

Report

Glycyl tRNA Synthetase Mutations in Charcot-Marie-Tooth Disease Type 2D and Distal Spinal Muscular Atrophy Type V

Anthony Antonellis,^{1,5} Rachel E. Ellsworth,¹ Nyamkhisig Sambuughin,⁴ Imke Puls,² Annette Abel,² Shih-Queen Lee-Lin,¹ Albena Jordanova,⁶ Ivo Kremensky,⁶ Kyproula Christodoulou,⁷ Lefkos T. Middleton,⁷ Kumaraswamy Sivakumar,⁴ Victor Ionasescu,⁸ Benoit Funalot,⁹ Jeffery M. Vance,¹⁰ Lev G. Goldfarb,³ Kenneth H. Fischbeck,² and Eric D. Green¹

¹Genome Technology Branch, National Human Genome Research Institute, ²Neurogenetics Branch and ³Clinical Neurogenetics Unit, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda; ⁴Barrow Neurological Institute, Phoenix; ⁵Graduate Genetics Program, The George Washington University, Washington, DC; ⁶Laboratory of Molecular Pathology, Sofia Medical University, Sofia; ⁷Molecular Genetics Department D, The Cyprus Institute of Neurology and Genetics, Nicosia; ⁸Division of Medical Genetics, Department of Pediatrics, University of Iowa, Iowa City; ⁹Department of Neurology and INSERM U573, Hôpital Sainte-Anne, Paris; and ¹⁰Center for Human Genetics, Institute for Genomic Sciences and Policy, Duke University, Durham, NC

Charcot-Marie-Tooth disease type 2D (CMT2D) and distal spinal muscular atrophy type V (dSMA-V) are axonal peripheral neuropathies inherited in an autosomal dominant fashion. Our previous genetic and physical mapping efforts localized the responsible gene(s) to a well-defined region on human chromosome 7p. Here, we report the identification of four disease-associated missense mutations in the glycyl tRNA synthetase gene in families with CMT2D and dSMA-V. This is the first example of an aminoacyl tRNA synthetase being implicated in a human genetic disease, which makes genes that encode these enzymes relevant candidates for other inherited neuropathies and motor neuron diseases.

Charcot-Marie-Tooth (CMT) disease constitutes a heterogeneous group of peripheral neuropathies estimated to affect 1 in 2,500 individuals (Skre 1974). The clinical features of CMT include muscular weakness and atrophy in the distal extremities, steppage gait, pes cavus, absent or diminished deep-tendon reflexes, and impaired sensation (Murakami et al. 1996). Through the measurement of motor nerve conductance velocities (MNCVs), CMT can be subdivided into two classes (Dyck and Lambert 1968). In CMT1, patients exhibit decreased MNCVs with demyelinating axons. In CMT2, patients exhibit normal MNCVs and no demyelination but have decreased amplitudes of evoked motor and sensory nerve responses. To date, six subtypes of CMT2 have been reported (CMT2A–F), with the genes responsible for

three of these now identified: (1) CMT2A: kinesin superfamily gene (*KIF1B*) (Zhao et al. 2001); (2) CMT2B: RAS-related GTP-binding protein 7 gene (*RAB7*) (Verhoeven et al. 2003); and (3) CMT2E: neurofilament light chain gene (*NEFL*) (Mersyanova et al. 2000).

CMT2D (MIM 601472) was first mapped to chromosome 7p in a large North American family exhibiting a peripheral neuropathy that was more pronounced in the upper extremities (family 1 in table 1) (Ionasescu et al. 1996). Subsequently, an additional North American family was identified with a CMT2 phenotype that mapped to 7p (family 2 in table 1) (Pericak-Vance et al. 1997). The CMT2D region overlaps an interval shown elsewhere to contain a locus for distal spinal muscular atrophy type V (dSMA-V) mapped in a Bulgarian kindred (family 3 in table 1) (Christodoulou et al. 1995). dSMA-V is a neuromuscular disorder with a phenotype similar to CMT2D; both diseases are associated with a more severe phenotype in the upper extremities that generally affects the thenar eminence and first dorsal interosseous muscle groups. The main characteristic that distinguishes these disorders is a distal sensory loss in patients with

Received January 28, 2003; accepted for publication February 20, 2003; electronically published April 10, 2003.

