Background: A 1.5-Mb microduplication containing the gene for peripheral myelin protein 22 (PMP22) on chromosome 17p11.2-12 is responsible for 75% of cases of the demyelinating form of Charcot-Marie-Tooth disease (CMT1A). Methods for molecular diagnosis of CMT1A use Southern blot and/or amplification by PCR of polymorphic poly(AC) repeats (microsatellites) located within the duplicated region, or the detection of junction fragments specific for the duplication. Difficulties with both strategies have led us to develop a new diagnostic strategy with highly polymorphic short tandem repeats (STRs) located inside the CMT1A duplicated region.

Methods: We tested 10 STRs located within the duplication for polymorphic behavior. Three STRs were selected and used to test a set of 130 unrelated CMT1A patients and were compared with nonduplicated controls. The study was then extended to a larger population of patients. Alleles of interest were sequenced. A manual protocol using polyacrylamide electrophoresis and silver staining and an automated capillary electrophoresis protocol to separate fluorescently labeled alleles were validated.

Results: We identified three new STRs covering 0.55 Mb in the center of the CMT1A duplication. One marker, 4A, is located inside the PMP22 gene. The two others, 9A and 9B, more telomerically positioned, have the highest observed heterozygosity reported to date for CMT1A markers: 0.80 for 9A, and 0.79 for 9B. Tetra- and pentanucleotide repeats offered clear amplification, accurate sizing, and easy quantification of intensities.

Conclusions: Combined use of the three STRs allows robust diagnosis with almost complete informativeness. In our routine diagnosis for CMT1A, they have replaced the use of other polymorphic markers, either in a manual adaptation or combined with fluorescence labeling and allele sizing on a DNA sequencer.

© 2001 American Association for Clinical Chemistry

Charcot-Marie-Tooth (CMT) disease is the most frequent inherited peripheral neuropathy, with an estimated prevalence of ~1 in 2500 (1). Two major forms of CMT can be identified electrophysiologically: one form shows defects in the formation or maintenance of myelin (CMT1) and the other primary axonal degeneration (CMT2) (2). A microduplication of 1.5 Mb containing the gene for peripheral myelin protein 22 (PMP22) on chromosome 17p11.2 is responsible for 75% of cases of the demyelinating form (CMT1A) (3–9). Diagnosis with markers located inside the duplication usually is carried out using restriction fragment length polymorphism (RFLP) probes (3, 4, 10) and poly(AC) repeats (11–13), and less frequently with fluorescence in situ analysis (14) or pulse-field gel electrophoresis (15). Molecular diagnosis often relies on the interpretation of differences in allele intensities, even with the most polymorphic markers that have been reported to date.

When poly(AC) repeats are used for molecular diagnosis, artifact bands produced by slippage of the polymerase enzyme may lead to difficulties in interpretation of dosage for different alleles. Other useful diagnostic methods have been developed after extensive investigation of
the two repeated 30-kb sequences (REPs) flanking the duplicated region (16). Recombination events leading to the CMT1A duplication occur almost entirely in four adjacent regions within 7.8 kb of the CMT1A-REP sequences, with a “hot spot” of recombination located in a 3.2-kb central segment defined by two restriction enzyme sites, EcoRI and SacI (17, 18). Characterization of the junction fragment resulting from recombination allows positive identification of the duplication with RFLP probes on Southern blots (17, 18).

Several methods with direct allele-specific amplification of the junction fragments by PCR have also been reported (19–22). However, to date, no PCR strategy covering all of the possible recombination events has been reported; thus, the overall efficiency of this approach is difficult to evaluate. Moreover, infrequent variation in the CMT1A-REP sequences (23–25), chimerism of the duplication junction fragments (26), and possible recombination outside the CMT1A-REP sequences (27, 28) make the methods based on identification of the CMT1A junction fragments prone to misinterpretation. We have developed a new strategy with polymorphic short tandem repeats (STRs) located inside the CMT1A duplicated region that allow amplification with very low or no stuttering, accurate sizing, and visual quantification of allele intensity, which offers the choice between inexpensive detection without labeling, using nondenaturing polyacrylamide gels and silver staining, or the use of fluorescent-labeled primers and separation on an automated sequencing apparatus.

