FIGURE 4: Electrophysiological recordings of the Kv4.3 channels. Whole cell patch-clamp recordings revealed endogenous currents in untransfected cells (A), large rapidly inactivating A-type potassium current in wild-type (WT) Kv4.3-transfected cells (B), and reduced current amplitudes in p.F227del Kv4.3-transfected cells (C). WT *(open black circles)* had a significantly larger current density than p.F227del *(red circles; ***p < 0.0005)*, and the current density of p.F227del was similar to that of controls (Ctrl; *filled black circles;* D).

NM_004980) is the third voltage-gated potassium channel gene discovered to cause human cerebellar ataxia. Point mutations[28–33] in *Shaker* Kv1.1-encoding *KCNA1* cause episodic ataxia type 1 (OMIM160120). Point mutations in *Shaw*-related Kv3.3-encoding *KCNC3* cause SCA13 (OMIM605259), with phenotypes ranging from neurodevelopmental disorders to adult onset neurodegeneration.[34,35] Although the exact mechanism is not yet known how mutations may disrupt Kv4.3 function in regulating neuronal excitability and how this may in turn lead to neurodegeneration, the discovery of *KCND3* as the causative gene in SCA19/22 adds to the growing evidence of the importance of fine tuning neuronal excitability in the health and survival of neurons and especially cerebellar neurons. Other voltage-gated K$^+$ channels have been proposed in neurodegenerative diseases.[36–38]

Patients with p.F227del in Kv4.3 all have a protracted clinical course, with slowly progressive cerebellar ataxia. As proposed, the phenotypes of dominant cerebellar ataxias due to mutations other than polyglutamine expansions are slow and less complicated despite earlier age at onset.[4] There are clinical differences between families A and B. The Chinese family presents with relatively pure cerebellar signs, which are consistent with the clinical classification of autosomal dominant cerebellar ataxia (ADCA) type III,[39] similar to SCA6, whereas the French family presents with signs consistent with ADCA type I.[39] The clinical progression of SCA22 appears even slower than that of SCA6, as evaluated with the clinical rating scale of ataxia SARA.[9] For individuals carrying the p.F227del mutation, the onset of ataxia was earlier (15–30 years) in the younger generation than in the older (30–50 years), which is also true for those with the p.G345V mutation, although the age at onset was later (50s in the older generation, compared to 30s–40s in the younger generation), suggesting that the G345V-associated phenotype may be milder. Of note, the disease is almost fully penetrant in the 6 families described in this report. It is our hope and expectation that additional patients and families will be identified with SCA19/22 to further define the clinical features.

assistance with specific experiments; and J. Xu for bioinformatics.

References

1. Klockgether T. Update on degenerative ataxias. Curr Opin Neurol 2011;24:339–345.

2. Soong B- W, Paulson HL. Spinocerebellar ataxias: an update. Curr Opin Neurol 2007;20:438–446.

3. Soong BW, LU YC, Choo KB, Lee HY. Frequency analysis of autosomal dominant ataxias in Taiwanese patients and clinical and molecular characterization of spinocerebellar ataxia type 6. Arch Neurol 2001;58:1105–1109.

4. Durr A. Autosomal dominant cerebellar ataxias: polyglutamine expansions and beyond. Lancet Neurol 2010;9:885–894.

5. Chung MY, Lu YC, Cheng NC, Soong BW. A novel autosomal dominant spinocerebellar ataxia (SCA22) linked to chromosome 1p21-q23. Brain 2003;126:1293–1299.

6. Verbeek DS, Schelhaas JH, Ippel EF, et al. Identification of a novel SCA locus (SCA19) in a Dutch autosomal dominant cerebellar ataxia family on chromosome region 1p21-q21. Hum Genet 2002; 111:388–393.

7. Schelhaas HJ, Verbeek DS, Van de Warrenburg BP, Sinke RJ. SCA19 and SCA22: evidence for one locus with a worldwide distribution. Brain 2004;127:E6; author reply E7.

8. Schmitz-Hubsch T, du Montcel ST, Baliko L, et al. Scale for the assessment and rating of ataxia: development of a new clinical scale. Neurology 2006;66:1717–1720.

9. Lee YC, Liao YC, Wang PS, et al. Comparison of cerebellar ataxias: a three-year prospective longitudinal assessment. Mov Disord 2011;26:2081–2087.

10. Sambrook J, Russell DW. Molecular cloning: a laboratory manual. 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2001.

11. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002;30:97–101.

12. 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.

13. Durbin RM, Altshuler D, Abecasis GR, et al. A map of human genome variation from population-scale sequencing. Nature 2010; 467:1061–1073.

14. Postma AV, Bezzina CR, de Vries JF, et al. Genomic organisation and chromosomal localisation of two members of the KCND ion channel family, KCND2 and KCND3. Hum Genet 2000;106: 614–619.

15. Norris AJ, Nerbonne JM. Molecular dissection of I(A) in cortical pyramidal neurons reveals three distinct components encoded by Kv4.2, Kv4.3, and Kv1.4 alpha-subunits. J Neurosci 2010;30: 5092–5101.

16. Niwa N, Nerbonne JM. Molecular determinants of cardiac transient outward potassium current (I(to)) expression and regulation. J Mol Cell Cardiol 2010;48:12–25.

17. Lee YC, Yu CT, Lin KP, et al. MPZ mutation G123S characterization: evidence for a complex pathogenesis in CMT disease. Neurology 2008;70:273–277.

18. Shi G, Kleinklaus AK, Marrion NV, Trimmer JS. Properties of Kv2.1 K+ channels expressed in transfected mammalian cells. J Biol Chem 1994;269:23204–23211.

19. Tsaur ML, Chou CC, Shih YH, Wang HL. Cloning, expression and CNS distribution of Kv4.3, an A-type K+ channel alpha subunit. FEBS Lett 1997;400:215–220.

20. Kong W, Po S, Yamagishi T, et al. Isolation and characterization of the human gene encoding Ito: further diversity by alternative mRNA splicing. Am J Physiol 1998;275:H1963–H1970.

21. Dilks D, Ling HP, Cockett M, et al. Cloning and expression of the human kv4.3 potassium channel. J Neurophysiol 1999;81: 1974–1977.

22. Isbrandt D, Leicher T, Waldschutz R, et al. Gene structures and expression profiles of three human KCND (Kv4) potassium channels mediating A-type currents I(TO) and I(SA). Genomics 2000; 64:144–154.

23. Hsu YH, Huang HY, Tsaur ML. Contrasting expression of Kv4.3, an A-type K+ channel, in migrating Purkinje cells and other post-migratory cerebellar neurons. Eur J Neurosci 2003;18:601–612.

24. Oberdick J, Baader SL, Schilling K. From zebra stripes to postal zones: deciphering patterns of gene expression in the cerebellum. Trends Neurosci 1998;21:383–390.

25. Giudicessi JR, Ye D, Tester DJ, et al. Transient outward current (I(to)) gain-of-function mutations in the KCND3-encoded Kv4.3 potassium channel and Brugada syndrome. Heart Rhythm 2011;8: 1024–1032.

26. Tombola F, Pathak MM, Gorostiza P, Isacoff EY. The twisted ion-permeation pathway of a resting voltage-sensing domain. Nature 2007;445:546–549.

27. Khalili-Araghi F, Tajkhorshid E, Roux B, Schulten K. Molecular dynamics investigation of the omega-current in the Kv1.2 voltage sensor domains. Biophys J 2012;102:258–267.

28. D'Adamo MC, Liu Z, Adelman JP, et al. Episodic ataxia type-1 mutations in the hKv1.1 cytoplasmic pore region alter the gating properties of the channel. EMBO J 1998;17:1200–1207.

29. D'Adamo MC, Imbrici P, Sponcichetti F, Pessia M. Mutations in the KCNA1 gene associated with episodic ataxia type-1 syndrome impair heteromeric voltage-gated K(+) channel function. FASEB J 1999;13:1335–1345.

30. Eunson LH, Rea R, Zuberi SM, et al. Clinical, genetic, and expression studies of mutations in the potassium channel gene KCNA1 reveal new phenotypic variability. Ann Neurol 2000;48:647–656.

31. Rea R, Spauschus A, Eunson LH, et al. Variable K(+) channel subunit dysfunction in inherited mutations of KCNA1. J Physiol 2002; 538:5–23.

32. Imbrici P, D'Adamo MC, Kullmann DM, Pessia M. Episodic ataxia type 1 mutations in the KCNA1 gene impair the fast inactivation properties of the human potassium channels Kv1.4-1.1/Kvbeta1.1 and Kv1.4-1.1/Kvbeta1.2. Eur J Neurosci 2006;24:3073–3083.

33. Shook SJ, Mamsa H, Jen JC, et al. Novel mutation in KCNA1 causes episodic ataxia with paroxysmal dyspnea. Muscle Nerve 2008;37:399–402.

34. Waters MF, Minassian NA, Stevanin G, et al. Mutations in voltage-gated potassium channel KCNC3 cause degenerative and developmental central nervous system phenotypes. Nat Genet 2006;38:447–451.

35. Figueroa KP, Waters MF, Garibyan V, et al. Frequency of KCNC3 DNA variants as causes of spinocerebellar ataxia 13 (SCA13). PLoS One 2011;6:e17811.

36. Ariano MA, Cepeda C, Calvert CR, et al. Striatal potassium channel dysfunction in Huntington's disease transgenic mice. J Neurophysiol 2005;93:2565–2574.

37. Angulo E, Noe V, Casado V, et al. Up-regulation of the Kv3.4 potassium channel subunit in early stages of Alzheimer's disease. J Neurochem 2004;91:547–557.

38. Baranauskas G, Tkatch T, Surmeier DJ. Delayed rectifier currents in rat globus pallidus neurons are attributable to Kv2.1 and Kv3.1/3.2 K(+) channels. J Neurosci 1999;19:6394–6404.

39. Harding AE. Clinical features and classification of inherited ataxias. Adv Neurol 1993;61:1–14.

The Neurogenomics View of Neurological Diseases

Shoji Tsuji, MD, PhD

The availability of high-throughput genome sequencing technologies is expected to revolutionize our understanding of not only hereditary neurological diseases but also sporadic neurological diseases. The molecular bases of sporadic diseases, particularly those of sporadic neurodegenerative diseases, largely remain unknown. As potential molecular bases, various mechanisms can be considered, which include those underlying apparently sporadic neurological diseases with low-penetrant mutations in the gene for hereditary diseases, sporadic diseases with de novo mutations, and sporadic diseases with variations in disease-susceptible genes. With unprecedentedly robust power, high-throughput genome sequencing technologies will enable us to explore all of these possibilities. These new technologies will soon be applied in clinical practice. It will be a new era of datacentric clinical practice.

The elucidation of the molecular bases of neurological diseases is fundamental to the development of disease-modifying and preventive therapies.[1] Over the past 3 decades, we have witnessed remarkable progress in the identification of the genes that cause hereditary neurological diseases (**Figure 1**).[2-4] This has been accomplished mainly on the basis of the research paradigm known as "positional cloning,"[5,6] which uses linkage studies to pinpoint the position of genes on chromosomes followed by the identification of the causative gene. The identification of causative genes has further made it possible to develop disease models for hereditary neurological diseases[7-10] and to develop therapeutic strategies.[11]

The majority of neurological diseases, however, are sporadic without any obvious familial occurrence. We are thus faced with the challenge of elucidating the molecular bases of sporadic diseases. Intriguingly, the clinical presentations and neuropathological findings of hereditary forms of neurodegenerative diseases are often indistinguishable from those of sporadic diseases, raising the possibility that common pathophysiologic pathways underlie both hereditary and sporadic neurodegenerative diseases.

In contrast to the molecular bases of hereditary neurological diseases, the molecular bases of sporadic neurological diseases, particularly those of sporadic neurodegenerative diseases, largely remain unknown. A potential clue to the molecular bases of sporadic neurological diseases may be the clinical observation that siblings and relatives of a patient with a neurological disease are at an increased risk of developing the same disease; this phenomenon has been observed with regard to Parkinson disease (PD)[12] and amyotrophic lateral sclerosis.[13] These clinical observations suggest the involvement of genetic factors in these diseases (Figure 1). Until recently, it has been difficult to elucidate the genetic factors underlying sporadic neurological diseases. Rapid advancements in genome science, particularly the availability of massively parallel sequencing technologies that use next-generation sequencers (NGSs), are revolutionizing the neurogenomics view of sporadic neurological diseases. The elucidation of the genomic variants underlying sporadic diseases is expected to provide some answers that will help us to develop disease-modifying and preventive therapies.