Address for correspondence and reprints: Dr. Eric D. Green, Building 50, Room 5222, 50 South Drive, MSC 8002, Bethesda, MD 20892-8002. E-mail: egreen@nhgri.nih.gov

© 2003 by The American Society of Human Genetics. All rights reserved.
0002-9297/2003/7205-0023\$15.00

Table 1**Families with CMT2D and dSMA-V Analyzed in This Study**

Family	Classification	Age at Onset (years)	Symptoms More Severe in Upper Extremities	Symptoms Prominent in Thenar Eminence and First Dorsal Interosseous	References
1	CMT2	16–30 ^a	+	ND	Ionasescu et al. 1996
2	CMT2	ND	ND	ND	Pericak-Vance et al. 1997
3	dSMA	17 ^b	+	+	Christodoulou et al. 1995
4	CMT2/dSMA	18 ^b	+	+	Sambuughin et al. 1998
5	dSMA	ND	+	+	The present study (see fig. 3D)

NOTE.—+ = finding present; ND = not determined.

^a Range in age at onset.

^b Median age at onset.

CMT2D. Genetic mapping of a Mongolian kindred (family 4 in table 1), with phenotypic features of both CMT2D and dSMA-V, implicated the same region on chromosome 7p, thereby raising the likely possibility that the two diseases are allelic (Sambuughin et al. 1998). In addition, we recently identified an Algerian Sephardic Jewish family presenting with a dSMA phenotype (family 5 in table 1). Affected members are present in multiple generations, and the pattern of inheritance is consistent with an autosomal dominant trait. A 53-year-old man and his 27-year-old daughter presented with symptoms of bilateral hand amyotrophy. Hand weakness began at ages 13 and 26 years, respectively, and physical examination revealed decreased muscle strength limited to thenar and dorsal interosseous muscles of both hands. Strength was deemed normal in all other muscles in both patients. Tendon reflexes were normal, and physical exam revealed no sensory impairment. Electromyography showed a pure motor neuropathy limited to the hands, with preserved conduction velocities; sensory conduction studies were normal. On the basis of these studies, we classified this family as having dSMA type V.

Our previous studies indicated that the CMT2D/dSMA-V critical region encompasses an ~1.25-Mb region on chromosome 7p14 (Ellsworth et al. 1999). To refine this interval, additional genotyping was performed on families 1, 2, and 4. We identified a single recombination event at marker 7p-3180 (a novel dinucleotide repeat at chr7:29490559–29490683 of the November 2002 build on the UCSC Genome Browser) in the unaffected offspring of an affected individual from family 4 (data not shown). This finding narrows the CMT2D/dSMA-V critical region to ~980 kb between markers 7p-3180 (telomeric) and *D7S632* (centromeric) (fig. 1A). Of note, families 1 and 2 bear the same haplotype across this region, raising the possibility that they carry the same CMT2D allele (data not shown).

Analysis of the CMT2D/dSMA-V critical region reveals the presence of 11 known genes (fig. 1A). These genes were analyzed for mutations by sequencing PCR-amplified exons generated from representative members of each family. Screening of the glycyI tRNA synthetase

gene (*GARS* [MIM 600287]) revealed mutations in all five families (fig. 2A). Specifically, the following four heterozygous missense mutations were detected: (1) in families 1 and 2, a 1236g→c variant that results in a predicted G240R amino acid change (note that this is consistent with the above haplotype data for these two families); (2) in family 3, a 904c→t variant that results in a predicted L129P amino acid change; (3) in family 4, a 730a→g variant that results in a predicted E71G amino acid change; and (4) in family 5, a 2094g→c nucleotide change that results in a predicted G526R amino acid change.

Analysis of available individuals from each pedigree revealed that the mutations consistently segregate with the disease, with representative examples shown in figure 3A, 3C, and 3D. We also tested human-rodent hybrid cell lines derived from family 1, one carrying the “affected” (i.e., CMT2D-associated) chromosome 7 haplotype and one carrying an “unaffected” chromosome 7 haplotype. Sequence analysis revealed that the G240R mutation resides only on the copy of chromosome 7 bearing the affected haplotype (fig. 3B).