**Materials and Methods**

**Choice of STRs**

We collected overlapping sequences of BAC clones directly submitted by the MIT Center for Genome Research, Cambridge MA (29) in GenBank. A contig of the CMT1A duplication was manually constructed using databank searches with terminal sequences of the clones using the BLAST program (30). The positions of the PMP22 gene and of the polymorphic DNA markers most in use were assigned by BLAST analysis. GenBank Accession Numbers for the BAC clones and localization of the markers are given in Fig. 1. The positions of the CMT1A-REP sequences were confirmed on the BAC clones using sequences of the cosmid clone c74F4 for the proximal CMT1A-REP (17) and c15H12 for the distal CMT1A-REP sequence (31). The arrangement of the sequenced clones and the positions of markers were consistent with the cosmid maps of the duplication (32, 33). BAC clones covered the CMT1A duplication without any gaps. The size of the contig was 1.43 Mb. We selected 10 potential STR sequences with (a) ≥5 repeats and (b) the possibility of finding unique specific primers for amplification (Table 1). Primers were tested for uniqueness using the BLAST program. A number was given to each STR according to its position from the proximal part of the duplication (e.g., 4A within the 0.3 and 0.4 Mb interval).

**Manual Adaptation**

PCR amplification was carried out in standard conditions in a final volume of 25 μL in the presence of 20 pM of each primer and 0.5 U of EuroblueTaq® ADN polymerase (Laboratoires Eurobio) for 35 cycles. Final reaction conditions were 67 mM Tris-HCl, 16 mM (NH₄)₂SO₄, and 0.1 mL/L Tween 20. STRs showing locus-specific amplification and a polymorphic pattern for five unrelated persons on a 2% agarose gel were chosen for subsequent analysis. Each PCR product (50 ng) was loaded on a nondenaturing 8% polyacrylamide gel. Gels 0.75 or 1 mm thick were used indifferently. Gels were stained using the Silver Stain reagent set from Bio-Rad Laboratories, according to the manufacturer’s instructions.
EVALUATION OF STRs

Samples from 130 unrelated patients carrying a CMT1A duplication identified with probes pVAW409R3 and pEW401 were tested with the selected STRs. For each of the duplications, the recombination event had been assigned to one of the four recombination regions in the CMT1A-REP sequences using probe pNEA102 on EcoRI+SaiI digestions and probe pJ5P on EcoRI+HincII digestions (17, 18).

A complementary study was carried out with the three selected STRs for nine unrelated CMT1A patients with absence of the EcoRI+SaiI junction fragment and with no dosage differences between the alleles using the pNEA102 CMT1A-REP probe. These patients, correctly identified with markers located inside of the duplication, would have been misdiagnosed with the pNEA102 probe. They represent 9 of 480 unrelated CMT1A patients (1.9%) identified in our laboratory.

Seventy controls showing two alleles with RFLP probes D17S122 and D17S61 were also analyzed. We also analyzed 50 patients with hereditary neuropathy with liability to pressure palsies (HNPP), who were hemizygous for the PMP22 region (34), as confirmed by the observation of a deletion-specific junction fragment obtained with probe pNEA102.

SEQUENCING OF ALLELES

Sequencing of alleles was performed mainly on genomic DNA from HNPP patients. For some alleles, genomic DNA from healthy controls was amplified and cloned in pGEM®-T Easy (Promega Corporation). Sufficient clones were randomly selected and sequenced to obtain the sequence of both alleles. Sequencing was carried out with plasmid-specific T7 and SP6 primers and using fluorescently labeled dideoxynucleotide triphosphate substrates (PRISMTM Ready Reactions Rhodamine Terminator reagent set; Applied Biosystems). Allele numbering was according to proposed guidelines (35).

