measures, which are commonly used biomarkers in CMT. Further analyzing cross-sectional data from a previous study,[38] we now calculated the correlation between MRI anterior calf compartment fat fraction (AC FF) and foot dorsiflexion MRC grade, and obtained a Spearman correlation ρ of −0.64 and p value of 0.00001 (n = 39). Furthermore, we found in a follow-up 4-year longitudinal study in the same cohort that progression of AC FF is worse in individuals with declining foot dorsiflexion MRC strength compared to individuals with stable MRC strength ($p = 0.07$ in a paired t test). Together, these data cross-validate the significance of MRC muscle strength as an outcome measure in CMT.

The studied cohort is sizable in rare diseases, but of modest sample size in comparison to common disease GWASs. Thus, challenges arise from statistical power and multiple testing,

as well as lack of availability of sizable confirmation cohorts. Only single SNP markers with large effect sizes would be detectable. The *SIPA1L2* markers carried large ORs of 10.8 to 12.8. Several approaches were taken to maximize statistical power: (1) to ensure allelic homogeneity of the duplication mutation, we performed CNV analysis in each participant and excluded 2.6% of samples; (2) we collapsed the ordinal MRC trait into 2 dichotomous extremes of the foot dorsiflexion phenotype to enrich for samples that presumably carry strong phenotype-associated genotypes; and (3) we performed association analysis with the GWAF analytical package, which uses a GEE model to adjust for family relations. This allowed us to include informative small families in the study, while controlling for confounding relatedness effects.

Although direct genetic replication of these results was not possible, several independent lines of evidence are in support of the reported finding. The Asian cohort, constrained by sample size (55 severe and 25 mild cases) and a different LD structure than European samples, carried 4 significant variants 70 kb upstream of exon 1. In the European sample, we ran a gene-based association test based on 85 SNPs and found *SIPA1L2* to be significantly associated. More importantly, 2 resampling-based approaches, including bootstrap analysis and permutation test, successfully confirmed the statistical significance of the 4 signal SNPs in *SIPA1L2* in the European cohort. Finally, *SIPA1L2* is expressed in axons and Schwann cells and we showed that it belongs to an SOX10-regulated gene network of key myelination genes, including *PMP22, MPZ, PMP2*, and *EGR2*,[21] all of which are CMT1-causing genes.

SIPA1L2 encodes a member of the signal-induced proliferation-associated 1 (SIPA1) protein family, which is also known as the spine-associated RapGAP (SPAR) family. This family includes 4 members: SIPA1 (SPA-1), SIPA1L1 (SPAR1), SIPA1L2 (SPAR2), and SIPA1L3 (SPAR3).[14] These proteins share 3 common conserved domains: an N-terminal RapGAP domain, a PDZ domain, and a C-terminal coiled-coil domain with a leucine zipper.[13] SIPA1 is reported to colocalize with and activate RAP1 and RAP2,[39] and is also involved in mitogen-induced cell cycle progression.[40] SIPA1L1 has a reported role in synaptic actin cytoskeleton regulation and dendritic spine morphology changes.[14] Finally, SIPA1L3 has a function in epithelial cell morphogenesis and establishment or maintenance of polarity,[41] and mutations in *SIPA1L3* are known to cause congenital cataract.[41,42] Current knowledge of the molecular function of SIPA1L2 is limited, but it is expressed throughout Schwann cell development, from embryonic to adult stages of peripheral nerves.[31,32] Studies suggest SIPA1L2 also has GTPase activating protein activity and functions in cerebellar and hippocampal granule cells.[13] In previous GWAS studies, *SIPA1L2* was identified as a risk locus (lead SNP rs10797576, chr1:232,664,611) for Parkinson disease in European[43] and Iranian populations.[44]

Importantly, our functional experiments support a role for *SIPA1L2* in Schwann cell biology. We identified MYH9 and actin as protein binding partners, both of which are expressed by myelinating Schwann cells.[45,46] The latter is of increasing interest, as actin network remodeling has recently been shown as a necessary step for myelination of axons.[47–51] Furthermore, axon myelination is regulated by the TF SOX10.[35] In Schwann cells, we demonstrate that knockdown of *SOX10* reduces expression of *SIPA1L2*. Most intriguing is that *SIPA1L2* knockdown in turn reduces expression of coexpressed and coregulated

myelination proteins, including PMP22. Pending future follow-up studies, this latter result provides a potential mechanism of action for therapeutic intervention.

In conclusion, this study revealed novel genetic mechanisms involved in the disease phenotype of CMT1A. Polymorphisms in *SIPA1L2* are shown to modify the severity of foot dorsiflexion in CMT1A patients of European ancestry. This conclusion is supported by an unbiased genome-wide scan as well as additional functional evidence. Although more replication studies using independent cohorts are required to further validate these findings, our results provide novel insights into the mechanism of CMT1A and open up a new direction for the development of therapeutic drugs.

Acknowledgment

This work was supported by the NIH (R01NS43174, U54NS065712 to M.E.S., S.S.S., S.Z.; R01NS075764, U54NS065712 to S.Z., M.A.S.; R01NS075269 to J.S.; U54HD090256; R01NS094388 to B.-O.C.), Charcot-Marie-Tooth Association, and Muscular Dystrophy Association. M.M.R. received support from the MRC (MRC Centre grant G0601943), NIH National Institute of Neurological Diseases and Stroke, and Office of Rare Diseases (U54NS065712). This research was also supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

We thank the participants and their families for their contributions to the study.