To rule out the possibility that the above sequence variants represent rare polymorphisms, we genotyped appropriate control populations by PCR amplification and DNA sequencing or RFLP analysis. None of these mutations were encountered in the *GARS* gene of any control sample. For all mutations, screening was performed with a mixed-ethnic population collected in North America (described in Struewing et al. [1995]); for the mutations found in the Mongolian, Bulgarian, and Sephardic Jewish families, additional screening was performed with suitably matched controls. Our results revealed that: (1) E71G was absent in 398 unrelated chromosomes (130 from a Mongolian population and 268 from the mixed-ethnic population); (2) L129P was absent in 376 unrelated chromosomes (200 from an Eastern European population and 176 from the mixed-ethnic population); (3) G240R was absent in 368 unrelated chromosomes from the mixed-ethnic population; and (4) G526R was absent in 360 unrelated chromosomes (160 from a Sephardic Jewish population and 200 from the mixed-ethnic population).

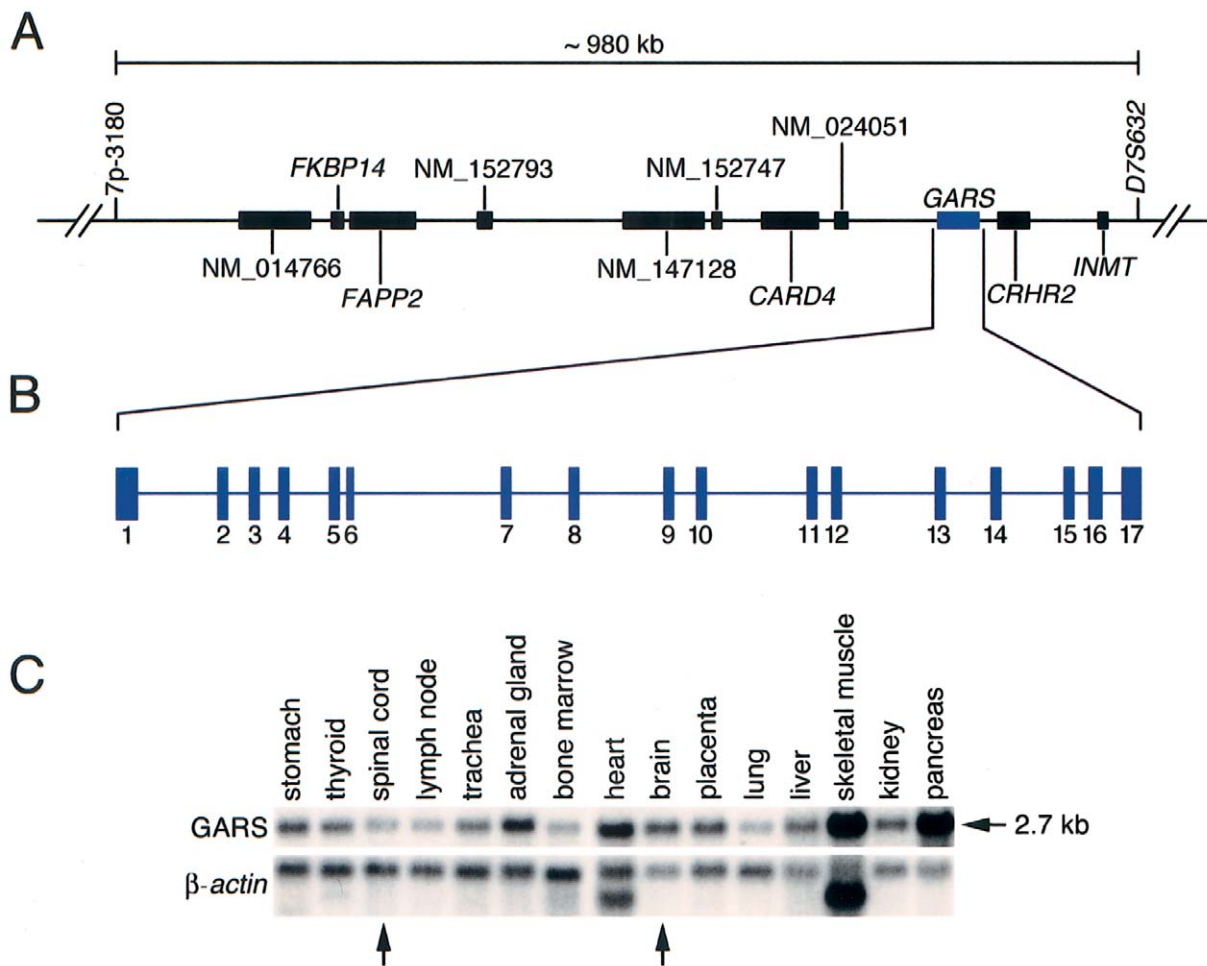


Figure 1 Localization, organization, and expression of the *GARS* gene. *A*, The interval of human chromosome 7p14 harboring the CMT2D/dSMA-V gene(s) is defined by genetic markers 7p-3180 and D7S632, spans ~980 kb, and contains the indicated 11 known genes (all with corresponding GenBank records). *B*, The *GARS* gene spans ~40 kb and contains 17 exons. *C*, Northern analysis reveals that the ~2.7-kb *GARS* transcript is ubiquitously expressed, with notable positive expression in the brain and spinal cord (arrows).