ALLELE FREQUENCIES

For the three selected STRs, the distribution of alleles for unrelated duplicated patients was compared with that observed for the same number of alleles in controls ($\chi^2$ test; 210 alleles for each sample). Because there was no significant difference ($P = 0.05$), all available results were pooled to calculate allele frequencies, with 486 (4A), 705 (9A), and 775 (9B) alleles, respectively.

FLUORESCENT LABELING AND SIZING ON AN AUTOMATED SEQUENCER

The forward primers for 9A and 9B were labeled with fluorescent 6-carboxyfluorescein (6-FAM), and 4A was labeled with hexachloro-6-carboxyfluorescein (HEX; MWG Biotech A.G.). PCR amplification conditions were as described above for unlabeled primers. Products were analyzed on an ABI Prism 310 using a 47-cm capillary filled with POP4 and using GeneScan® 2.1 software (all products from Applied Biosystems).

Results

The 10 selected repeated sequences were successfully amplified by PCR using a standard protocol with an
annealing temperature of 60°C. Separation of PCR products on a 2% agarose gel revealed that three STRs, 4A, 9A, and 9B, were polymorphic. The annealing temperatures were further optimized to 62, 62, and 66°C, respectively.

Complete concordance (100%) was found between results obtained with the combination of RFLP probes pVAW409R3/pEW401/pNEA102 and with the three-STR (4A, 9A, and 9B) combination. In particular, all nine duplications for which EcoRI+Sacl junction fragments were absent with the CMT1A-REP probe pNEA102 and that showed no dosage differences between alleles with this probe were also correctly identified with the three-STR combination.

STR 4A is ideally located in the first intron of the PMP22 gene, 3.5 kb upstream of exon 2. It has a complex structure (Table 2). The distribution of alleles was bimodal (Fig. 2) with a maximum value for the 14.2-repeat (ATCT)3AT(ATCT)5(ACCT)6 with length 118 bp, and another for the 17-repeat (ATCT)11(ACCT)6 with length 128 bp. The eight alleles characterized to date had sizes between 116 and 136 bp. Sequencing of four alleles revealed that at least two different sequences underlie the most common allele, 14.2 (118 bp), probably with equal frequency. This was also the case for the rarer allele, 15 (Table 2). From allele 16 on, the shorter (ACCT)6 tract seemed to remain constant, whereas the longer ATCT

<table>
<thead>
<tr>
<th>Table 2. Allele composition and frequency.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Locus 4A</strong></td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>14.2</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>15.2</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>17</td>
</tr>
<tr>
<td>18</td>
</tr>
<tr>
<td>19</td>
</tr>
</tbody>
</table>

| **B. Locus 9A**                            | Length, bp | Repeat sequence | No. of alleles | Frequency (n = 705) |
| 8 | 137 | ND | 2 | 0.003 |
| 9 | 142 | ND | 1 | 0.001 |
| 10 | 147 | ND | 7 | 0.010 |
| 11 | 152 | ND | 25 | 0.035 |
| 12 | 157 | (CAATA)12 | 75 | 0.106 |
| 13 | 162 | (CAATA)13 | 216 | 0.306 |
| 14 | 167 | (CAATA)14 | 179 | 0.254 |
| 15 | 172 | ND | 107 | 0.152 |
| 16 | 177 | (CAATA)16 | 62 | 0.088 |
| 17 | 182 | ND | 24 | 0.034 |
| 18 | 187 | ND | 7 | 0.010 |

| **C. Locus 9B**                            | Length, bp | Repeat sequence | No. of alleles | Frequency (n = 772) |
| 7 | 95 | ND | 4 | 0.005 |
| 8 | 99 | ND | 3 | 0.004 |
| 9 | 103 | (TTTC)9 | 14 | 0.018 |
| 10 | 107 | ND | 11 | 0.014 |
| 11 | 111 | (TTTC)11 | 186 | 0.241 |
| 12 | 115 | (TTTC)12 | 247 | 0.320 |
| 13 | 119 | ND | 107 | 0.139 |
| 14 | 123 | (TTTC)14 | 128 | 0.166 |
| 15 | 127 | (TTTC)15 | 55 | 0.071 |
| 16 | 131 | (TTTC)16 | 16 | 0.021 |
| 17 | 135 | ND | 1 | 0.001 |

a ND, not done.
repeat was variable. Because of an insertion of a dinucleotide in the ATCT tract, the difference in length of the shorter alleles was only two nucleotides, but these were well resolved by electrophoresis on polyacrylamide gel (Fig. 3) and by capillary electrophoresis (not shown). The heterozygosity of this locus for healthy controls was 0.69.