Author Affiliations: Department of Neurology, Graduate School of Medicine, University of Tokyo, and Medical Genome Center, University of Tokyo Hospital, Japan.
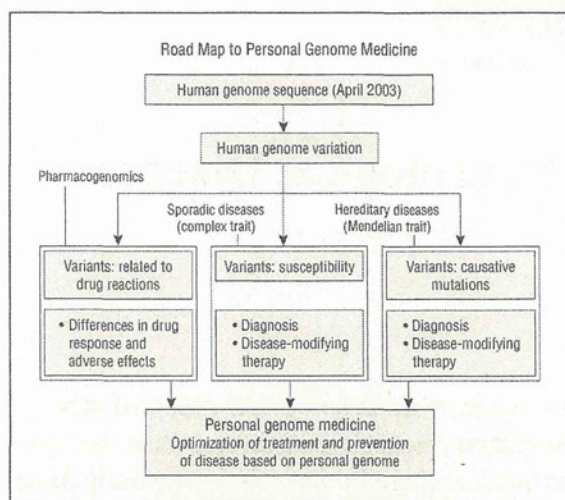
Figure 1. Diagram showing the road map to personal genome medicine. Since the completion of the human genome sequence in 2003, the research focus in human genetics has moved to how human genome variations affect human health. Human genome variations are considered to be associated not only with hereditary diseases but also with sporadic diseases. In addition, human genome variations are also associated with differences in drug responses and adverse effects. Optimization of treatment and prevention based on personal genome information will soon be a realistic paradigm in clinical practice.

Another important field is pharmacogenomics, in which genomic variations underlie differences in drug responses and adverse drug effects (Figure 1). This field is currently being introduced into clinical practice.

Thus, it will be essential to better understand how human genome variations affect our health with regard to diseases with Mendelian or complex traits, as well as with regard to pharmacogenomics. Herein, the neurogenomics view of neurological diseases and the future directions of clinical practice are discussed.

HIGH-THROUGHPUT GENOME SEQUENCING TECHNOLOGIES

Emerging new technologies for nucleotide sequencing have brought about a remarkable revolution in analyses of the human genome sequence. Compared with a conventional technology (namely, the Sanger method),[14,15] the throughput of massively parallel sequencing that uses NGSs[16] is increasing dramatically, with the current throughput at 600 GB per run, which means that a sufficient amount of sequence data can be obtained for whole-genome sequencing of at least 4 individuals.[17] In typical experiments, billions of short reads (100-150 base pairs [bp]) are obtained. These short reads are aligned to human genome reference sequences, and sequence variations are called through computational analyses.

Currently, 2 types of sequencing strategy (namely, whole-exome and whole-genome sequence analyses) are used. Because the cost of whole-genome sequencing is still considerably high, it is not easy to conduct whole-genome sequencing for a large number of individuals. In whole-exome sequence analysis, the enrichment of exonic sequences using oligonucleotide "baits," which is followed by sequencing, has been preferentially used. With this strategy, all exonic sequences in the human genome can be efficiently enriched.[18-20] With this approach, more than 90% of target regions can be enriched, and these enriched genomic regions are then subjected to massively parallel sequencing using NGSs. This approach is currently being used a lot for the identification of disease-relevant variants[21-31] and even for diagnostic purposes.[32-35]

Given the ever-increasing throughput of NGSs and the dramatically decreasing costs, it will soon be a realistic approach to conduct whole-genome sequencing for various research applications (Figure 2).[36-40] Studies have shown that there are more than 3 million variations in the human genome of each individual. In one study,[40] among the 3.3 million single-nucleotide polymorphisms (SNPs), 8996 known nonsynonymous SNPs and 1573 novel nonsynonymous SNPs were identified. Interestingly, 32 alleles exactly matched mutations previously registered in the Human Gene Mutation Database. In addition, 345 insertions/deletions were observed to overlap in a coding sequence and may alter protein function.[40] These findings indicate that, among the numerous candidate variations, it will be a challenge to determine which variations are relevant to diseases.

Given the enormous number of short read sequences (~100 bp), informatics analyses, including mapping to reference sequences and indentifying variations, require a huge computational power.[41-45] Furthermore, mutations can be variable, including single base substitutions, insertions/deletions, and structural variations. It is difficult to efficiently identify all the variations using currently available NGSs and software. For example, expansions of repeat motifs identified in frontotemporal dementia and amyotrophic lateral sclerosis[46] are difficult to identify using NGSs.

As already stated, most of the currently available NGSs produce billions of short reads of 100 to 150 bp. This is the limitation in analyzing various structural variations, some of which may be relevant to neurological diseases. Very recently, single-molecule sequencing technology has become available from Pacific Biosciences; this type of technology enables the acquisition of nucleotide sequences as large as 10 kilobases.[47,48] Another single-molecule sequencing technology using nanopores, which allows for the acquisition of much longer sequences,[49] will soon become available.

EFFECT OF HIGH-THROUGHPUT GENOME SEQUENCING ON UNDERSTANDING THE MOLECULAR BASES OF HEREDITARY NEUROLOGICAL DISEASES

The strategies for identifying causative genes for hereditary diseases have been well established.[5,6] The chromosomal localization of the disease-causing genes is pinpointed by linkage analysis using polymorphic DNA markers.[50-52] Although a number of genes have been identified by applying these technologies, more than 50% of the genes causing familial amyotrophic lateral sclerosis remain to be identified.[53] In families with hereditary diseases, the availability of affected and unaffected individuals is often limited owing to small family sizes and the small number of family members with a confirmed clinical and/or a pathological diagnosis. These circumstances pose a challenge to positional cloning because the candidate regions cannot be narrowed down to small regions that are sufficient for identifying the
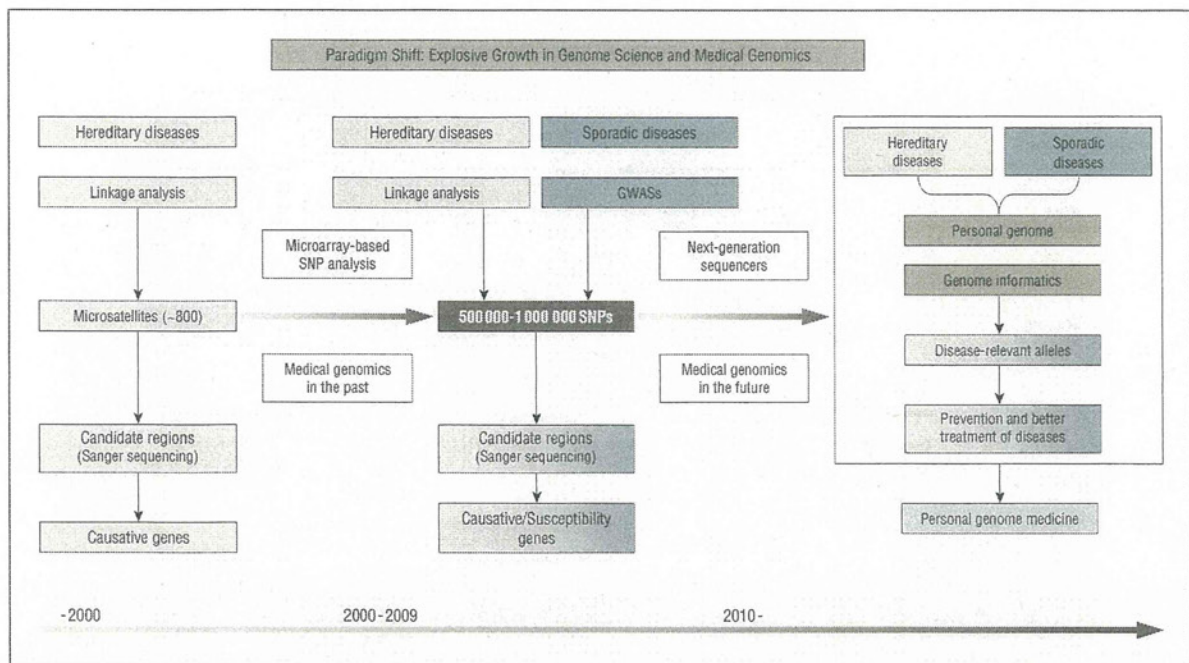
Figure 2. Diagram showing the paradigm shift (ie, the explosive growth in genome science and medical genomics). Over the past decade, genome-wide association studies (GWASs) using common single-nucleotide polymorphisms (SNPs) have been conducted to identify genomic variations in sporadic neurological diseases. The theoretical framework of GWASs is the common disease–common variants hypothesis. Although GWASs have successfully revealed numerous susceptibility genes for common diseases such as diabetes mellitus, as well as neurodegenerative diseases, the odds ratios associated with these risk alleles are generally low and account for only a small proportion of estimated heritability. The availability of high-throughput genome sequencing technologies will enable us to identify all the genomic variants, and eventually those of disease-relevant alleles based on the common disease–multiple rare variants hypothesis.

causative genes by sequencing individual genes in the candidate regions. Despite these difficult circumstances, the availability of NGSs with unbelievably high throughput has made the identification of causative genes possible.[31,54,55] Given the large capacity of NGSs, the most essential step (and the bottleneck) is now the collection of as many samples from patients and their families as possible based on well-characterized clinical information, including the correct diagnosis, regardless of family size or number.

EFFECT OF HIGH-THROUGHPUT GENOME SEQUENCING ON UNDERSTANDING THE MOLECULAR BASES OF SPORADIC NEUROLOGICAL DISEASES

The elucidation of the molecular bases of sporadic neurological diseases is now a big challenge. We need to take various mechanisms into account as the molecular bases of sporadic neurological diseases, which include (1) apparently sporadic diseases with low-penetrant mutations in the gene for hereditary diseases, (2) sporadic diseases with de novo mutations, (3) sporadic diseases with variations in disease-susceptible genes, and (4) sporadic diseases with other mechanisms. These different molecular bases are reviewed.

APPARENTLY SPORADIC NEUROLOGICAL DISEASES WITH LOW-PENETRANT MUTATIONS IN THE GENE FOR HEREDITARY DISEASES

There are numerous examples of low-penetrant mutations in apparently sporadic cases of neurological diseases. Sporadic cases of amyotrophic lateral sclerosis due

to low-penetrant *SOD1* mutations have been well characterized.[56-61] In prion diseases, patients with V180I or M232R mutations in the prion protein (*PRNP*) gene rarely have a family history of prion diseases, indicating that these patients are usually diagnosed as having sporadic Creutzfeldt-Jakob disease.[62]

SPORADIC NEUROLOGICAL DISEASES WITH DE NOVO MUTATIONS

Alternating hemiplegia of childhood is a rare neurological disorder characterized by early-onset episodes of hemiplegia, dystonia, various paroxysmal symptoms, and developmental impairments. Almost all cases are sporadic, but the concordance of alternating hemiplegia of childhood in monozygotic twins and the dominant transmission in a family with a milder phenotype have been reported. With this background information, Rosenwich et al[63] conducted whole-exome sequencing of 3 proband-parent trios to identify a disease-associated gene and then examined whether mutations in the gene were also present in the remaining patients and their healthy parents. Whole-exome sequencing indeed showed 3 heterozygous de novo missense mutations.[63] Similar approaches have been used for a number of diseases, including severe epileptic encephalopathy,[64] autism, and schizophrenia.[65] The rationale for these approaches is based on the hypothesis that patients with severe phenotypes associated with reduced reproductive fitness may harbor de novo mutations.[65,66]