The human *GARS* protein is encoded by a 17-exon gene that spans ~40 kb on chromosome 7p14 (fig. 1*B*) and that is expressed in a ubiquitous fashion (fig. 1*C*), including the brain and spinal cord (tissues relevant for neurodegenerative diseases). Examination of the four CMT2D/dSMA-V-associated mutations reveals striking conservation of the altered amino acids. Specifically, in all but one case, the variant amino acid is identical in organisms distributed across wide taxonomic spans, from primates to yeast (fig. 2*C*; see the “Electronic-Database Information” section for GenBank accession numbers for multispecies amino acid sequences). The one exception is the human residue G240 (a glycine). In roundworm, there is an alanine at this position (albeit these two amino acids are quite similar in terms of side chain size and charge).

The above genetic data implicate mutations in the *GARS* gene as the cause of the neurodegenerative dis-

orders, CMT2D and dSMA-V. *GARS* is a member of the family of aminoacyl tRNA synthetases responsible for charging tRNAs with their cognate amino acids. The functional holoenzyme exists as a homodimer (reviewed in Freist et al. [1996]) and contains three major functional domains (fig. 2*B*): (1) the WHEP-TRS domain (residues 13–63; pfam00458 in the NCBI Conserved Domain Database) for conjugation with other aminoacyl tRNA synthetases in enzyme complexes; (2) the core catalytic domain (residues 92–168 and 241–324; pfam 00587) for ligation; and (3) the anticodon-binding domain (residues 557–655; pfam03129) for recognition of glycine-specific tRNAs. One of the identified mutations (L129P) falls within the catalytic core (fig. 2*B*), whereas another (G240R) lies one residue upstream of the catalytic core (fig. 2*B*) and two residues upstream of a highly conserved LRPETAQ sequence. This stretch of amino acids is fully conserved between human and *Ther-*