References

1. Skre H. Genetic and clinical aspects of Charcot-Marie-Tooth's disease. Clin Genet 1974;6:98–118. [PubMed: 4430158]
2. Fridman V, Bundy B, Reilly MM, et al. CMT subtypes and disease burden in patients enrolled in the Inherited Neuropathies Consortium natural history study: a cross-sectional analysis. J Neurol Neurosurg Psychiatry 2015;86:873–878. [PubMed: 25430934]
3. Lupski JR, de Oca-Luna RM, Slaugenhaupt S, et al. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. Cell 1991;66: 219–232. [PubMed: 1677316]
4. van Paassen BW, van der Kooi AJ, van Spaendonck-Zwarts KY, et al. PMP22 related neuropathies: Charcot-Marie-Tooth disease type 1A and hereditary neuropathy with liability to pressure palsies. Orphanet J Rare Dis 2014;9:38. [PubMed: 24646194]
5. Sturtz FG, Latour P, Mocquard Y, et al. Clinical and electrophysiological phenotype of a homozygously duplicated Charcot-Marie-Tooth (type 1A) disease. Eur Neurol 1997;38:26–30.
6. LeGuern E, Gouider R, Mabin D, et al. Patients homozygous for the 17p11.2 duplication in Charcot-Marie-Tooth type 1A disease. Ann Neurol 1997;41:104–108. [PubMed: 9005872]
7. Liu P, Gelowani V, Zhang F, et al. Mechanism, prevalence, and more severe neuropathy phenotype of the Charcot-Marie-Tooth type 1A triplication. Am J Hum Genet 2014;94:462–469. [PubMed: 24530202]
8. Kim SM, Lee J, Yoon BR, et al. Severe phenotypes in a Charcot-Marie-Tooth 1A patient with PMP22 triplication. J Hum Genet 2015; 60:103–106. [PubMed: 25500726]
9. Meggouh F, de Visser M, Arts WF, et al. Early onset neuropathy in a compound form of Charcot-Marie-Tooth disease. Ann Neurol 2005; 57:589–591. [PubMed: 15786462]
10. Sinkiewicz-Darol E, Lacerda AF, Kostera-Pruszczyk A, et al. The LITAF/SIMPLE I92V sequence variant results in an earlier age of onset of CMT1A/HNPP diseases. Neurogenetics 2015;16:27–32. [PubMed: 25342198]
11. Nam SH, Kanwal S, Nam DE, et al. Association of miR-149 polymorphism with onset age and severity in Charcot-Marie-Tooth disease type 1A. Neuromuscul Disord 2018;28:502–507. [PubMed: 29729827]

12. Dolnik A, Kanwal N, Mackert S, et al. Sipa1l3/SPAR3 is targeted to postsynaptic specializations and interacts with the Fezzin ProSA-PiP1/Lzts3. J Neurochem 2016;136:28–35. [PubMed: 26364583]
13. Spilker C, Acuna Sanhueza GA, Bockers TM, et al. SPAR2, a novel SPAR-related protein with GAP activity for Rap1 and Rap2. J Neurochem 2008;104:187–201. [PubMed: 17961154]
14. Spilker C, Kreutz MR. RapGAPs in brain: multipurpose players in neuronal Rap signalling. Eur J Neurosci 2010;32:1–9. [PubMed: 20576033]
15. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–909. [PubMed: 16862161]
16. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575. [PubMed: 17701901]
17. Chen MH, Yang Q. GWAF: an R package for genome-wide association analyses with family data. Bioinformatics 2010;26:580–581. [PubMed: 20040588]
18. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010; 26:2336–2337. [PubMed: 20634204]
19. Mishra A, Macgregor S. VEGAS2: software for more flexible gene-based testing. Twin Res Hum Genet 2015;18:86–91. [PubMed: 25518859]
20. Mishra A, MacGregor S. A novel approach for pathway analysis of GWAS data highlights role of BMP signaling and muscle cell differentiation in colorectal cancer susceptibility. Twin Res Hum Genet 2017;20:1–9. [PubMed: 28105966]
21. Srinivasan R, Sun G, Keles S, et al. Genome-wide analysis of EGR2/-SOX10 binding in myelinating peripheral nerve. Nucleic Acids Res 2012;40:6449–6460. [PubMed: 22492709]
22. Hung HA, Sun G, Keles S, Svaren J. Dynamic regulation of Schwann cell enhancers after peripheral nerve injury. J Biol Chem 2015;290: 6937–6950. [PubMed: 25614629]
23. Lopez-Anido C, Sun G, Koenning M, et al. Differential Sox10 genomic occupancy in myelinating glia. Glia 2015;63:1897–1914. [PubMed: 25974668]
24. Yu Y, Chen Y, Kim B, et al. Olig2 targets chromatin remodelers to enhancers to initiate oligodendrocyte differentiation. Cell 2013;152: 248–261. [PubMed: 23332759]
25. Ma KH, Hung HA, Svaren J. Epigenomic regulation of Schwann cell reprogramming in peripheral nerve injury. J Neurosci 2016;36: 9135–9147. [PubMed: 27581455]
26. Speir ML, Zweig AS, Rosenbloom KR, et al. The UCSC Genome Browser database: 2016 update. Nucleic Acids Res 2016;44(D1): D717–D725. [PubMed: 26590259]
27. Zhang F, Seeman P, Liu P, et al. Mechanisms for nonrecurrent genomic rearrangements associated with CMT1A or HNPP: rare CNVs as a cause for missing heritability. Am J Hum Genet 2010; 86:892–903. [PubMed: 20493460]
28. Clarke GM, Anderson CA, Pettersson FH, et al. Basic statistical analysis in genetic case-control studies. Nat Protoc 2011;6:121–133. [PubMed: 21293453]
29. Panagiotou OA, Ioannidis JP, Genome-Wide Significance Project. What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. Int J Epidemiol 2012; 41:273–286. [PubMed: 22253303]
30. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012;22:1790–1797. [PubMed: 22955989]
31. Sennett R, Wang Z, Rezza A, et al. An integrated transcriptome atlas of embryonic hair follicle progenitors, their niche, and the developing skin. Dev Cell 2015;34:577–591. [PubMed: 26256211]
32. Le N, Nagarajan R, Wang JY, et al. Analysis of congenital hypomyelinating Egr2Lo/Lo nerves identifies Sox2 as an inhibitor of Schwann cell differentiation and myelination. Proc Natl Acad Sci U S A 2005; 102:2596–2601. [PubMed: 15695336]
33. Barrette B, Calvo E, Vallieres N, Lacroix S. Transcriptional profiling of the injured sciatic nerve of mice carrying the Wld(S) mutant gene: identification of genes involved in neuroprotection, neuroinflammation, and nerve regeneration. Brain Behav Immun 2010;24:1254–1267. [PubMed: 20688153]