Twin studies in which differences in the phenotypes of monozygotic and dizygotic twins were compared have
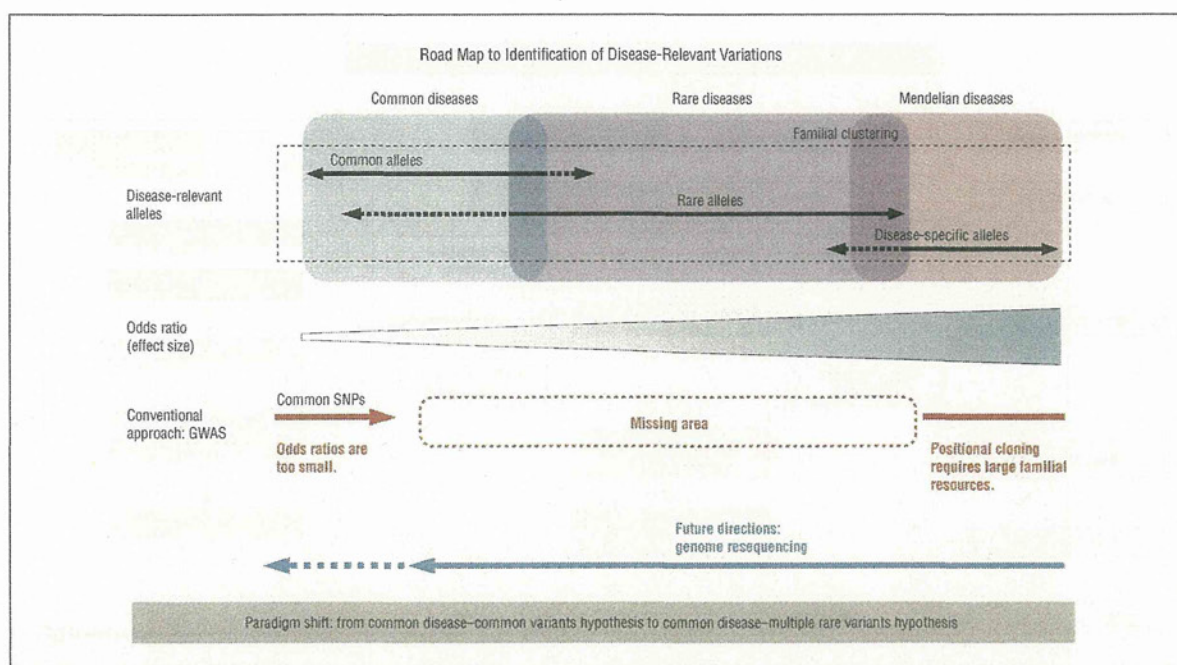
Figure 3. Diagram showing the road map to the identification of disease-relevant variations. Shifting the paradigm from the common disease–common variants hypothesis to the common disease–multiple rare variants hypothesis will lead to the elucidation of the molecular bases of sporadic neurological diseases. Relatively rare sporadic neurological diseases will be good candidates for identifying disease-relevant alleles with large effect sizes because, depending on the effect sizes, the sample sizes can be small.

long been conducted to delineate the involvement of genetic factors. Therefore, the comparison of whole-genome sequences of discordant monozygotic twins is expected to accelerate the discovery of genomic variations responsible for the disease phenotypes.[67,68]

SPORADIC NEUROLOGICAL DISEASES WITH VARIATIONS IN DISEASE-SUSCEPTIBILITY GENES

Over the past decade, genome-wide association studies (GWASs) using common SNPs have been conducted to identify genomic variations associated with sporadic neurological diseases. The theoretical framework of GWASs is the "common disease–common variants" hypothesis, in which common diseases are attributable in part to allelic variants present in more than 5% of the population.[69-71] Although GWASs have successfully revealed numerous susceptibility genes for common diseases such as diabetes mellitus, as well as neurodegenerative diseases, the odds ratios associated with these risk alleles are generally low and account for only a small proportion of estimated heritability.[72-75]

In GWASs, the general finding that the odds ratios associated with risk alleles identified for disease susceptibility are low indicates that GWASs based on the common disease–common variants hypothesis are not effective in identifying genetic risks with large effect sizes. The current experience with GWASs strongly suggests that rarer variants that are difficult to detect by GWASs may account for the "missing" heritability.[17,74] Such rare variants may have large effect sizes as genetic risk factors for diseases. Thus, the paradigm should be shifted from the "com-

mon disease–common variants" hypothesis to the "common disease–multiple rare variants" hypothesis to identify disease-relevant alleles with large effect sizes (**Figure 3**).

An excellent example of rare variants with substantially large effect sizes is the recent discovery of the glucocerebrosidase (*GBA*) gene as a robust genetic risk factor for PD.[76,77] A population-based study[78] coupled with genealogy information demonstrated that the estimated risk ratio for PD for siblings of patients with PD was significantly high, indicating that genetic factors substantially contribute to the development of sporadic PD. Recent clinical observations[79] have suggested the association of sporadic PD with heterozygous mutations in the *GBA* gene encoding the enzyme that is deficient in patients with Gaucher disease, an autosomal recessive lysosomal storage disease. Furthermore, the comorbidity of PD and Gaucher disease was previously described.[80] We conducted an extensive resequencing analysis of *GBA* in patients with PD and controls, and we found that *GBA* variants that are pathogenic for Gaucher disease confer a robust susceptibility to sporadic PD and even account for the familial clustering of PD.[77] The combined carrier frequency of the "pathogenic variants" was as high as 9.4% in patients with PD and significantly higher than that in controls (0.37%), with a markedly high odds ratio of 28.0 (95% CI, 7.3-238.3) for patients with PD compared with controls.

We can draw the following conclusions from the discovery of the major disease-susceptibility gene (*GBA*) with a large effect size: (1) a genetic factor with a large effect size has been discovered in sporadic PD; (2) in accordance with the large effect size, there is a tendency of familial clustering (multiplex families such as affected siblings); and (3) the disease-relevant allele could not be

identified by GWASs using common SNPs and was identified only by nucleotide sequence analysis. These conclusions strongly encourage us to search for disease-susceptibility genes with large effect sizes based on the common disease–multiple rare variants hypothesis. Although the majority of rare missense variants have been suggested to be functionally deleterious in humans,[81] it remains controversial whether a comparison of allele frequencies of rare variants (in particular, missense variants) is a sufficient method for identifying variants associated with diseases. Functional annotation of all the variants obtained by comprehensive genome sequencing will no doubt increase the robust power for detecting significant associations of variants with diseases.

SPORADIC NEUROLOGICAL DISEASES WITH OTHER MECHANISMS

Besides the mechanisms already mentioned, there may be others underlying sporadic neurological diseases. The involvement of somatic mutations occurring in certain cell lineages in sporadic neurological diseases is a potentially interesting mechanism. Such a mechanism in certain types of cancer is well established.[82] The involvement of epigenetics in the development of sporadic neurodegenerative diseases is also a potentially attractive mechanism.[83,84] Recently, there have been an increasing number of studies suggesting that "prion-like" processes (ie, the propagation of misfolded proteins leading to abnormal aggregation) may be involved in the pathogenesis of sporadic neurodegenerative diseases.[85,86] In the field of autoimmune diseases such as multiple sclerosis, the involvement of genetic factors is well characterized. The application of massively parallel sequencing to extensively characterize T-cell receptor repertoires[87,88] and immunoglobulin heavy chain genes,[89] along with sequence-based typing of HLAs,[90,91] will provide new insights into the molecular bases of autoimmune diseases.

As discussed in this review, the availability of robust technologies using NGSs will revolutionize our research paradigms for exploring the molecular bases of hereditary and sporadic neurological diseases. Furthermore, these technologies will soon be applied in clinical practice. It will be a new era of datacentric clinical practice. Are we prepared for this new era?

REFERENCES

1. Trojanowski JQ, Vandeerstichele H, Korecka M, et al; Alzheimer's Disease Neuroimaging Initiative. Update on the biomarker core of the Alzheimer's Disease Neuroimaging Initiative subjects. *Alzheimers Dement.* 2010;6(3):230-238.
2. Gusella JF, Wexler NS, Conneally PM, et al. A polymorphic DNA marker genetically linked to Huntington's disease. *Nature.* 1983;306(5940):234-238.
3. The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell.* 1993;72(6):971-983.
4. Koide R, Ikeuchi T, Onodera O, et al. Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). *Nat Genet.* 1994;6(1):9-13.
5. Collins FS. Positional cloning: let's not call it reverse anymore. *Nat Genet.* 1992;1(1):3-6.
6. Collins FS. Positional cloning moves from perditional to traditional. *Nat Genet.* 1995;9(4):347-350.
7. Sato T, Miura M, Yamada M, et al. Severe neurological phenotypes of Q129 DRPLA transgenic mice serendipitously created by en masse expansion of CAG repeats in Q76 DRPLA mice. *Hum Mol Genet.* 2009;18(4):723-736.
8. Suzuki K, Zhou J, Sato T, et al. DRPLA transgenic mouse substrains carrying single copy of full-length mutant human DRPLA gene with variable sizes of expanded CAG repeats exhibit CAG repeat length- and age-dependent changes in behavioral abnormalities and gene expression profiles. *Neurobiol Dis.* 2012;46(2):336-350.
9. Katsuno M, Adachi H, Kume A, et al. Testosterone reduction prevents phenotypic expression in a transgenic mouse model of spinal and bulbar muscular atrophy. *Neuron.* 2002;35(5):843-854.
10. Mangiarini L, Sathasivam K, Seller M, et al. Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell.* 1996;87(3):493-506.
11. Katsuno M, Banno H, Suzuki K, et al; Japan SBMA Interventional Trial for TAP-144-SR (JASMITT) study group. Efficacy and safety of leuprorelin in patients with spinal and bulbar muscular atrophy (JASMITT study): a multicentre, randomised, double-blind, placebo-controlled trial. *Lancet Neurol.* 2010;9(9):875-884.
12. Sveinbjörnsdottir S, Hicks AA, Jonsson T, et al. Familial aggregation of Parkinson's disease in Iceland. *N Engl J Med.* 2000;343(24):1765-1770.
13. Fang F, Kamel F, Lichtenstein P, et al. Familial aggregation of amyotrophic lateral sclerosis. *Ann Neurol.* 2009;66(1):94-99.
14. Sanger F, Donelson JE, Coulson AR, Kössel H, Fischer D. Use of DNA polymerase I primed by a synthetic oligonucleotide to determine a nucleotide sequence in phage fl DNA. *Proc Natl Acad Sci U S A.* 1973;70(4):1209-1213.
15. Sanger F, Air GM, Barrell BG, et al. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature.* 1977;265(5596):687-695.
16. Metzker ML. Sequencing technologies—the next generation. *Nat Rev Genet.* 2010;11(1):31-46.
17. Tsuji S. Genetics of neurodegenerative diseases: insights from high-throughput resequencing. *Hum Mol Genet.* 2010;19(R1):R65-R70.
18. Okou DT, Steinberg KM, Middle C, Cutler DJ, Albert TJ, Zwick ME. Microarray-based genomic selection for high-throughput resequencing. *Nat Methods.* 2007;4(11):907-909.
19. Fayssoux RS, Tally W, Sanfilippo JA, et al. Spinal injuries after falls from hunting tree stands. *Spine J.* 2008;8(3):522-528.
20. Hodges E, Xuan Z, Balija V, et al. Genome-wide in situ exon capture for selective resequencing. *Nat Genet.* 2007;39(12):1522-1527.
21. Volpi L, Roversi G, Colombo EA, et al. Targeted next-generation sequencing appoints c16orf57 as clericuzio-type poikiloderma with neutropenia gene. *Am J Hum Genet.* 2010;86(1):72-76.
22. Hedges DJ, Burges D, Powell E, et al. Exome sequencing of a multigenerational human pedigree. *PLoS One.* 2009;4(12):e8232.
23. Zaghloul NA, Katsanis N. Functional modules, mutational load and human genetic disease. *Trends Genet.* 2010;26(4):168-176.
24. Biesecker LG. Exome sequencing makes medical genomics a reality. *Nat Genet.* 2010;42(1):13-14.
25. Ng SB, Buckingham KJ, Lee C, et al. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet.* 2010;42(1):30-35.
26. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A.* 2009;106(45):19096-19101.
27. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature.* 2009;461(7261):272-276.
28. Maher B. Exome sequencing takes centre stage in cancer profiling. *Nature.* 2009;459(7244):146-147.
29. Ng PC, Levy S, Huang J, et al. Genetic variation in an individual human exome. *PLoS Genet.* 2008;4(8):e1000160.
30. Ishiura H, Fukuda Y, Mitsui J, et al. Posterior column ataxia with retinitis pigmentosa in a Japanese family with a novel mutation in FLVCR1. *Neurogenetics.* 2011;12(2):117-121.
31. Ishiura H, Sako W, Yoshida M, et al. The TRK-fused gene is mutated in hereditary motor and sensory neuropathy with proximal dominant involvement. *Am J Hum Genet.* 2012;91(2):320-329.
32. Coppola G, Geschwind DH. Genomic medicine enters the neurology clinic. *Neurology.* 2012;79(2):112-114.

33. Sailer A, Scholz SW, Gibbs JR, et al. Exome sequencing in an SCA14 family demonstrates its utility in diagnosing heterogeneous diseases. *Neurology*. 2012; 79(2):127-131.

34. Pierson TM, Adams DA, Markello T, et al; NISC Comparative Sequencing Program. Exome sequencing as a diagnostic tool in a case of undiagnosed juvenile-onset GM1-gangliosidosis. *Neurology*. 2012;79(2):123-126.

35. Landouré G, Sullivan JM, Johnson JO, et al. Exome sequencing identifies a novel TRPV4 mutation in a CMT2C family. *Neurology*. 2012;79(2):192-194.