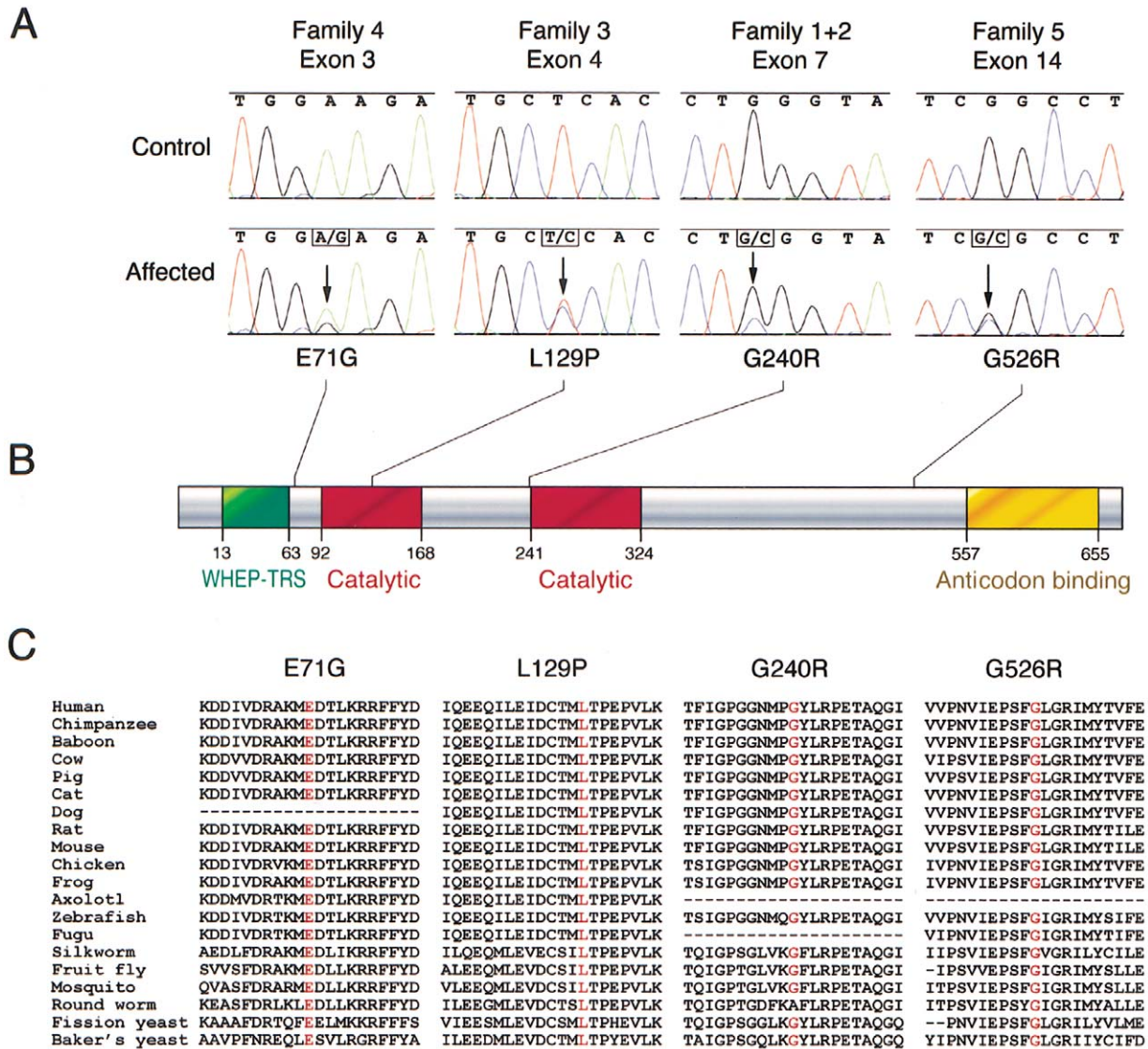


Figure 2 Characterization, localization, and conservation of GARS mutations. *A*, Representative sections of sequence chromatograms are shown for the regions encompassing the identified GARS mutations in the indicated families. Arrows denote each mutation (present in a heterozygous state), with the resulting amino acid changes depicted below along with an indication of their relative positions in the GARS protein. *B*, The known functional domains of the GARS protein are indicated in green (WHEP-TRS enzyme conjugation domain), red (core catalytic domain), and yellow (Gly-tRNA anticodon binding domain). *C*, For each of the four detected mutations, the variant amino acid is shown along with the flanking GARS sequence in multiple, evolutionarily diverse species. Note that each specific amino acid change is given at the top, with the relevant position depicted in red for the sequence in each species. Dashes indicate where sequence data were unavailable.

mus thermophilus; in the latter species, it composes a loop and helix structure essential for forming the glycine-binding pocket (reviewed in Freist et al. [1996]). The positions of these two mutations in conjunction with the presence of mutations across the GARS protein suggest a corresponding loss or decrease in enzyme activity. It is possible that these missense mutations act in a dominant negative fashion, such that the presence of a mutant subunit within dimers (i.e., mutant/wild type, mutant/mutant) greatly reduces overall GARS activity.

To investigate further the possible functional impact of these mutations, we attempted to model the affected amino acids on three-dimensional structures available for the *T. thermophilus* GARS protein. Human GARS exhibits high overall similarity to these structures; however, the absolute identity between the human and *T. thermophilus* GARS is only 27%–29% (A. Baxevanis, personal communication). Although some of the identified mutations fall within structured regions, standard threading techniques (e.g., Bryant and Lawrence 1993) to examine