34. Arthur-Farraj PJ, Latouche M, Wilton DK, et al. c-Jun reprograms Schwann cells of injured nerves to generate a repair cell essential for regeneration. Neuron 2012;75:633–647. [PubMed: 22920255]
35. Svaren J, Meijer D. The molecular machinery of myelin gene transcription in Schwann cells. Glia 2008;56:1541–1551. [PubMed: 18803322]
36. Zorick TS, Syroid DE, Arroyo E, et al. The transcription factors SCIP and Krox-20 mark distinct stages and cell fates in Schwann cell differentiation. Mol Cell Neurosci 1996;8:129–145.
37. Murphy SM, Herrmann DN, McDermott MP, et al. Reliability of the CMT neuropathy score (second version) in Charcot-Marie-Tooth disease. J Peripher Nerv Syst 2011;16:191–198. [PubMed: 22003934]
38. Morrow JM, Sinclair CD, Fischmann A, et al. MRI biomarker assessment of neuromuscular disease progression: a prospective observational cohort study. Lancet Neurol 2016;15:65–77. [PubMed: 26549782]
39. Kurachi H, Wada Y, Tsukamoto N, et al. Human SPA-1 gene product selectively expressed in lymphoid tissues is a specific GTPase-activating protein for Rap1 and Rap2. Segregate expression profiles from a rap1GAP gene product. J Biol Chem 1997;272: 28081–28088. [PubMed: 9346962]
40. Hattori M, Tsukamoto N, Nur-e-Kamal MS, et al. Molecular cloning of a novel mitogen-inducible nuclear protein with a Ran GTPase-activating domain that affects cell cycle progression. Mol Cell Biol 1995;15:552–560. [PubMed: 7799964]
41. Greenlees R, Mihelec M, Yousoof S, et al. Mutations in SIPA1L3 cause eye defects through disruption of cell polarity and cytoskeleton organization. Hum Mol Genet 2015;24:5789–5804. [PubMed: 26231217]
42. Evers C, Paramasivam N, Hinderhofer K, et al. SIPA1L3 identified by linkage analysis and whole-exome sequencing as a novel gene for autosomal recessive congenital cataract. Eur J Hum Genet 2015;23: 1627–1633. [PubMed: 25804400]
43. Nalls MA, Pankratz N, Lill CM, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nat Genet 2014;46:989–993. [PubMed: 25064009]
44. Safaralizadeh T, Jamshidi J, Esmaili Shandiz E, et al. SIPA1L2, MIR4697, GCH1 and VPS13C loci and risk of Parkinson's diseases in Iranian population: a case-control study. J Neurol Sci 2016;369:1–4. [PubMed: 27653855]
45. Trapp BD, Andrews SB, Wong A, et al. Co-localization of the myelin-associated glycoprotein and the microfilament components, F-actin and spectrin, in Schwann cells of myelinated nerve fibres. J Neurocytol 1989;18:47–60. [PubMed: 2468742]
46. Patzig J, Jahn O, Tenzer S, et al. Quantitative and integrative proteome analysis of peripheral nerve myelin identifies novel myelin proteins and candidate neuropathy loci. J Neurosci 2011;31: 16369–16386. [PubMed: 22072688]
47. Jin F, Dong B, Georgiou J, et al. N-WASp is required for Schwann cell cytoskeletal dynamics, normal myelin gene expression and peripheral nerve myelination. Development 2011;138: 1329–1337. [PubMed: 21385763]
48. Novak N, Bar V, Sabanay H, et al. N-WASP is required for membrane wrapping and myelination by Schwann cells. J Cell Biol 2011;192: 243–250. [PubMed: 21263026]
49. Montani L, Buerki-Thurnherr T, de Faria JP, et al. Profilin 1 is required for peripheral nervous system myelination. Development 2014;141:1553–1561. [PubMed: 24598164]
50. Nawaz S, Sanchez P, Schmitt S, et al. Actin filament turnover drives leading edge growth during myelin sheath formation in the central nervous system. Dev Cell 2015;34:139–151. [PubMed: 26166299]
51. Samanta J, Salzer JL. Myelination: actin disassembly leads the way. Dev Cell 2015;34:129–130. [PubMed: 26218317]