36. Schuster SC, Miller W, Ratan A, et al. Complete Khoisan and Bantu genomes from southern Africa. *Nature*. 2010;463(7283):943-947.

37. Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008;456 (7218):53-59.

38. Wang J, Wang W, Li R, et al. The diploid genome sequence of an Asian individual. *Nature*. 2008;456(7218):60-65.

39. Ahn SM, Kim TH, Lee S, et al. The first Korean genome sequence and analysis: full genome sequencing for a socio-ethnic group. *Genome Res*. 2009;19(9):1622-1629.

40. Wheeler DA, Srinivasan M, Egholm M, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature*. 2008;452(7189):872-876.

41. Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res*. 2008;18(11):1851-1858.

42. Li R, Li Y, Kristiansen K, Wang J. SOAP: short oligonucleotide alignment program. *Bioinformatics*. 2008;24(5):713-714.

43. Wang J, Mu Q. Soap-HT-BLAST: high throughput BLAST based on Web services. *Bioinformatics*. 2003;19(14):1863-1864.

44. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26(5):589-595.

45. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.

46. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*. 2011;72(2):245-256.

47. Eid J, Fehr A, Gray J, et al. Real-time DNA sequencing from single polymerase molecules. *Science*. 2009;323(5910):133-138.

48. Flusberg BA, Webster DR, Lee JH, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods*. 2010;7(6):461-465.

49. Olasagasti F, Lieberman KR, Benner S, et al. Replication of individual DNA molecules under electronic control using a protein nanopore. *Nat Nanotechnol*. 2010; 5(11):798-806.

50. Lathrop GM, Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. *Proc Natl Acad Sci U S A*. 1984;81(11):3443-3446.

51. Gudbjartsson DF, Thorvaldsson T, Kong A, Gunnarsson G, Ingolfsdottir A. Allegro version 2. *Nat Genet*. 2005;37(10):1015-1016.

52. Fukuda Y, Nakahara Y, Date H, et al. SNP HiTLink: a high-throughput linkage analysis system employing dense SNP data. *BMC Bioinformatics*. 2009;10:121.

53. Wu CH, Fallini C, Ticozzi N, et al. Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature*. 2012;488(7412):499-503.

54. Doi H, Yoshida K, Yasuda T, et al. Exome sequencing reveals a homozygous SYT14 mutation in adult-onset, autosomal-recessive spinocerebellar ataxia with psychomotor retardation. *Am J Hum Genet*. 2011;89(2):320-327.

55. Ku CS, Naidoo N, Pawitan Y. Revisiting Mendelian disorders through exome sequencing. *Hum Genet*. 2011;129(4):351-370.

56. Suthers G, Laing N, Wilton S, Dorosz S, Waddy H. "Sporadic" motoneuron disease due to familial SOD1 mutation with low penetrance. *Lancet*. 1994;344 (8939-8940):1773.

57. Murakami T, Warita H, Hayashi T, et al. A novel SOD1 gene mutation in familial ALS with low penetrance in females. *J Neurol Sci*. 2001;189(1-2):45-47.

58. Rezania K, Yan J, Dellefave L, et al. A rare Cu/Zn superoxide dismutase mutation causing familial amyotrophic lateral sclerosis with variable age of onset, incomplete penetrance and a sensory neuropathy. *Amyotroph Lateral Scler Other Motor Neuron Disord*. 2003;4(3):162-166.

59. Andersen PM, Restagno G, Stewart HG, Chiò A. Disease penetrance in amyotrophic lateral sclerosis associated with mutations in the SOD1 gene. *Ann Neurol*. 2004;55(2):298-299, author reply 299.

60. Nogales-Gadea G, Garcia-Arumi E, Andreu AL, Cervera C, Gamez J. A novel exon 5 mutation (N139H) in the SOD1 gene in a Spanish family associated with incomplete penetrance. *J Neurol Sci*. 2004;219(1-2):1-6.

61. Majounie E, Renton AE, Mok K, et al; Chromosome 9-ALS/FTD Consortium; French research network on FTLD/FTLD/ALS; ITALSGEN Consortium. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol*. 2012; 11(4):323-330.

62. Nozaki I, Hamaguchi T, Sanjo N, et al. Prospective 10-year surveillance of human prion diseases in Japan. *Brain*. 2010;133(10):3043-3057.

63. Rosewich H, Thiele H, Ohlenbusch A, et al. Heterozygous de-novo mutations in ATP1A3 in patients with alternating hemiplegia of childhood: a whole-exome sequencing gene-identification study. *Lancet Neurol*. 2012;11(9):764-773.

64. Veeramah KR, O'Brien JE, Meisler MH, et al. De novo pathogenic SCN8A mutation identified by whole-genome sequencing of a family quartet affected by infantile epileptic encephalopathy and SUDEP. *Am J Hum Genet*. 2012;90(3):502-510.

65. Awadalla P, Gauthier J, Myers RA, et al. Direct measure of the de novo mutation rate in autism and schizophrenia cohorts. *Am J Hum Genet*. 2010;87(3):316-324.

66. Rivière JB, van Bon BW, Hoischen A, et al. De novo mutations in the actin genes ACTB and ACTG1 cause Baraitser-Winter syndrome. *Nat Genet*. 2012;44(4): 440-444, S1-S2.

67. Sasaki H, Emi M, Iijima H, et al. Copy number loss of (src homology 2 domain containing)-transforming protein 2 (SHC2) gene: discordant loss in monozygotic twins and frequent loss in patients with multiple system atrophy. *Mol Brain*. 2011;4:24.

68. Baranzini SE, Mudge J, van Velkinburgh JC, et al. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. *Nature*. 2010; 464(7293):1351-1356.

69. Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet*. 2001;17(9):502-510.

70. Collins FS, Guyer MS, Charkravarti A. Variations on a theme: cataloging human DNA sequence variation. *Science*. 1997;278(5343):1580-1581.

71. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet*. 2001;69(1):124-137.

72. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A*. 2009;106(23):9362-9367.

73. Visscher PM. Sizing up human height variation. *Nat Genet*. 2008;40(5):489-490.

74. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461(7265):747-753.

75. Satake W, Nakabayashi Y, Mizuta I, et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat Genet*. 2009;41(12):1303-1307.

76. Sidransky E, Nalls MA, Aasly JO, et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N Engl J Med*. 2009;361(17):1651-1661.

77. Mitsui J, Mizuta I, Toyoda A, et al. Mutations for Gaucher disease confer high susceptibility to Parkinson disease. *Arch Neurol*. 2009;66(5):571-576.

78. Shino MY, McGuire V, Van Den Eeden SK, et al. Familial aggregation of Parkinson's disease in a multiethnic community-based case-control study. *Mov Disord*. 2010;25(15):2587-2594.

79. Goker-Alpan O, Schiffmann R, LaMarca ME, Nussbaum RL, McInerney-Leo A, Sidransky E. Parkinsonism among Gaucher disease carriers. *J Med Genet*. 2004; 41(12):937-940.

80. Neudorfer O, Giladi N, Elstein D, et al. Occurrence of Parkinson's syndrome in type I Gaucher disease. *QJM*. 1996;89(9):691-694.

81. Kryukov GV, Pennacchio LA, Sunyaev SR. Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. *Am J Hum Genet*. 2007;80(4):727-739.

82. Knudson AG Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A*. 1971;68(4):820-823.

83. Bihaqi SW, Schumacher A, Maloney B, Lahiri DK, Zawia NH. Do epigenetic pathways initiate late onset Alzheimer disease (LOAD): towards a new paradigm. *Curr Alzheimer Res*. 2012;9(5):574-588.

84. Matsumoto L, Takuma H, Tamaoka A, et al. CpG demethylation enhances alpha-synuclein expression and affects the pathogenesis of Parkinson's disease. *PLoS One*. 2010;5(11):e15522.

85. Stöhr J, Watts JC, Mensinger ZL, et al. Purified and synthetic Alzheimer's amyloid beta (Aβ) prions. *Proc Natl Acad Sci U S A*. 2012;109(27):11025-11030.

86. de Calignon A, Polydoro M, Suárez-Calvet M, et al. Propagation of tau pathology in a model of early Alzheimer's disease. *Neuron*. 2012;73(4):685-697.

87. Freeman JD, Warren RL, Webb JR, Nelson BH, Holt RA. Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. *Genome Res*. 2009; 19(10):1817-1824.

88. Warren RL, Freeman JD, Zeng T, et al. Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res*. 2011;21(5):790-797.

89. Logan AC, Gao H, Wang C, et al. High-throughput VDJ sequencing for quantification of minimal residual disease in chronic lymphocytic leukemia and immune reconstitution assessment. *Proc Natl Acad Sci U S A*. 2011;108(52): 21194-21199.

90. Holcomb CL, Höglund B, Anderson MW, et al. A multi-site study using high-resolution HLA genotyping by next generation sequencing. *Tissue Antigens*. 2011; 77(3):206-217.

91. Erlich RL, Jia X, Anderson S, et al. Next-generation sequencing for HLA typing of class I loci. *BMC Genomics*. 2011;12:42.

CSF1R Mutations Identified in Three Families With Autosomal Dominantly Inherited Leukoencephalopathy

Jun Mitsui,[1] Takashi Matsukawa,[1] Hiroyuki Ishiura,[1] Koichiro Higasa,[2] Jun Yoshimura,[2] Taro L. Saito,[2] Budrul Ahsan,[1] Yuji Takahashi,[1] Jun Goto,[1] Atsushi Iwata,[1,3] Yuki Niimi,[4] Yuuichi Riku,[4] Yoji Goto,[4] Kazuo Mano,[4] Mari Yoshida,[5] Shinichi Morishita,[2] and Shoji Tsuji[1,6,7]*

[1]Department of Neurology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[2]Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan

[3]Department of Molecular Neuroscience on Neurodegeneration, The University of Tokyo, Tokyo, Japan

[4]Department of Neurology, Nagoya Daiichi Red Cross Hospital, Aichi, Japan

[5]Department of Neuropathology, Institute for Medical Science of Aging, Aichi Medical University, Aichi, Japan

[6]Medical Genome Center, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[7]Division of Applied Genetics, National Institute of Genetics, Shizuoka, Japan

Genetic and phenotypic heterogeneities are considerably high in adult-onset leukoencephalopathy, in which comprehensive mutational analyses of the candidate genes by conventional methods are too laborious. We applied exome sequencing to conduct a comprehensive mutational analysis of genes for autosomal dominant leukoencephalopathies. Genomic DNA samples from four patients of three families with autosomal dominantly inherited adult-onset leukodystrophy were subjected to exome sequencing. On the basis of the results, 21 patients with adult-onset sporadic leukodystrophy and one patient with pathologically proven HDLS were additionally screened for *CSF1R* mutations. Exome sequencing identified heterozygous *CSF1R* mutations (p.I794T and p.R777W) in two families. I794T has recently been reported as a causative mutation for hereditary diffuse leukoencephalopathy with spheroids (HDLS), and R777W is a novel mutation. Although mutational analysis of *CSF1R* in 21 sporadic cases revealed no mutations, another novel *CSF1R* mutation, p.C653Y, was identified in one patient with autopsy-proven HDSL. These variants were located in the PTK domain where the causative mutations cluster. Functional prediction of the mutant *CSF1R* as well as cross-species conservation of the affected amino acids supports the notion that these variants are pathogenic for HDLS. Exome sequencing is useful for a comprehensive mutational analysis of causative genes for hereditary leukoencephalopathies, and *CSF1R* should be considered a candidate gene for patients with autosomal dominant leukoencephalopathies. © 2012 Wiley Periodicals, Inc.

Key words: hereditary diffuse leukoencephalopathy with spheroids; *CSF1R*; exome sequencing; molecular diagnosis

INTRODUCTION

Hereditary diffuse leukoencephalopathy with spheroids (HDLS) is a rare autosomal dominant disease characterized neuropatholog-

ically by myelin loss and the presence of axonal spheroids. The clinical presentations of HDLS are characterized by a mean age at onset of 39 ± 15 years (range: 8–78 years) and insidiously progressive behavioral, cognitive, and/or motor dysfunctions (parkinsonism, ataxia, pyramidal dysfunctions, and epilepsy) [Axelsson et al., 1984; Wider et al., 2009]. White matter abnormalities are predominantly found in the frontal and parietal lobes accompanied by evolving cortical atrophy on magnetic resonance imaging (MRI). Because neither the clinical symptoms nor the MRI findings are specific, until recently, diagnosis of HDLS has depended solely on histopathological examination [Axelsson et al., 1984] or brain biopsy [Mateen et al., 2010]. The causative gene for HDLS has recently been identified to be the colony stimulating factor 1 receptor gene (CSF1R) [Rademakers et al., 2011].