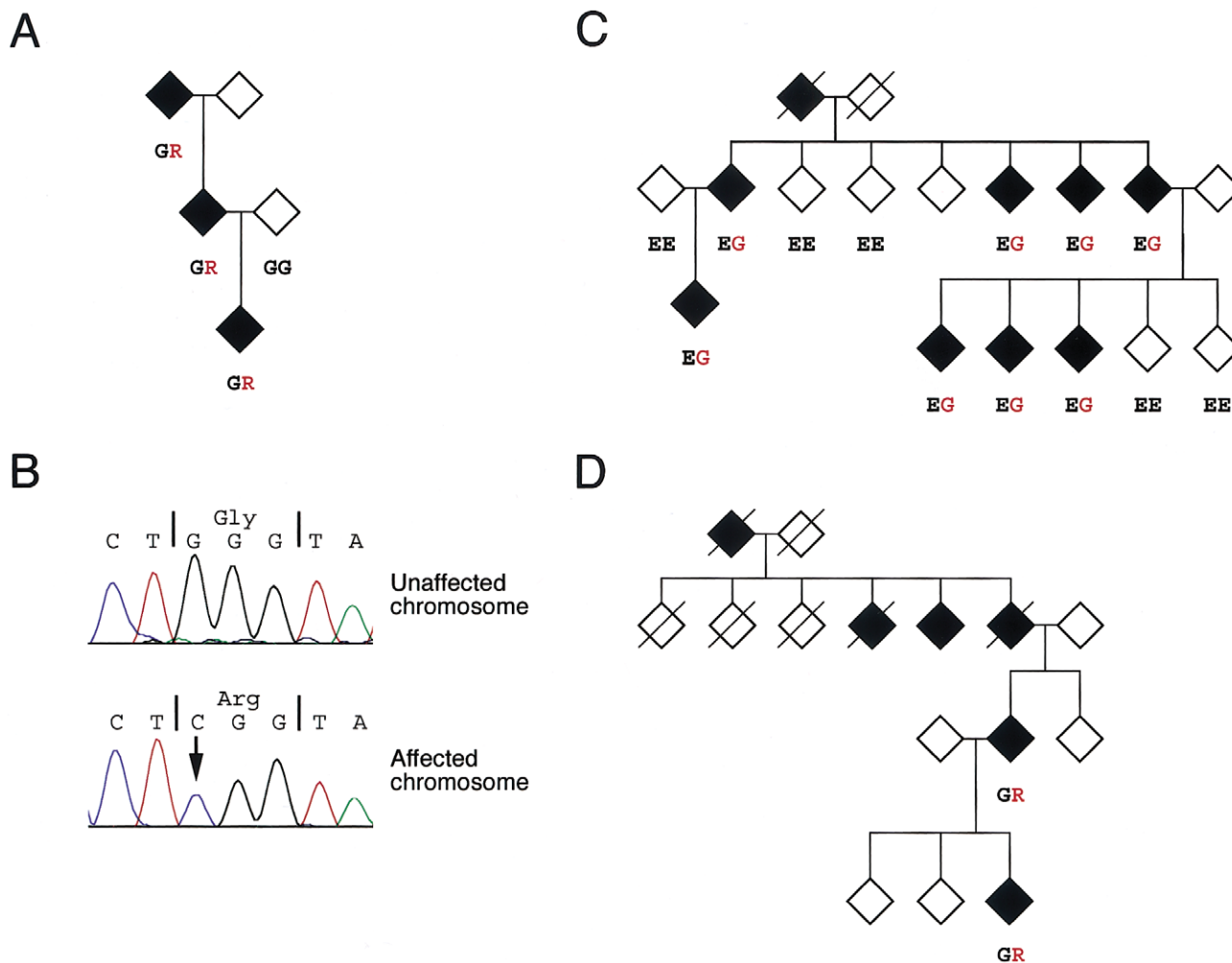


Figure 3 Segregation of *GARS* mutations with CMT2D/dSMA-V. *A*, Representative branch of family 1, showing segregation of the G240R mutation (amino acid change shown in red) with CMT2D. *B*, Sequence analysis of the G240R mutation in separated affected and unaffected copies of human chromosome 7 from family 1 (see text). Note that the 2094c allele (arrow) is present only on the affected chromosome (lower panel). *C*, Representative branch of family 4 showing segregation of the E71G mutation (amino acid change shown in red) with CMT2D/dSMA-V. *D*, Full pedigree of family 5. Note the autosomal dominant mode of inheritance and the presence of the G526R mutation in both individuals with dSMA-V (DNA from other family members was unavailable).

the effect of these mutations on the thermodynamic stability of *GARS* cannot be applied. As a rule, a minimum of 40% sequence identity is required for the results to be biologically meaningful (A. Baxeavanis, personal communication).