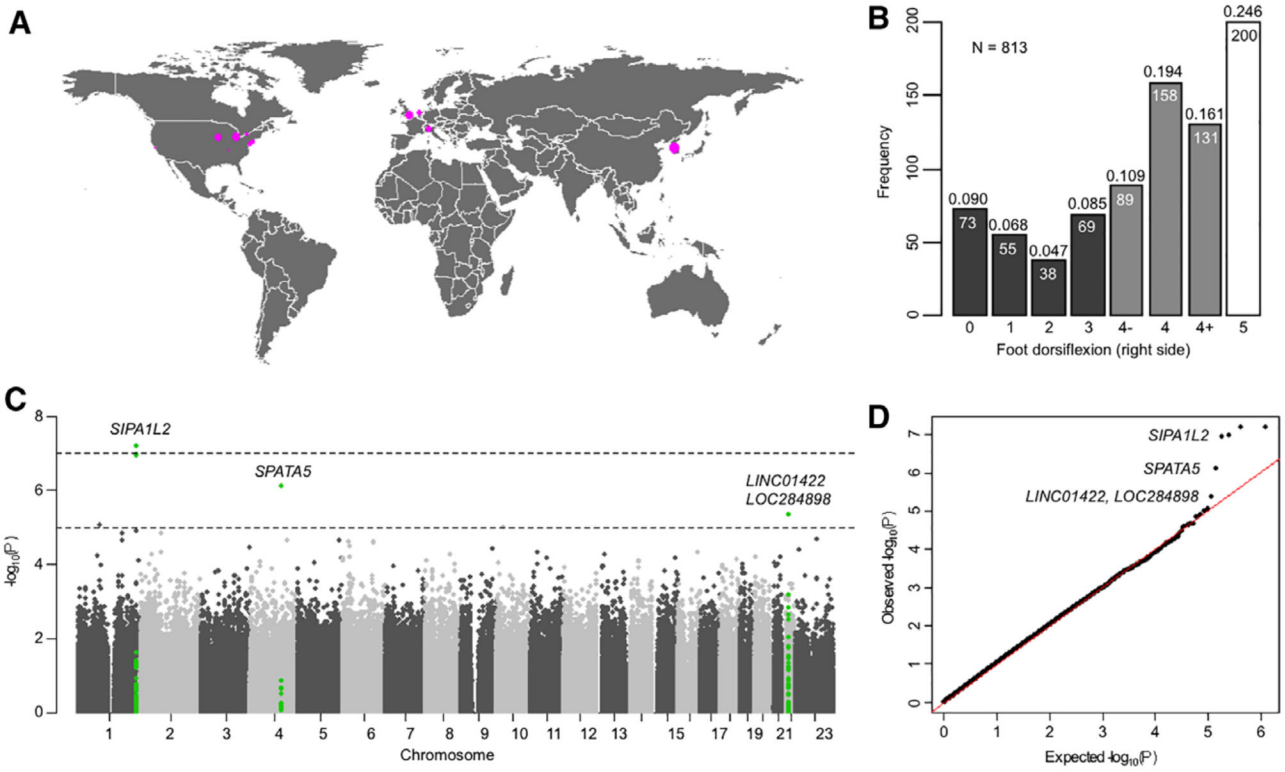

FIGURE 1:

Genome-wide association analysis of foot dorsiflexion in Charcot-Marie-Tooth disease type 1A patients. (A) Locations of participating clinical sites. The study includes 971 samples from 12 clinical sites worldwide. The size of the circles is proportional to the number of samples from that site. (B) Distribution of Medical Research Council scores for foot dorsiflexion (right side) in n = 813 samples. (C, D) Genome-wide association analysis of foot dorsiflexion in the European cohort (n = 330). (C) Manhattan plot shows genome-wide significance on chromosome 1. Association analysis was performed with the R package GWAF to compare marker allele frequencies between severe cases (foot dorsiflexion = 0–3, n = 183) and mild cases (foot dorsiflexion = 5, n = 147), accounting for family relatedness. The dashed lines show the genome-wide significance threshold ($p = 1 \times 10^{-7}$) and the suggestive significance threshold ($p = 1 \times 10^{-5}$). Association peaks (100 kb around the lead single nucleotide polymorphisms) are highlighted in green. (D) Quantile-quantile plot compares the observed distribution of p values with the expected uniform distribution. Genomic inflation factor (λ) = 1.04.