Leukoencephalopathies are clinically and genetically heterogeneous diseases and a number of diseases need to be considered for the differential diagnosis. The following are among the leukoencephalopathies with autosomal dominant inheritance: cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) [Joutel et al., 1996], lamin B1 duplications [Padiath et al., 2006], Alexander disease [Brenner et al., 2001], several types of cerebral amyloid angiopathy and small vessel diseases [Ghiso et al., 1986; Levy et al., 1990; Van Broeckhoven et al., 1990; Sherrington et al., 1995; Vidal et al., 1999; Paloneva et al., 2000; Brenner et al., 2001; Gould et al., 2006; Padiath et al., 2006; Richards et al., 2007], and HDLS. The following are among the leukoencephalopathies with autosomal recessive inheritance: cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy (CARASIL) [Hara et al., 2009], vanishing white matter disease [Leegwater et al., 2001; van der Knaap et al., 2002], Nasu-Hakola disease [Paloneva et al., 2000; Kondo et al., 2002], Krabbe disease [Sakai et al., 1994], and metachromatic leukodystrophy [Polten et al., 1991]. Hence, establishing the diagnosis of these conditions is often difficult in clinical practice, necessitating comprehensive mutational analysis of a substantial number of causative genes for leukoencephalopathies [Köhler, 2010]. Here, we applied exome sequencing employing massively parallel sequencers for a comprehensive mutational analysis of causative genes for leukoencephalopathies, and identified cases with mutations in CSF1R.

METHODS

Exome Sequencing Analysis

Written informed consent was obtained from all the participants. Genomic DNA samples were extracted from peripheral blood leukocytes following standard procedures. We conducted exome sequencing in four individuals of three Japanese families with autosomal dominantly inherited adult-onset leukodystrophy of unknown etiology (II-6 and III-1 of Family 1; III-3 of Family 2; and III-1 of Family 3, Fig. 1A). Exonic sequences were enriched using a TruSeq Exome Enrichment kit (Illumina, San Diego, CA) and subjected to massively parallel sequence analysis employing an Illumina Genome Analyzer IIx platform (Illumina) following the manufacturer's instructions. Burrows Wheeler Aligner [Li and Durbin, 2009] and Samtools [Li et al., 2009] were used with the

default settings for alignment of raw reads and variation detection. All the genomic variants were filtered against dbSNP (build 135) [http://www.ncbi.nlm.nih.gov/snp]. Aligned short reads were viewed using the University of Tokyo Genome Browser (UTGB) [Saito et al., 2009]. Variant confirmations were conducted by direct nucleotide sequence analysis using an ABI 3100 Genetic Analyzer (Life Technologies, Foster City, CA).

Screening for CSF1R Mutations in Patients With Sporadic Leukoencephalopathy and in a Case of Pathologically Diagnosed HDSL

To investigate the frequency of CSF1R mutations in patients with sporadic leukoencephalopathies, we further conducted the direct nucleotide sequence analysis of CSF1R in 21 Japanese patients with the clinical diagnosis of sporadic leukoencephalopathies. Because all the previously reported mutations in patients with HDLS are solely located in the protein tyrosine kinase (PTK) domain of CSF1R, we screened for mutations of CSF1R exons 12–20. In addition, the result of mutational analysis of CSF1R for one pathologically proven case of HDLS (III-1 of Family 4, Fig. 1A) was also included in this study.

In Silico Analyses

To predict the impact of amino acid substitutions on protein activity, we conducted a battery of in silico analyses using Polymorphism phenotyping v2 (Polyphen-2) [Adzhubei et al., 2010]; SIFT [Ng and Henikoff, 2001]; MutationTaster [Schwarz et al., 2010]; and MUPro [Cheng et al., 2006], along with species conservation analysis using UCSC Genome Browser [Karolchik et al., 2007].

RESULTS

Exome Sequencing

We obtained mean coverage depths of 109.4, 100.1, 137.7, and 195.0 in II-6 and III-1 of Family 1, III-3 of Family 2, and III-1 of Family 3, respectively, which are sufficient for examining the exons for mutations. All the nonsynonymous and splice-site single-nucleotide variants (SNVs), or insertions and deletions in coding sequences (coding indels), which were not registered in dbSNP135, were obtained (hereafter collectively called "novel variants"). The results of the exome sequence analysis are summarized in Table I. We then screened for variants involving previously known causative genes of autosomal dominant adult-onset leukoencephalopathies (NOTCH3, LMNB1, GFAP, APP, PSEN1, PSEN2, CST3, ITM2B, TREX1, COL4A1, and CSF1R). We solely found a heterozygous p.I794T mutation in CSF1R that was shared among the affected individuals (III-1 and II-6) of Family 1 (Table II). This mutation is identical to that previously identified as the causative mutation for HDLS [Rademakers et al., 2011]. We, furthermore, identified a novel CSF1R mutation (p.R777W) in III-3 of Family 2 (Fig. 2). There were no other novel variants in the known causative genes either in Family 1 or in Family 2 (Table II). No candidate variants were identified in the known causative genes for autosomal dominant leukoencephalopathies in III-1 of Family 3.
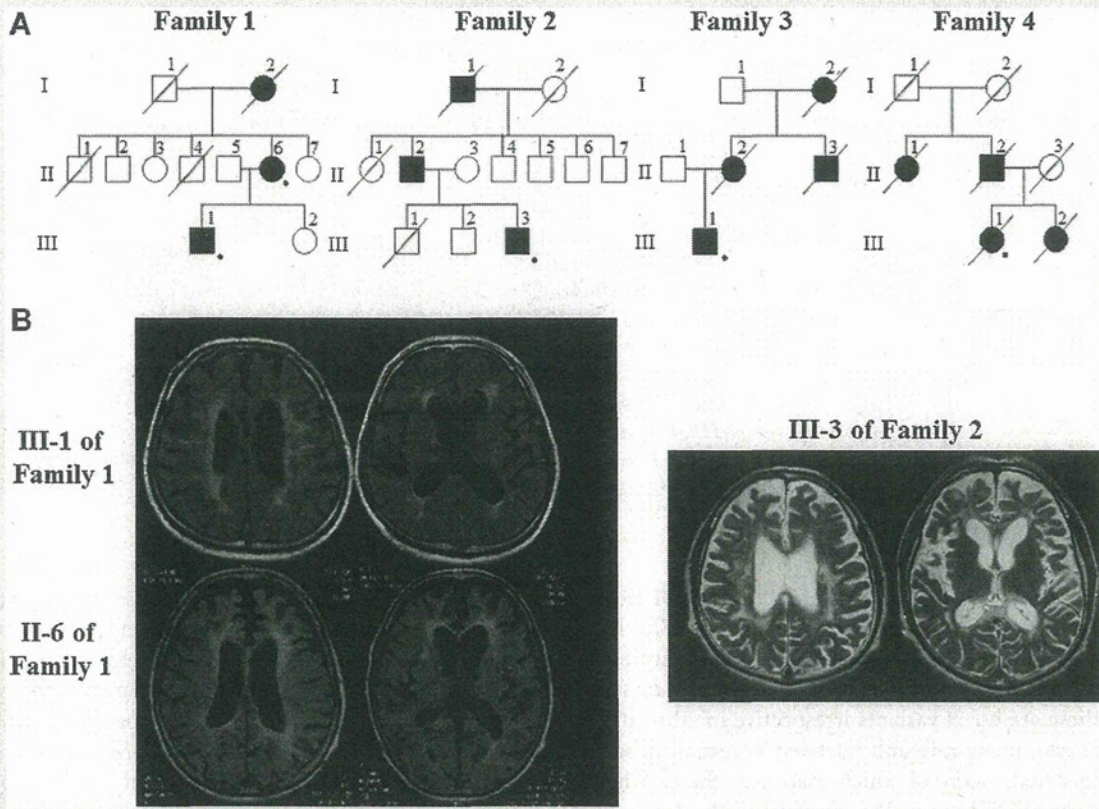
FIG. 1. Pedigree charts and neuroimaging findings. A: Males, squares; female, circles; affected individuals, solid symbols; unaffected individuals, open symbols; available genomic DNAs, dot. B: Axial fluid attenuated inversion recovery (FLAIR) magnetic resonance images of individual III-1 of Family 1 at the age of 54 and those of individual II-6 of Family 1 at the age of 70, showing atrophy in frontal, parietal, and medial temporal lobes, with periventricular confluent hyperintensities. Axial T2-weighted magnetic resonance images of individual III-3 of Family 2 at the age of 41, showing atrophy in frontal, parietal and temporal lobes, with confluent hyperintensities in periventricular regions including corpus callosum.

Screening for *CSF1R* Mutations in Patients With Sporadic Leukoencephalopathy and in a Patient With Pathologically Diagnosed HDSL

Considering the possibilities of reduced penetrance, insufficient information on the family history, or de novo mutations in *CSF1R* in patients with sporadic leukoencephalopathies, we investigated the frequencies of *CSF1R* mutations in 21 patients with adult-onset sporadic leukoencephalopathy. In this sporadic case series,

however, no mutations were identified. We further screened for *CSF1R* mutations in a patient with pathologically diagnosed HDLS (III-1 of Family 4, Fig. 1), which revealed another novel *CSF1R* mutation (p.C653Y).

Implications of Mutant *CSF1R* as Cause of HDSL

None of the identified variants (C653Y, R777W, and I794T) were identified either in dbSNP135, 1,000 Genomes (http://

TABLE I. Summary of Exome Sequence Analysis Results

Family	Individual	Read	Mapped read	Coverage	Proportion of target bases covered > 10-fold read depth (%)	SNVs (total)	Unknown NS/SS SNVs	Indels (total)	Unknown coding indels
1	II-6	67.2 M	61.8 M (91.8%)	109.4	89.1	227,491	354	20,965	32
	III-1	61.4 M	56.5 M (92.0%)	100.1	88.2	218,020	349	20,261	30
2	III-3	77.9 M	64.9 M (83.4%)	137.7	80.8	193,528	396	18,271	36
3	III-1	94.4 M	91.9 M (97.5%)	195.0	80.8	253,999	292	24,951	31

SNVs, single nucleotide variants; NS, nonsynonymous; SS, splice-site.

TABLE II. Nonsynonymous Variants in Known Causative Genes of Autosomal Dominant Adult-Onset Leukoencephalopathies (*NOTCH3*, *LMNB1*, *GFAP*, *APP*, *PSEN1*, *PSEN2*, *CST3*, *ITM2B*, *TREX1*, *COL4A1*, and *CSF1R*)

Family	Individual	Variants registered in dbSBP135	Novel variants (not registered in dbSNP135)
1	II-6	*CSF1R*, rs10079250	*CSF1R*, I794T
		NOTCH3, rs1044009	
		COL4A1, rs536174	
		COL4A1, rs9515185	
	III-1	*CSF1R*, rs10079250	*CSF1R*, I794T
		NOTCH3, rs1044009	
		COL4A1, rs3742207	
		COL4A1, rs536174	
2	III-3	*NOTCH3*, rs1044009	*CSF1R*, R777W
		COL4A1, rs3742207	
		COL4A1, rs536174	
3	III-1	*NOTCH3*, rs1044009	None
		COL4A1, rs536174	

www.1000genomes.org/, accessed at June 2012), NHLBI Exome Sequencing Project (http://evs.gs.washington.edu/EVS/, EVS-v0.0.13), or 720 samples (individuals with other neurological diseases from the Japanese population) of inhouse exome database, showing that these are novel variants irrespective of ethnicity. In the above databases, many rare and relatively common missense variants are registered, many of which map near the two newly identified variants. Nonetheless, the absence of the two newly identified variants in these databases is strongly indicative of pathogenicity. As the family members were unavailable for segregation analyses of the mutation except for Family 1, we then investigated cross-species conservation of the mutated amino acids and functional prediction analyses to obtain supporting evidence for the pathogenicity of the mutant CSF1R identified in this study. As shown in Figure 3B, a strong conservation of the affected amino acids across species was well demonstrated for all three variants. In silico functional prediction analyses unanimously predicted all the identified variants to be "Probably damaging" by Polyphen-2, "Damaging" by SIFT, "Disease-causing" by MutationTaster, and "Stability decreasing" by MUPro (Fig. 3C). Intriguingly, they were all located in the PTK domain similarly to previously reported mutations (Fig. 3A). Taken together, we considered that all the identified variants (one known causative mutation and two novel mutations) were pathogenic for HDLS.