The mechanism by which mutations in such a ubiquitously expressed gene lead to such a highly specific phenotype (i.e., peripheral neuropathy/neuronopathy) requires additional investigation. There are two possibilities, in the likely event that the molecular pathology is due to decreased enzyme activity. First, the phenotype may arise owing to a general defect in translation efficiency, since every protein-bearing glycine would be affected. In the environment of a typical cell, this defect

may not give rise to a detectable phenotype. However, cells bearing long axons may be more prone to ensuing pathology owing to a reduction of protein products reaching axon termini. Second, there may be a specific need for more glycine-rich proteins in neurons affected in CMT2D and dSMA-V. Also intriguing is the fact that the pathologic mechanism must account for the more prominent phenotype in specific muscle groups of the hand. In addition to studies designed to resolve these issues, it will be of interest to perform *GARS* mutation screening in: (1) a broader group of patients with CMT2 and dSMA to assess the frequency of mutations in these disorders; and (2) patients with neuropathology affecting the thenar eminence in a predominant fashion, for

example, as seen in carpal tunnel syndrome (Sternbach 1999).

These findings represent the first example of a defect in an aminoacyl tRNA synthetase being directly associated with a human genetic disease. It is interesting that two mutant forms of the superoxide dismutase 1 gene (*SOD1* [MIM 147450]), the gene implicated in the neurodegenerative disease amyotrophic lateral sclerosis 1 (*ALS1* [MIM 105400]), have been shown to interact directly with lysyl tRNA synthetase (*KARS*) (Kunst et al. 1997). Since these interactions are not observed with wild-type *SOD1*, this raises the possibility that *KARS* inhibition may be involved in the pathology of *ALS1*. Thus, the genes for all aminoacyl tRNA synthetases should be considered relevant candidates for inherited neuropathies and motor neuron diseases.

Acknowledgments

We thank the members of the families for their participation in this study. We also thank: Larry Brody for providing screening populations, the Washington University Genome Sequencing Center for sequencing, Don Hadley for sample collection, Veneta Georgieva for family 3 pedigree data, Andy Baxevasian for protein modeling, and Robert Nussbaum and Francis Collins for critical review of the manuscript. This work was supported in part by grant NS 26630 (to J.M.V.). All studies performed herein were approved by the National Institute of Neurological Disorders and Stroke Institutional Review Board, with informed consent obtained from all subjects.

Electronic-Database Information

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

GenBank, <http://www.ncbi.nih.gov/GenBank/> (for *GARS* cDNA [accession number NM_002047], *GARS* protein [accession number NP_002038], *FKBP14* [accession number NM_017946], *FAPP2* [accession number NM_032639], *CARD4* [accession number NM_006092], *CRHR2* [accession number NM_001883], and *INMT* [accession number NM_006774]). Multispecies amino acid sequences were derived from the following accession numbers: chimpanzee [*Pan troglodytes*, accession number AC091720], baboon [*Papio anubis*, accession number AC091658], cow [*Bos taurus*, accession number AC092083], pig [*Sus scrofa*, accession number AC091755], cat [*Felis catus*, accession number AC092009], dog [*Canis familiaris*, accession numbers BM538632 and AC092088], rat [*Rattus norvegicus*, accession number AC091711], mouse [*Mus musculus*, accession number AAH21747.1], chicken [*Gallus gallus*, accession number AC092081], frog [*Xenopus laevis*, accession numbers BJ031186, BQ730838, BE680980, and BE669219], axolotl [*Ambystoma mexicanum*, accession number BI818038], zebrafish [*Danio rerio*, accession number AC099322], fugu

[*Takifugu rubripes*, accession number AC098806], silkworm [*Bombyx mori*, accession number Q04451], fruit fly [*Drosophila melanogaster*, accession number AAF49668.1], mosquito [*Anopheles gambiae*, accession number EAA10181.1], roundworm [*Caenorhabditis elegans*, accession number NP_498093.1], fission yeast [*Schizosaccharomyces pombe*, accession number NP_593935.1], and baker's yeast [*Saccharomyces cerevisiae*, accession number NP_009679.1])
 NCBI Conserved Domain Database, <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=cdd>
 Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for *CMT2D*, *GARS*, *SOD1*, and *ALS1*)
 UCSC Genome Browser, <http://genome.ucsc.edu> (for human genome)