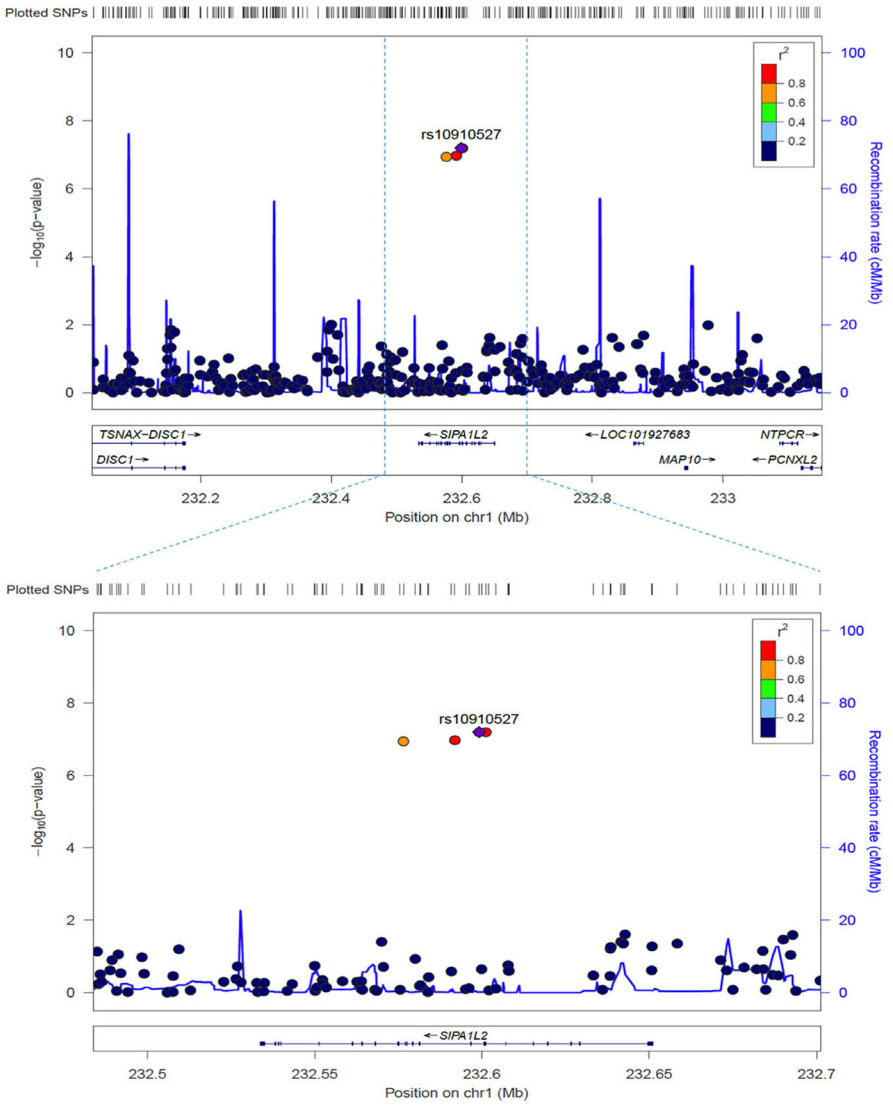

FIGURE 2:

Regional association plots of genome-wide significant signals on chromosome 1. The genome-wide significant single nucleotide polymorphisms (SNPs) are located in the *SIPA1L2* gene. The purple dot shows the most significantly associated SNP (rs10910527 at chr1:232,599,189). The color scheme shows R^2 values of rs10910527 with other variants (1000 Genomes Project, November 2014, EUR).

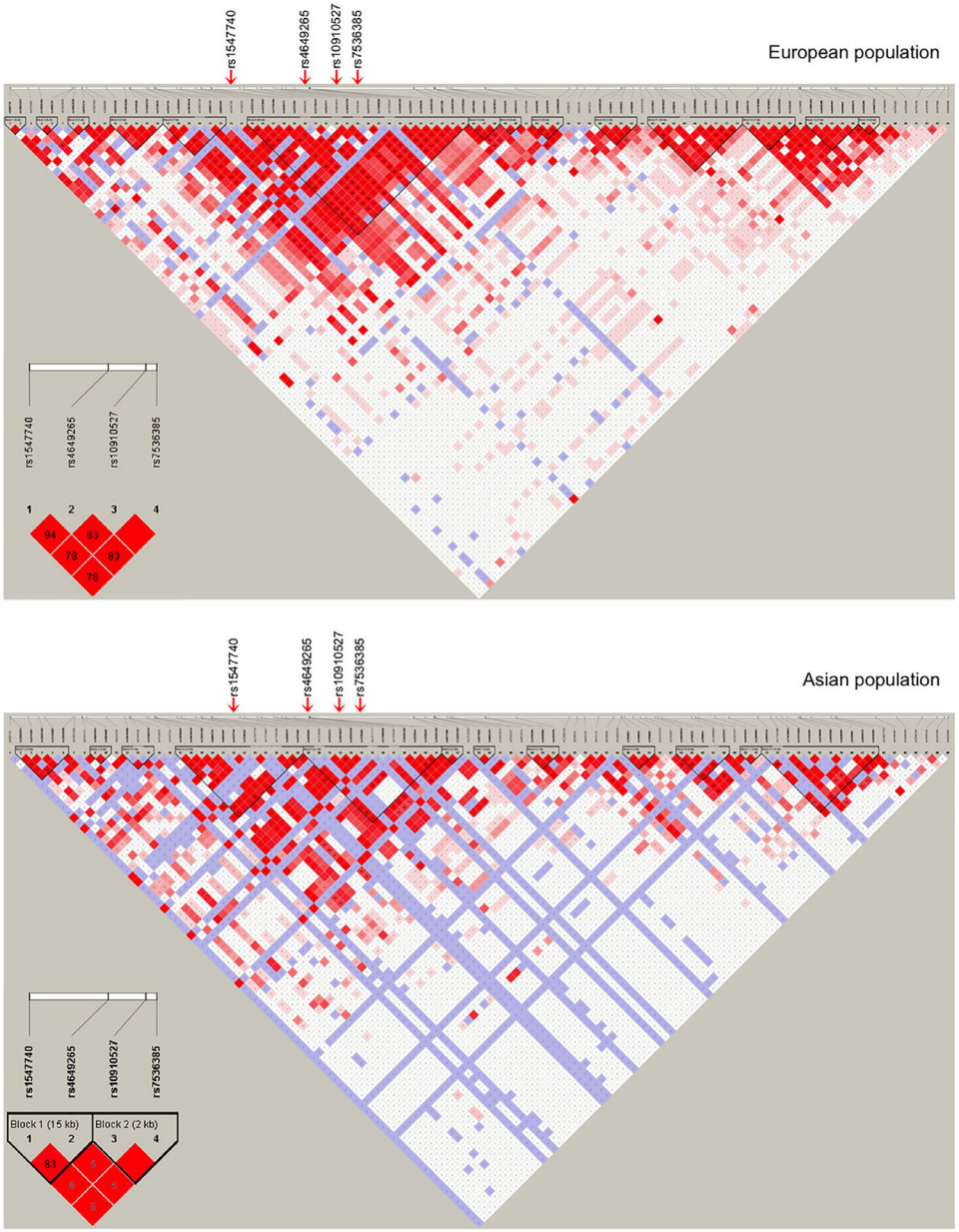

FIGURE 3:

Linkage disequilibrium plots of the *SIPA1L2* locus in the European cohort and the Asian cohort. The color scheme represents D′, and the values show R^2 between single nucleotide polymorphisms.