Clinical Features

In total, three families with adult-onset leukoencephalopathy with *CSF1R* mutations were identified, including one family with the pathological diagnosis of HDSL. Detailed clinical information was available for six individuals from the three families. A summary of the clinical characteristics is shown in Table III. The mean age at onset of the six patients was 51.0 ± 10.0 years (range: 38–65 years), and the mean age at death of the three deceased patients was 55.7 ± 10.2 years (range: 44–63 years). Initial symptoms substantially varied within and across the families. Notably, the proband (III-3) of Family 2 developed alcoholism 4 years before his admis-

sion to our hospital, and substance abuse such as alcoholism has often been reported in HDLS [Axelsson et al., 1984; van der Knaap et al., 2000]. In the course of disease development, personality and behavioral changes, and dementia were highly prevalent. Parkinsonism was observed in one of the three patients, and seizures were present in two of the five patients. Brain MR images of individuals III-1 and II-6 of Family 1, and III-3 of Family 2 are shown in Figure 1B, which showed atrophy in the frontal, parietal, and medial temporal lobes, with periventricular confluent hyperintensities on fluid attenuated inversion recovery (FLAIR) or T2-weighted images. Detailed information on the medical history of the patients is provided in Supplementary Text.

DISCUSSION

Various mutations in *CSF1R* have been described in 15 families with HDLS [Rademakers et al., 2011; Kinoshita et al., 2012], and all the causative mutations are located in the PTK domain of CSF1R (Fig. 3A). We herein identified *CSF1R* mutations in three families with autosomal dominantly inherited leukoencephalopathy (two with leukoencephalopathy of unknown etiology and one with autopsy-proven HDLS). The I794T mutation is the same as that identified in a previously reported family in the United States (Family SC) [Van Gerpen et al., 2008], and the novel mutations identified (C653Y and R777W) are also located at the PTK domain. Taken together with strong conservation of the affected amino acids across species (Fig. 3B), and in silico prediction of functional impairment associated with the mutations (Fig. 3C), all the identified mutations in this study are considered to be pathogenic.

HDLS is indeed a rare hereditary disease, and it is likely unrecognized owing to the nonspecific MRI findings, which are common to ischemic changes or other causes of white matter diseases, and also owing to the quite variable clinical presentations. The affected individuals (III-1 and II-6) of Family 1 started to show cognitive impairments in their fifth or sixth decades, which relatively rapidly worsened year by year, and their clinical diagnosis was atypical Alzheimer disease. The examined member (III-3) of Family 2
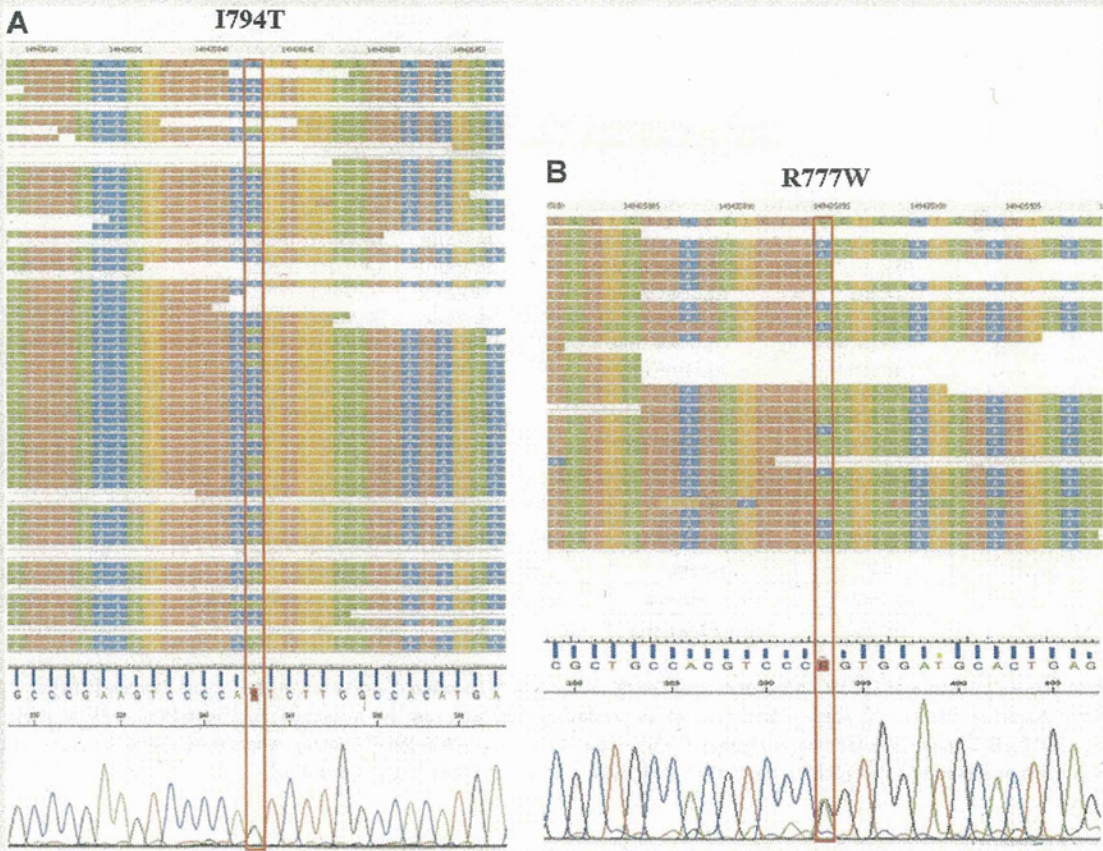
FIG. 2. Identification of causative mutation in *CSF1R*. Aligned short reads and the corresponding electropherograms of direct nucleotide sequence analysis are shown (A: I794T mutation; B: R777W mutation).

gradually developed alcoholism in his late thirties, and around the age of 40, his cognitive impairment and gait disturbance worsened rapidly. His clinical diagnosis was cerebral small vascular disease with unknown etiology. It should be emphasized that the affected individuals in Families 1 and 2 did not have a diagnosis of HDLS at the outset, and that only through exome sequencing was *CSF1R* identified. Thus *CSF1R* should be considered a candidate gene for autosomal dominant leukoencephalopathies regardless of whether biopsy has been obtained to look for spheroids. To data, the diagnosis of HDLS has been made solely by neuropathological

TABLE III. Clinical Characteristics of Six Individuals From Three Japanese Families With *CSF1R* Mutations

| Family	Individual	*CSF1R* mutation	Sex	Onset age	Death age	Initial symptom	Clinical features during course of disease development				
							Personality and behavioral changes	Dementia	Depression	Parkinsonism	Seizures
1	II-6	I794T	Female	60	Alive	Forgetfulness	+	+	−	−	−
	III-1	I794T	Male	52	60	Apathy	+	+	+	+	−
2	II-2	Not tested	Male	65	Alive	Depression	+	+	Not described	Not described	+
	III-3	R777W	Male	38	Alive	Alcoholism	+	+	+	−	−
4	III-1	C653Y	Female	48	63	Repetitive behavior	+	+	Not described	Not described	+
	III-2	Not tested	Female	43	44	Body weight loss	Not described	Not described	Not described	Not described	Not described
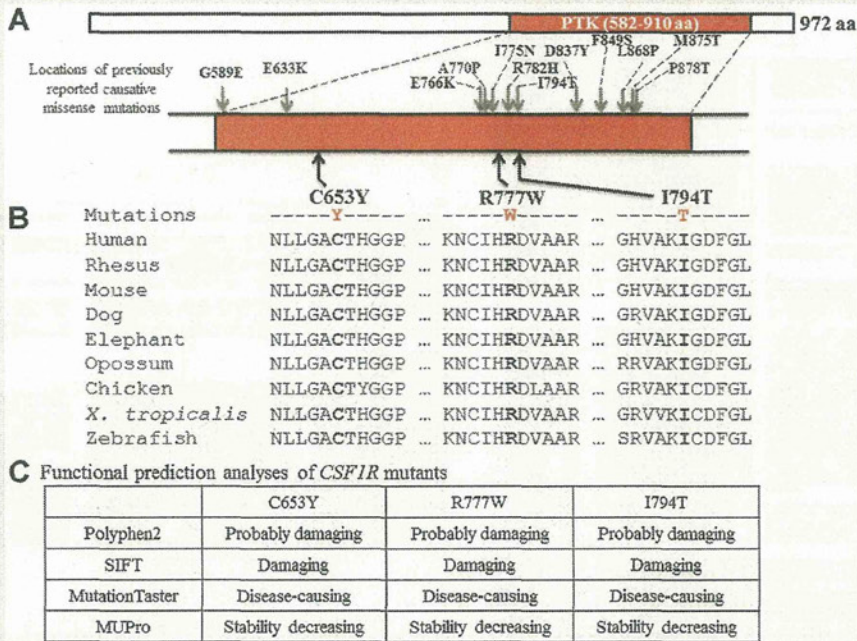
FIG. 3. Protein domain structure of CSF1R with summary of *CSF1R* mutations. A: Location of protein tyrosine kinase domain (PTK) of CSF1R, with summary of mutations identified in the present study as well as previously reported causative missense mutations [Rademakers et al., 2011; Kinoshita et al., 2012]. B: Comparative genomic analysis of multiple species for the parts of the PTK domain where the mutations occur is shown. C: Summary of in silico analyses of pathogenicity prediction (Polyphen-2, SIFT, MutationTaster, and MUPro).

findings. With the availability of mutational analysis of *CSF1R*, the clinical spectrum of patients with *CSF1R* mutations and genotype–phenotype correlations should be thoroughly investigated.

There are numerous genes related to leukoencephalopathies, for which it is difficult to focus on particular genes for the mutational analysis depending solely on phenotypes. Targeted sequencing would be as effective as exome sequencing and certainly a lot less expensive currently. On the other hand, a growing number of causative genes have been identified in Mendelian diseases, and comprehensive mutational analysis of causative genes may necessitate updating of a method for target sequencing on a continuous basis, which is often difficult in clinical practice. Considering this situation, exome sequencing has become a common method of molecular diagnosis of Mendelian diseases. On the other hand, we are encountering an increasing number of very rare variants that are not necessarily pathogenic. Because exome sequencing provides virtually all the variants of genes that are relevant to a particular disease group, that is, leukoencephalopathies in this study, knowledge of allele frequencies of variants in a specific phenotyped population is indeed quite helpful for interpreting which of those variants are likely to be pathogenic.

REFERENCES

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. Nat Methods 7(4):248–249.

Axelsson R, Röyttä M, Sourander P, Akesson HO, Andersen O. 1984. Hereditary diffuse leucoencephalopathy with spheroids. Acta Psychiatr Scand Suppl 314:1–65.

Brenner M, Johnson AB, Boespflug-Tanguy O, Rodriguez D, Goldman JE, Messing A. 2001. Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease. Nat Genet 27(1):117–120.

Cheng J, Randall A, Baldi P. 2006. Prediction of protein stability changes for single-site mutations using support vector machines. Proteins 62(4): 1125–1132.

Ghiso J, Jensson O, Frangione B. 1986. Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variant of gamma-trace basic protein (cystatin C). Proc Natl Acad Sci USA 83(9):2974–2978.

Gould DB, Phalan FC, van Mil SE, Sundberg JP, Vahedi K, Massin P, Bousser MG, Heutink P, Miner JH, Tournier-Lasserve E, John SW. 2006. Role of COL4A1 in small-vessel disease and hemorrhagic stroke. N Engl J Med 354(14):1489–1496.

Hara K, Shiga A, Fukutake T, Nozaki H, Miyashita A, Yokoseki A, Kawata H, Koyama A, Arima K, Takahashi T, Ikeda M, Shiota H, Tamura M, Shimoe Y, Hirayama M, Arisato T, Yanagawa S, Tanaka A, Nakano I, Ikeda S, Yoshida Y, Yamamoto T, Ikeuchi T, Kuwano R, Nishizawa M,

Tsuji S, Onodera O. 2009. Association of HTRA1 mutations and familial ischemic cerebral small-vessel disease. N Engl J Med 360(17):1729–1739.

Joutel A, Corpechot C, Ducros A, Vahedi K, Chabriat H, Mouton P, Alamowitch S, Domenga V, Cécillion M, Marechal E, Maciazek J, Vayssiere C, Cruaud C, Cabanis EA, Ruchoux MM, Weissenbach J, Bach JF, Bousser MG, Tournier-Lasserve E. 1996. Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. Nature 383(6602):707–710.

Karolchik D, Bejerano G, Hinrichs AS, Kuhn RM, Miller W, Rosenbloom KR, Zweig AS, Haussler D, Kent WJ. 2007. Comparative genomic analysis using the UCSC genome browser. Methods Mol Biol 395:17–34.