References

- Bryant SH, Lawrence CE (1993) An empirical energy function for threading protein sequence through the folding motif. *Proteins* 16:92–112
- Christodoulou K, Kyriakides T, Hristova AH, Georgiou DM, Kalaydjieva L, Yshpekova B, Ivanova T, Weber JL, Middleton LT (1995) Mapping of a distal form of spinal muscular atrophy with upper limb predominance to chromosome 7p. *Hum Mol Genet* 4:1629–1632
- Dyck PJ, Lambert EH (1968) Lower motor and primary sensory neuron diseases with peroneal muscular atrophy. *Arch Neurol* 18:619–625
- Ellsworth RE, Ionasescu V, Searby C, Sheffield VC, Braden VV, Kucaba TA, McPherson JD, Marra MA, Green ED (1999) The *CMT2D* locus: refined genetic position and construction of a bacterial clone-based physical map. *Genome Res* 9:568–574
- Freist W, Logan DT, Gauss DH (1996) Glycyl-tRNA synthetase. *Biol Chem Hoppe Seyler* 377:343–356
- Ionasescu V, Searby C, Sheffield VC, Roklina T, Nishimura D, Ionasescu R (1996) Autosomal dominant Charcot-Marie-Tooth axonal neuropathy mapped on chromosome 7p (*CMT2D*). *Hum Mol Genet* 5:1373–1375
- Kunst CB, Mezey E, Brownstein MJ, Patterson D (1997) Mutations in *SOD1* associated with amyotrophic lateral sclerosis cause novel protein interactions. *Nat Genet* 15:91–94
- Mersyanova IV, Perepelov AV, Polyakov AV, Sitnikov VF, Dadali EL, Oparin RB, Petrin AN, Evgrafov OV (2000) A new variant of Charcot-Marie-Tooth disease type 2 is probably the result of a mutation in the neurofilament-light gene. *Am J Hum Genet* 67:37–46
- Murakami T, Garcia CA, Reiter LT, Lupski JR (1996) Charcot-Marie-Tooth disease and related neuropathies. *Medicine* 75:233–250
- Pericak-Vance MA, Speer MC, Lennon F, West SG, Menold MM, Stajich JM, Wolpert CM, Slotterbeck BD, Saito M, Tim RW, Rozear MP, Middleton LT, Tsuji S, Vance JM (1997) Confirmation of a second locus for *CMT2* and evidence for additional genetic heterogeneity. *Neurogenetics* 1:89–93
- Sambuughin N, Sivakumar K, Selenge B, Lee HS, Friedlich D, Baasanjav D, Dalakas MC, Goldfarb LG (1998) Autosomal dominant distal spinal muscular atrophy type V (dSMA-V)

- and Charcot-Marie-Tooth disease type 2D (CMT2D) segregate within a single large kindred and map to a refined region on chromosome 7p15. *J Neurol Sci* 161:23–28
- Skre H (1974) Genetic and clinical aspects of Charcot-Marie-Tooth's disease. *Clin Genet* 6:98–118
- Sternbach G (1999) The carpal tunnel syndrome. *J Emerg Med* 17:519–523
- Struewing JP, Abeliovich D, Peretz T, Avishai N, Kaback MM, Collins FS, Brody LC (1995) The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. *Nat Genet* 11:198–200
- Verhoeven K, De Jonghe P, Coen K, Verpoorten N, Auer-Grumbach M, Kwon JM, FitzPatrick D, Schmedding E, De Vriendt E, Jacobs A, Van Gerwen V, Wagner K, Hartung HP, Timmerman V (2003) Mutations in the small GTP-ase late endosomal protein RAB7 cause Charcot-Marie-Tooth type 2B neuropathy. *Am J Hum Genet* 72:722–727
- Zhao C, Takita J, Tanaka Y, Setou M, Nakagawa T, Takeda S, Yang HW, Terada S, Nakata T, Takei Y, Saito M, Tsuji S, Hayashi Y, Hirokawa N (2001) Charcot-Marie-Tooth disease type 2A caused by mutation in a microtubule motor *KIF1B β* . *Cell* 105:587–597