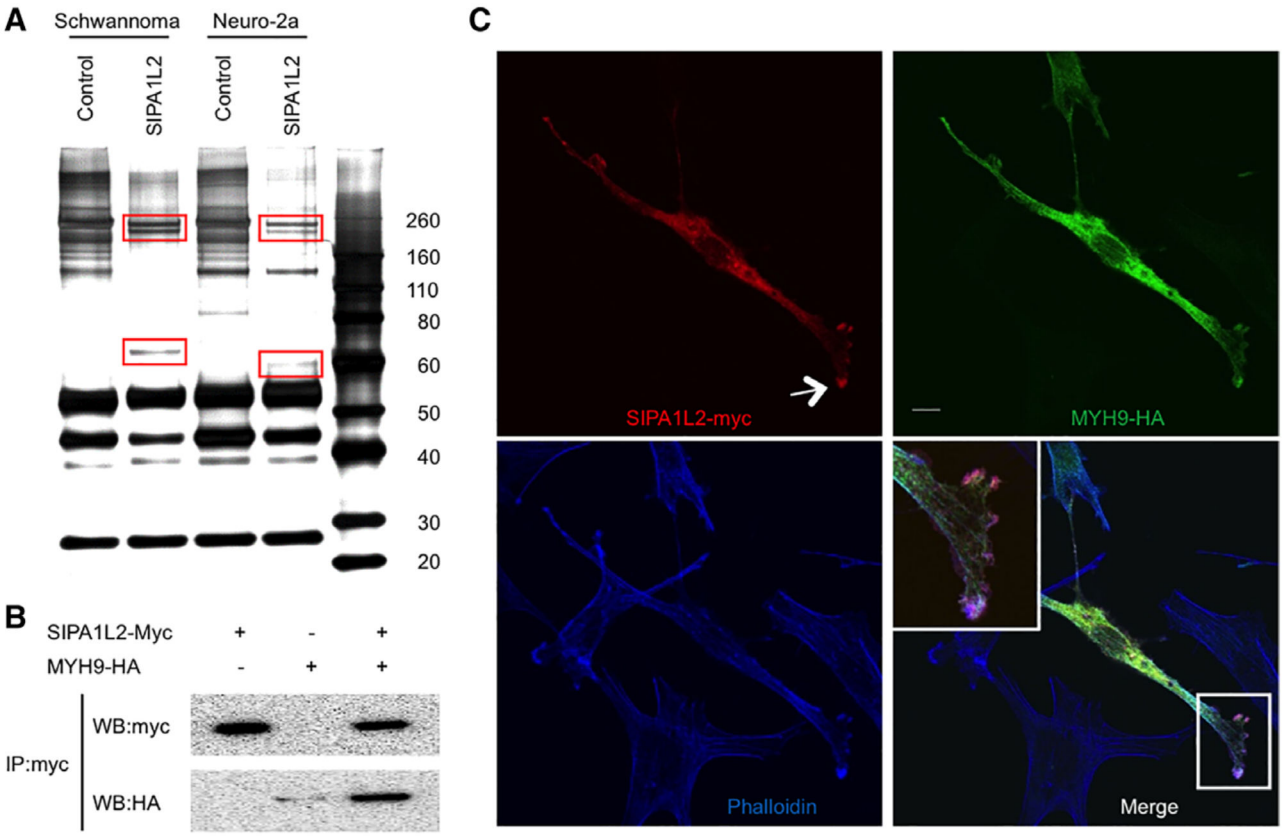

FIGURE 4:

Identification of SIPA1L2 protein binding partners. (A) Silver staining of sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel showing bands from cells immunoprecipitated with SIPA1L2-Myc. Red rectangles indicate excised bands submitted for mass spectrometry analysis. MYH9 and β-actin were identified by mass spectrometry. (B) Coimmunoprecipitation assay in Schwannoma cells cotransfected with *SIPA1L2-Myc* and *MYH9-HA* validated the interaction between SIPA1L2 and MYH9. (C) Confocal images of colocalization between SIPA1L2, MYH9, and F-actin. In Schwannoma cells triple-stained with anti-Myc (red), anti-HA (green), and phalloidin (cyan), colocalization (merge) was enriched in the tips of growth cones. The insert highlights their colocalization.