Kinoshita M, Yoshida K, Oyanagi K, Hashimoto T, Ikeda S. 2012. Hereditary diffuse leukoencephalopathy with axonal spheroids caused by R782H mutation in CSF1R: Case report. J Neurol Sci 318(1–2):115–118.

Köhler W. 2010. Leukodystrophies with late disease onset: An update. Curr Opin Neurol 23(3):234–241.

Kondo T, Takahashi K, Kohara N, Takahashi Y, Hayashi S, Takahashi H, Matsuo H, Yamazaki M, Inoue K, Miyamoto K, Yamamura T. 2002. Heterogeneity of presenile dementia with bone cysts (Nasu–Hakola disease): Three genetic forms. Neurology 59(7):1105–1107.

Leegwater PA, Vermeulen G, Könst AA, Naidu S, Mulders J, Visser A, Kersbergen P, Mobach D, Fonds D, van Berkel CG, Lemmers RJ, Frants RR, Oudejans CB, Schutgens RB, Pronk JC, van der Knaap MS. 2001. Subunits of the translation initiation factor eIF2B are mutant in leukoencephalopathy with vanishing white matter. Nat Genet 29(4):383–388.

Levy E, Carman MD, Fernandez-Madrid IJ, Power MD, Lieberburg I, van Duinen SG, Bots GT, Luyendijk W, Frangione B. 1990. Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. Science 248(4959):1124–1126.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25(14):1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAM tools. Bioinformatics 25(16):2078–2079.

Mateen FJ, Keegan BM, Krecke K, Parisi JE, Trenerry MR, Pittock SJ. 2010. Sporadic leucodystrophy with neuroaxonal spheroids: Persistence of DWI changes and neurocognitive profiles: A case study. J Neurol Neurosurg Psychiatry 81(6):619–622.

Ng PC, Henikoff S. 2001. Predicting deleterious amino acid substitutions. Genome Res 11(5):863–874.

Padiath QS, Saigoh K, Schiffmann R, Asahara H, Yamada T, Koeppen A, Hogan K, Ptácek LJ, Fu YH. 2006. Lamin B1 duplications cause autosomal dominant leukodystrophy. Nat Genet 38(10):1114–1123.

Paloneva J, Kestilä M, Wu J, Salminen A, Böhling T, Ruotsalainen V, Hakola P, Bakker AB, Phillips JH, Pekkarinen P, Lanier LL, Timonen T, Peltonen L. 2000. Loss-of-function mutations in TYROBP (DAP12) result in a presenile dementia with bone cysts. Nat Genet 25(3):357–361.

Polten A, Fluharty AL, Fluharty CB, Kappler J, von Figura K, Gieselmann V. 1991. Molecular basis of different forms of metachromatic leukodystrophy. N Engl J Med 324(1):18–22.

Rademakers R, Baker M, Nicholson AM, Rutherford NJ, Finch N, Soto-Ortolaza A, Lash J, Wider C, Wojtas A, DeJesus-Hernandez M, Adamson

J, Kouri N, Sundal C, Shuster EA, Aasly J, MacKenzie J, Roeber S, Kretzschmar HA, Boeve BF, Knopman DS, Petersen RC, Cairns NJ, Ghetti B, Spina S, Garbern J, Tselis AC, Uitti R, Das P, Van Gerpen JA, Meschia JF, Levy S, Broderick DF, Graff-Radford N, Ross OA, Miller BB, Swerdlow RH, Dickson DW, Wszolek ZK. 2011. Mutations in the colony stimulating factor 1 receptor (CSF1R) gene cause hereditary diffuse leukoencephalopathy with spheroids. Nat Genet 44(2):200–205.

Richards A, van den Maagdenberg AM, Jen JC, Kavanagh D, Bertram P, Spitzer D, Liszewski MK, Barilla-Labarca ML, Terwindt GM, Kasai Y, McLellan M, Grand MG, Vanmolkot KR, de Vries B, Wan J, Kane MJ, Mamsa H, Schäfer R, Stam AH, Haan J, de Jong PT, Storimans CW, van Schooneveld MJ, Oosterhuis JA, Gschwendter A, Dichgans M, Kotschet KE, Hodgkinson S, Hardy TA, Delatycki MB, Hajj-Ali RA, Kothari PH, Nelson SF, Frants RR, Baloh RW, Ferrari MD, Atkinson JP. 2007. C-terminal truncations in human 3′–5′ DNA exonuclease TREX1 cause autosomal dominant retinal vasculopathy with cerebral leukodystrophy. Nat Genet 39(9):1068–1070.

Saito TL, Yoshimura J, Sasaki S, Ahsan B, Sasaki A, Kuroshu R, Morishita S. 2009. UTGB toolkit for personalized genome browsers. Bioinformatics 25(15):1856–1861.

Sakai N, Inui K, Fujii N, Fukushima H, Nishimoto J, Yanagihara I, Isegawa Y, Iwamatsu A, Okada S. 1994. Krabbe disease: Isolation and characterization of a full-length cDNA for human galactocerebrosidase. Biochem Biophys Res Commun 198(2):485–491.

Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. 2010. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 7(8):575–576.

Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, Chi H, Lin C, Li G, Holman K, Tsuda T, Mar L, Foncin JF, Bruni AC, Montesi MP, Sorbi S, Rainero I, Pinessi L, Nee L, Chumakov I, Pollen D, Brookes A, Sanseau P, Polinsky RJ, Wasco W, Da Silva HA, Haines JL, Perkicak-Vance MA, Tanzi RE, Roses AD, Fraser PE, Rommens JM, St George-Hyslop PH. 1995. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature 375(6534):754–760.

Van Broeckhoven C, Haan J, Bakker E, Hardy JA, Van Hul W, Wehnert A, Vegter-Van der Vlis M, Roos RA. 1990. Amyloid beta protein precursor gene and hereditary cerebral hemorrhage with amyloidosis (Dutch). Science 248(4959):1120–1122.

van der Knaap MS, Naidu S, Kleinschmidt-Demasters BK, Kamphorst W, Weinstein HC. 2000. Autosomal dominant diffuse leukoencephalopathy with neuroaxonal spheroids. Neurology 54(2):463–468.

van der Knaap MS, Leegwater PA, Könst AA, Visser A, Naidu S, Oudejans CB, Schutgens RB, Pronk JC. 2002. Mutations in each of the five subunits of translation initiation factor eIF2B can cause leukoencephalopathy with vanishing white matter. Ann Neurol 51(2):264–270.

Van Gerpen JA, Wider C, Broderick DF, Dickson DW, Brown LA, Wszolek ZK. 2008. Insights into the dynamics of hereditary diffuse leukoencephalopathy with axonal spheroids. Neurology 71(12):925–929.

Vidal R, Frangione B, Rostagno A, Mead S, Révész T, Plant G, Ghiso J. 1999. A stop-codon mutation in the BRI gene associated with familial British dementia. Nature 399(6738):776–781.

Wider C, Van Gerpen JA, DeArmond S, Shuster EA, Dickson DW, Wszolek ZK. 2009. Leukoencephalopathy with spheroids (HDLS) and pigmentary leukodystrophy (POLD): A single entity? Neurology 72(22):1953–1959.

REPORT

The TRK-Fused Gene Is Mutated in Hereditary Motor and Sensory Neuropathy with Proximal Dominant Involvement

Hiroyuki Ishiura,[1] Wataru Sako,[3] Mari Yoshida,[4] Toshitaka Kawarai,[3] Osamu Tanabe,[3,5] Jun Goto,[1] Yuji Takahashi,[1] Hidetoshi Date,[1] Jun Mitsui,[1] Budrul Ahsan,[1] Yaeko Ichikawa,[1] Atsushi Iwata,[1] Hiide Yoshino,[6] Yuishin Izumi,[3] Koji Fujita,[3] Kouji Maeda,[3] Satoshi Goto,[3] Hidetaka Koizumi,[3] Ryoma Morigaki,[3] Masako Ikemura,[7] Naoko Yamauchi,[7] Shigeo Murayama,[8] Garth A. Nicholson,[9] Hidefumi Ito,[10] Gen Sobue,[11] Masanori Nakagawa,[12] Ryuji Kaji,[3,*] and Shoji Tsuji[1,2,13,*]

Hereditary motor and sensory neuropathy with proximal dominant involvement (HMSN-P) is an autosomal-dominant neurodegenerative disorder characterized by widespread fasciculations, proximal-predominant muscle weakness, and atrophy followed by distal sensory involvement. To date, large families affected by HMSN-P have been reported from two different regions in Japan. Linkage and haplotype analyses of two previously reported families and two new families with the use of high-density SNP arrays further defined the minimum candidate region of 3.3 Mb in chromosomal region 3q12. Exome sequencing showed an identical c.854C>T (p.Pro285-Leu) mutation in the TRK-fused gene (*TFG*) in the four families. Detailed haplotype analysis suggested two independent origins of the mutation. Pathological studies of an autopsied patient revealed TFG- and ubiquitin-immunopositive cytoplasmic inclusions in the spinal and cortical motor neurons. Fragmentation of the Golgi apparatus, a frequent finding in amyotrophic lateral sclerosis, was also observed in the motor neurons with inclusion bodies. Moreover, TAR DNA-binding protein 43 kDa (TDP-43)-positive cytoplasmic inclusions were also demonstrated. In cultured cells expressing mutant TFG, cytoplasmic aggregation of TDP-43 was demonstrated. These findings indicate that formation of TFG-containing cytoplasmic inclusions and concomitant mislocalization of TDP-43 underlie motor neuron degeneration in HMSN-P. Pathological overlap of proteinopathies involving TFG and TDP-43 highlights a new pathway leading to motor neuron degeneration.

Hereditary motor and sensory neuropathy with proximal dominant involvement (HMSN-P [MIM 604484]) is an autosomal-dominant disease characterized by predominantly proximal muscle weakness and atrophy followed by distal sensory disturbances.[1] HMSN-P was first described in patients from the Okinawa Islands of Japan, where more than 100 people are estimated to be affected.[2] Two Brazilian HMSN-P-affected families of Okinawan ancestry have also been reported.[3,4]

The disease onset is usually in the 40s and is followed by a slowly progressive course. Painful muscle cramps and abundant fasciculations are observed, particularly in the early stage of the disease. In contrast to the clinical presentations of other hereditary motor and sensory neuropathies (HMSNs) presenting with predominantly distal motor weakness reflecting axonal-length dependence, the clinical presentation of HMSN-P is unique in that it involves proximal predominant weakness with widespread fasciculations resembling those of amyotrophic lateral sclerosis (ALS).[5] Distal sensory loss is accompanied later in the disease course, but the degree of the sensory involvement varies among patients. Neuropathological findings revealed severe neuronal loss and gliosis in the spinal anterior horns and mild neuronal loss and gliosis in the hypoglossal and facial nuclei of the brainstem, which indicates that the primary pathological feature of HMSN-P is a motor neuronopathy involving motor neurons, but not a motor neuropathy involving axons.[1,5] The posterior column, corticospinal tract, and spinocerebellar tract showed loss of myelinated fibers and gliosis. Neuronal loss and gliosis were found in Clarke's nucleus. Dorsal root ganglia showed mild to marked neuronal loss.[1,5] These observations suggest that HMSN-P shares neuropathological findings in part with those observed in familial ALS.[6]

Previous studies on Okinawan kindreds mapped the disease locus to chromosome 3q.[1] Subsequently, we identified two large families (families 1 and 2 in Figure 1A) affected by quite a similar phenotype in the Kansai area of Japan, located in the middle of the main island of Japan and far distant from the Okinawa Islands. We mapped the
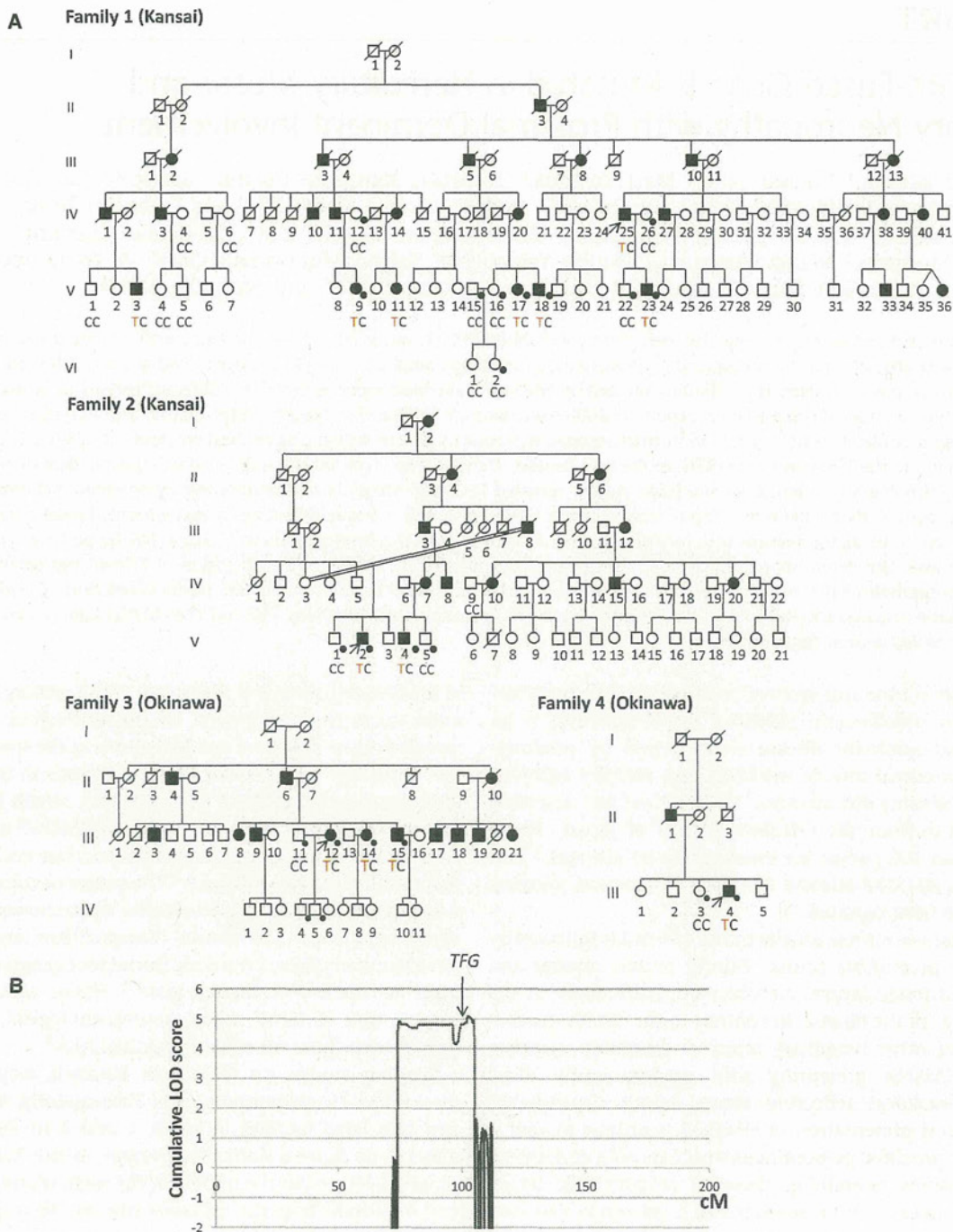
Figure 1. Pedigree Charts and Linkage Analysis
(A) Pedigree charts of families 1 and 2 (Kansai kindreds) and families 3 and 4 (Okinawan kindreds) are shown. Squares and circles indicate males and females, respectively. Affected persons are designated with filled symbols. A diagonal line through a symbol represents a deceased person. A person with an arrow is an index patient. Genotypes of *TFG* c.854 are shown in individuals in whom genomic DNA was analyzed. Individuals genotyped with SNP arrays for linkage analysis and haplotype reconstruction are indicated by dots.
(B) Cumulative parametric multipoint LOD scores on chromosome 3 of all the families are shown.

disease locus to chromosome 3q,[7] overlapping with the previously defined locus, which strongly indicates that these diseases are indeed identical.

In addition to the large Kansai HMSN-P-affected families, we found two new Okinawan HMSN-P-affected lies, we found two new Okinawan HMSN-P-affected

families (families 3 and 4 in Figure 1A) in our study. In total, 9 affected and 15 unaffected individuals from the Kansai area and four affected and four unaffected individuals from the Okinawa Islands were enrolled in the study. Written informed consent was obtained from

Table 1. Clinical Characteristics of Patients with HMSN-P from Families 1 and 2 from Kansai and Families 3 and 4 from Okinawa

| | Families 1 and 2 | Family 3 | | | Family 4 |
		III-12	III-14	III-15	III-4
Age at examination (years)	40s–50s	54	52	50	54
Age at onset (years)	37.5 ± 8	44	40	early 20s	41
Initial symptoms	shoulder dislocation and difficulty walking	proximal leg weakness	painful cramps	painful cramps and fasciculation	painful cramps and calf atrophy
Motor					
Proximal muscle weakness and atrophy	+	+	mild	+	+
Painful cramps	+	+	+	+	+
Fasciculations	+	+	+	+	+
Motor ability	bedridden after 10–20 years from disease onset	unable to walk; wheelchair	only mild difficulty climbing stairs	walk with effort	unable to walk; wheelchair
Bulbar symptoms	– ~ +	–	–	–	–
Sensory					
Dysesthesia	+	+	mild	+	+
Decreased tactile sensation	+	+	–	mild	+
Decreased vibratory sensation	+	mild	mild	mild	+
Reflexes					
Tendon reflexes	diminished	diminished	diminished	diminished	diminished
Pathological reflexes	–	–	–	–	–
Laboratory Tests and Electrophysiological Findings					
Serum creatine kinase level	270 ± 101 IU/l	761 IU/l	not measured	625 IU/l	399 IU/l
Hyperglycemia	4/13 patients	–	–	–	+
Hyperlipidemia	3/13 patients	+	–	+	+
Nerve conduction study	motor and sensory axonal degeneration	motor and sensory axonal degeneration	not examined	not examined	motor and sensory axonal degeneration
Needle electromyography	neurogenic changes with fibrillation potentials and positive sharp waves	neurogenic changes with fibrillation potentials and positive sharp waves	not examined	not examined	not examined

The clinical characteristics of the patients from families 1 and 2 were summarized in accordance with the previous studies.[5,6]

all participants. This study was approved by the institutional review boards at the University of Tokyo and the Tokushima University Hospital. Genomic DNA was extracted from peripheral-blood leukocytes or an autopsied brain according to standard procedures.

The clinical presentations of the patients from the four families are summarized in Table 1 and Table S1, available online. Characteristic painful cramps and fasciculations were noted at the initial stage of the disease in all the patients from the four families. Whereas some of the patients showed painful cramps in their 20s, the ages of onset of motor weakness (41.6 ± 2.9 years old) were quite uniform. These patients presented slowly progressive, predominantly proximal weakness and atrophy with diminished tendon reflexes in the lower extremities. Sensory impairment was generally mild. Indeed, one patient (III-4 in family 4) has been diagnosed with very slowly progressive ALS. Although frontotemporal dementia (FTD) is an occasionally observed clinical presentation in patients with ALS, dementia was not observed in these patients. Laboratory tests showed mildly elevated serum creatine kinase levels. Electrophysiological studies showed similar results in all the patients investigated and revealed a decreased number of motor units with abundant positive sharp waves, fibrillation, and fasciculation potentials. Sensory-nerve action potentials of the sural nerve were lost in the later stage of the disease. All these clinical findings were similar to those described in previous reports.[1,3,4]

To further narrow the candidate region, we conducted detailed genotyping by employing the Genome-Wide Human SNP array 6.0 (Affymetrix). Multipoint parametric linkage analysis and haplotype reconstruction were performed with the pipeline software SNP-HiTLink[8] and Allegro v.2[9] (Figure 1A). In addition to the SNP genotyping, we also used newly discovered polymorphic dinucleotide repeats for haplotype comparison (microsatellite marker 1 [MS1], chr3: 101,901,207–101,901,249; and MS2, chr3: 102,157,749–102,157,795 in hg18) around *TFG* (see Table S2 for primer sequences). The genome-wide linkage study revealed only one chromosome 3 region showing a cumulative LOD score exceeding 3.0 (Figure 1B), confirming the result of our previous study.[7] An obligate recombination event was observed between rs4894942 and rs1104964, thus further refining the telomeric boundary of the candidate region in Kansai families (Figure 2A). The Okinawan families (families 3 and 4) shared an extended disease haplotype spanning 3.3 Mb, consistent with a founder effect reported in the Okinawan HMSN-P-affected families,[1] thus defining the 3.3 Mb region as the minimum candidate region.

We then performed exon capture (Sequence Capture Human Exome 2.1 M Array [NimbleGen]) of the index patient from family 3 and subsequent passively parallel sequencing by using two lanes of GAIIx (100 bp single end [Illumina]) and a one-fifth slide of SOLiD 4 (50 bp single end [Life Technologies]). GAIIx and SOLiD4 yielded 2.60 and 2.76 Gb of uniquely mapped reads,[10] respectively. The average coverages were 29.0× and 26.8× in GAIIx and SOLiD4, respectively (Table S3 and Figure S1). In summary, 175,236 single nucleotide variants (SNVs) and 25,987 small insertions/deletions were called.[11] The numbers of exonic and splice-site variants were 14,189 and 127, respectively. In the minimum candidate region of 3.3 Mb, only 11 exonic SNVs were found, and only one was novel (i.e., not found in dbSNP) and nonsynonymous. Direct nucleotide-sequence analysis confirmed the presence of heterozygous SNV c.854C>T (p.Pro285Leu) in TRK-fused gene (*TFG* [NM_006070.5]) in all the patients from families 3 and 4 (Figure 3A and Figure S2[12]). Intriguingly, direct nucleotide-sequence analysis of all *TFG* exons (see Table S4 for primer sequences) of one patient from each of families 1 and 2 from the Kansai area revealed an identical c.854C>T (p.Pro285Leu) *TFG* mutation cosegregating with the disease (Figure 1A and Figure 3A). The base substitution was not observed in 482 Japanese controls (964 chromosomes), dbSNP, the 1000 Genomes Project Database, or the Exome Sequencing Project Database. Pro285 is located in the P/Q-rich domain in the C-terminal region of TFG (Figure 3B) and is evolutionarily conserved (Figure 3C). PolyPhen predicts it to be "probably damaging." Because some of the exonic sequences were not sufficiently covered by exome sequencing (i.e., their read depths were no more than 10×) (Figure S1), direct nucleotide-sequence analysis was further conducted for these exonic sequences (Table S5). However, it did not reveal any other novel

nonsynonymous variants, confirming that c.854C>T (p.Pro285Leu) is the only mutation exclusively present in the candidate region of 3.3 Mb. All together, we concluded that it was the disease-causing mutation.

Because we found an identical mutation in both Kansai (families 1 and 2) and Okinawan (families 3 and 4) families, we then compared the haplotypes with the c.854C>T (p.Pro285Leu) mutation in the Kansai and Okinawan families in detail. To obtain high-resolution haplotypes, we included custom-made markers, including MS1 and MS2, and new SNVs identified by our exome analysis, in addition to the high-density SNPs used in the linkage analysis. The two Kansai families shared as long as 24.0 Mb of haplotype, and the two Okinawan families shared 3.3 Mb, strongly supporting a common ancestry in each region. When the haplotypes of the Kansai and Okinawan families were compared, it turned out that these families do not share the same haplotype because the markers nearest to *TFG* are discordant at markers 48.5 kb centromeric and 677 bp telomeric to the mutation within a haploblock (Figure 2B). Although the possibility of rare recombination events just distal to the mutation cannot be completely excluded, as suggested by the population-based recombination map (Figure 2B), these findings strongly support the interpretation that the mutations have independent origins and provide further evidence that *TFG* contains the causative mutation for this disease.

Mutational analyses of *TFG* were further conducted in patients with other diseases affecting lower motor neurons (including familial ALS [n = 18], axonal HMSN [n = 26], and hereditary motor neuropathy [n = 3]) and revealed no mutations in *TFG*, indicating that c.854C>T (p.Pro285Leu) in *TFG* is highly specific to HMSN-P.

In this study, we identified in all four families a single variant that appears to have developed on two different haplotypes. The mutation disrupts the PXXP motif, also known as the Src homology 3 (SH3) domain, which might affect protein-protein interactions. In addition, substitution of leucine for proline is expected to markedly alter the protein's secondary structure, which might substantially compromise the physiological functions of TFG.

By employing the primers shown in Table S6, we obtained full-length cDNAs by PCR amplification of the cDNAs prepared from a cDNA library of the human fetal brain (Clontech). During this process, four species of cDNA were identified (Figure S3A). To determine the relative abundance of these cDNA species, we used the primers shown in Table S7 to conduct fragment analysis of the RT-PCR products obtained from RNAs extracted from various tissues; these primers were designed to discriminate four cDNA species on the basis of the size of the PCR products. The analysis revealed that TFG is ubiquitously expressed, including in the spinal cord and dorsal root ganglia, which are the affected sites of HMSN-P (Figure S3B).

Neuropathological studies were performed in a *TFG*-mutation-positive patient (IV-25 in family 1) who died of