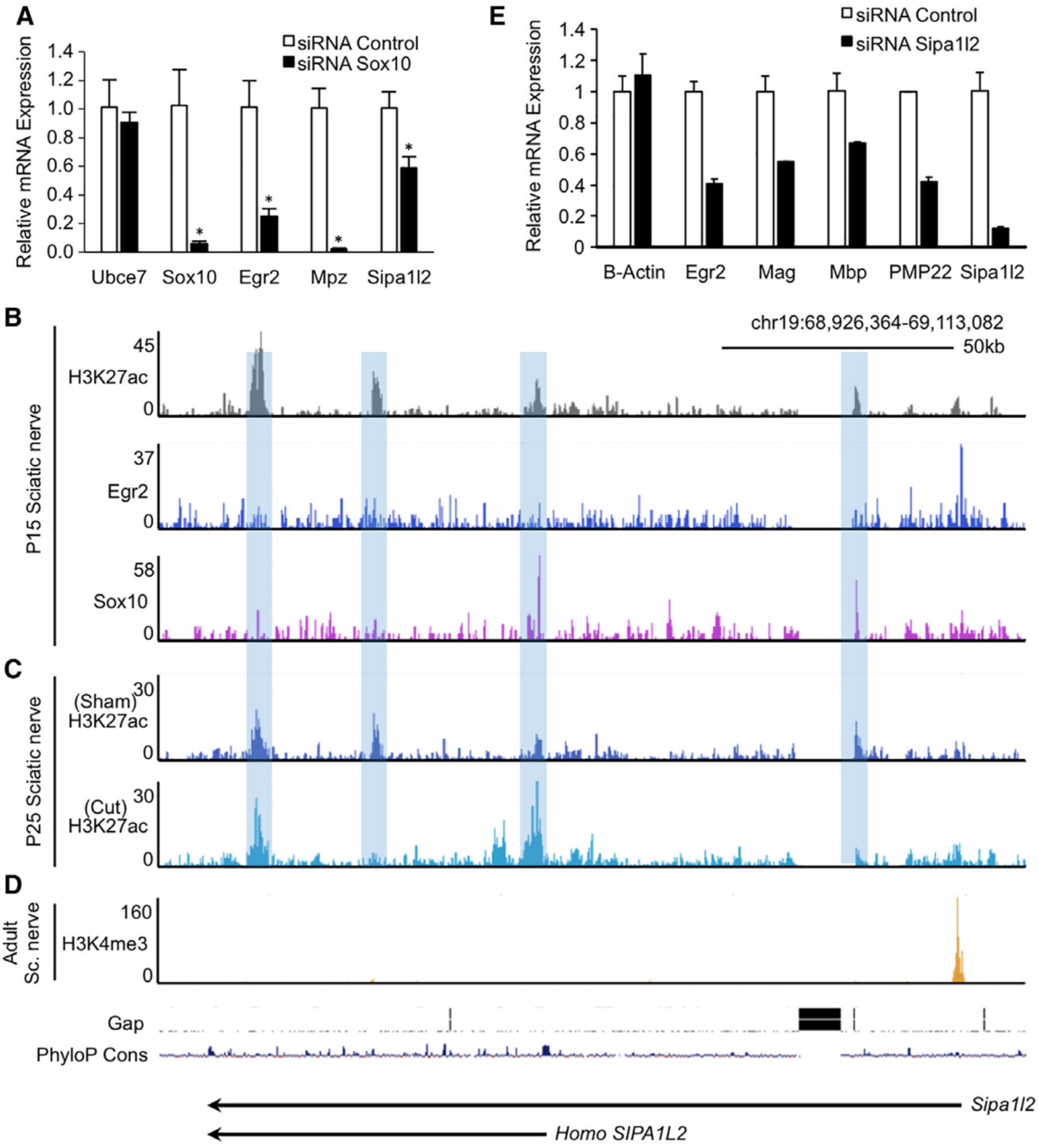

FIGURE 5:

Putative regulatory elements in the *Sipa1l2* gene locus. (A) S16 cells were transfected with Sox10 siRNA, and quantitative real-time polymerase chain reaction (RT-PCR) was performed to assess levels of Sipa1l2 mRNA (n = 3) normalized to B-actin mRNA. Ubce7 mRNA was measured as a negative control, along with the mRNA for the myelin-related genes *Egr2* and *Mpz*. Probability values are from the Student 2-tailed *t* test. *Statistically significant. (B) ChIP-Seq profiles at the *Sipa1l2* gene reveal regions bound by EGR2 and SOX10[21,23] and H3K27ac-marked enhancers in P15 rat sciatic nerve. (C) ChIP-Seq analysis depicts dynamic changes in the active enhancer mark, histone H3K27 acetylation, in P25 rat sham-injured versus with cut sciatic nerve (3 days after transection).[22] (D) ChIP-Seq

analysis of H3K4me3 enrichment in adult sciatic nerve[25] identifies active promoters of expressed genes. Gaps in the rn5 rat genome build are included as a track (black bars), as well as conservation scores (PhyloP Cons). (E) S16 cells were transfected with siRNA for Sipa1l2, and the mRNA levels of several myelin genes were determined by quantitative RT-PCR.

TABLE 1.

Characteristics of Patients Enrolled in the Study

Characteristic	European Cohort	Asian Cohort	Extreme-Phenotype Subset, European Population
Number of patients	734	237	330
Female participants, n (%)	407 (55.4%)	113 (47.7%)	176 (53.3%)
Age at exam, yr, mean (SD)	41.8 (17.4)	36.3 (17.9)	43.2 (17.6)
Age at exam, yr, range	4–86	2–83	5–80
Number of families	550	155	265

TABLE 2.

Most Significant SNPs Associated with Foot Dorsiflexion in European Cohort

SNP	Chr	BP (hg19)	Effect Allele	Ref Allele	Gene	Beta	SE	*p*	OR	95% CI for OR
rs10910527	1	232,599,189	T	C	*SIPA1L2*	2.549	0.471	6.30E-08	12.794	5.083–32.206
rs7536385	1	232,601,208	G	A	*SIPA1L2*	2.549	0.471	6.30E-08	12.794	5.083–32.206
rs4649265	1	232,591,925	C	A	*SIPA1L2*	2.387	0.449	1.06E-07	10.881	4.513–26.234
rs1547740	1	232,576,594	T	G	*SIPA1L2*	2.385	0.450	1.13E-07	10.859	4.495–26.233
rs303143	4	124,155,197	G	A	*SPATA5*	−3.407	0.689	7.60E-07	0.033	0.009–0.128
rs2213767	22	27,370,782	T	C		1.396	0.304	4.37E-06	4.039	2.226–7.329
rs641712	1	86,488,232	A	G	*COL24A1*	−0.881	0.198	8.51E-06	0.414	0.281–0.611
rs11007056	10	28,652,892	A	G		−1.036	0.234	9.88E-06	0.355	0.224–0.561
rs16859600	1	233,928,834	A	G		1.285	0.294	1.24E-05	3.615	2.031–6.432
rs10158853	1	177,940,612	G	A		−1.036	0.239	1.46E-05	0.355	0.222–0.567
rs2366812	2	85,954,544	G	T		−0.917	0.212	1.46E-05	0.400	0.264–0.606
rs5924090	23	86,954,001	A	C		1.025	0.241	2.14E-05	2.787	1.738–4.470
rs10913485	1	177,965,223	C	T		−1.016	0.240	2.22E-05	0.362	0.226–0.579
rs17066814	5	165,825,034	G	A		1.640	0.387	2.24E-05	5.155	2.414–11.007
rs17025973	4	150,168,192	C	T		1.934	0.457	2.34E-05	6.917	2.824–16.941
rs17026072	4	150,202,270	G	A		1.934	0.457	2.34E-05	6.917	2.824–16.941
rs7743652	6	25,502,076	A	G	*LRRC16A*	0.764	0.181	2.45E-05	2.147	1.506–3.061
rs482014	6	125,365,295	T	C	*RNF217*	0.915	0.217	2.52E-05	2.497	1.632–3.82
rs9375383	6	125,354,154	C	A	*RNF217*	0.939	0.223	2.64E-05	2.557	1.652–3.959
rs3094204	6	31,091,992	T	C	*PSORS1C1*	0.726	0.175	3.32E-05	2.067	1.467–2.912
rs1705984	3	194,266,725	A	C		−1.035	0.251	3.61E-05	0.355	0.217–0.581
rs1536510	9	126,935,882	T	C		1.079	0.262	3.75E-05	2.942	1.76–4.916
rs5909373	23	17,566,980	A	C	*NHS*	−0.795	0.194	4.02E-05	0.452	0.309–0.66
rs16955656	16	81,644,493	T	G	*CMIP*	1.498	0.368	4.68E-05	4.473	2.174–9.201
rs1341296	11	29,363,716	C	T		−0.822	0.202	4.70E-05	0.440	0.296–0.653
rs6759087	2	85,948,793	A	G		−1.016	0.250	4.74E-05	0.362	0.222–0.591

BP = base pair; CI = confidence interval; OR = odds ratio; SE = standard error; SNP = single nucleotide polymorphism.

TABLE 3.

Genotype of Top Associated SNP rs10910527 in the European Cohort

	rs10910527 Genotype			
Foot Dorsiflexion	**C/C**	**C/T**	**MAF**	**1000 Genomes, EUR MAF**
0–3	165	18	0.0492	0.0398
4	208	13	0.0294	
5	145	2	0.0068	

MAF = minor allele frequency; SNP = single nucleotide polymorphism.

TABLE 4.

RegulomeDB Scores of SNPs in Linkage Disequilibrium with rs10910527 in 1000 Genomes EUR

SNP	Genomic Position	RegulomeDB Score	R^2 with rs10910527
rs10910527[a]	chr1:232599189	5	1
rs12044960	chr1:232598925	6	1
rs1326291	chr1:232599550	5	1
rs6674476	chr1:232600403	5	1
rs7536385[a]	chr1:232601208	6	1
rs7531995	chr1:232601291	6	1
rs12568689	chr1:232601355	7	1
rs4233063	chr1:232597629	5	0.9221
rs2148327	chr1:232597084	7	0.9221
rs2357067	chr1:232594857	5	0.9221
rs59313883	chr1:232593475	5	0.9221
rs5781697	chr1:232592930	5	0.9221
rs10797565	chr1:232593206	6	0.8963
rs4649378	chr1:232594264	5	0.871
rs59806150	chr1:232599226	5	0.8448
rs7529880	chr1:232598503	7	0.8448
rs6695556	chr1:232592298	6	0.8448
rs4649265[a]	chr1:232591925	6	0.8448
rs58383309	chr1:232603970	4	0.8195
rs6662895	chr1:232590092	3a	0.8195
rs11586498	chr1:232584375	6	0.8195
rs12410125	chr1:232598385	5	0.7934
rs4649266	chr1:232606283	5	0.7444
rs4649268	chr1:232606593	5	0.7444
rs10495327	chr1:232608616	2b	0.7444
rs16857370	chr1:232618974	5	0.7444
rs1536324	chr1:232583227	4	0.7426
rs57063260	chr1:232584462	5	0.7423
rs12036020	chr1:232576998	5	0.7171
rs1547740[a]	chr1:232576594	2b	0.6914
rs75068467	chr1:232597527	5	0.6407
rs1359628	chr1:232596004	5	0.6407

[a]SNPs associated with foot dorsiflexion strength in genome-wide association studies.

SNP = single nucleotide polymorphism.