For STR 9A, the sequenced alleles had a uniformly repeated pentamer (CAATA)$_n$ motif. Eleven alleles were observed in a discrete symmetrical distribution around allele 13 (162 bp), starting from 8 repeats (137 bp) up to 18 (187 bp; Table 2 and Fig. 2). The heterozygosity for healthy controls was 0.80. An example of transmission of STR 9A alleles in a CMT1A family is given in Fig. 4.

The repeat motif of STR 9B was composed of a tetranucleotide, (TTTC)$_n$, and variation gave a continuous set of 11 alleles. The number of repeats ranged from 7 (95 bp) to 17 (135 bp). The maximum frequency was found for allele 12 (115 bp), followed by alleles 11 and 14 (Table 2 and Fig. 2). Observed heterozygosity for healthy controls was 0.79. Examples of fragments obtained for this locus are presented in Fig. 5.

An example including all three loci is given in Fig. 6 in a family study where both parents and two children carry the duplication and where the third child carries two duplications. The family has been described in more detail elsewhere (36). The histograms in Fig. 2 summarize the distributions for the three STRs. An example of semiautomated separation is presented in Fig. 7.

**Discussion**

Four trinucleotide, five tetranucleotide, and one pentanucleotide repeats were tested for potential polymorphic behavior. Three of these repeat sequences were polymorphic: two, 4A and 9B, were tetrameric repeats, and 9A was a pentameric repeat. Because screening was performed on an agarose gel with limited resolution, only highly polymorphic STRs with large variations of alleles were selected. The trinucleotide repeat STR 7A with sequence (GAG)$_2$(GAC)(GAG)$_2$, for which we observed
two different alleles, probably is also polymorphic, but it was not studied in further detail.

The STRs had maximum allele distributions of 14 repeats (4A and 9A) and 12 repeats (9B), respectively. For the three STRs, the lowest repeat number was 7 or 8 and ranged up to 17–19 repeats. By contrast, STRs 1A (five repeats), 5A (eight repeats), 6B (nine repeats), 8A (six repeats), 10A (six repeats) and 12A (seven repeats) did not show a large variation of alleles. Because the sequenced clones in the databank most probably represent a frequent allele, candidate STRs should be selected with 12–14 repeats to find other highly polymorphic STRs. The most polymorphic STRs have simple motif repeats (9A and 9B), but compound repeats such as 4A must also be considered when testing for polymorphisms.

The three markers cover ~0.55 Mb at the center of the CMT1A region, including the PMP22 gene sequence. The localization of STR 4A inside the genomic sequence of PMP22 is particularly valuable. At least two different sequences underlie the most frequent allele, 14.2 (118 bp), probably with equal frequency. Sequencing of this STR may help to establish the presence of two copies of the PMP22 gene when no informativeness is found at the PMP22 gene locus. On the distal side, STRs 9A and 9B are separated by only 20 kb and are located 130–150 kb upstream of the group of markers D17S125, D17S839, D17S1358, and D17S61.

The observed heterozygosity of 0.69 at locus 4A is close to the highest heterozygosities reported for poly(AC)s of the CMT1A region: D17S122, 0.74 (3); D17S793, 0.70; D17S921, 0.74 (11, 12); and D17S1358, 0.74 (13). STRs 9A and 9B are to date the most polymorphic loci within the CMT1A duplication, with observed heterozygosities of 0.80 and 0.79. When we studied 130 unrelated duplications, 78% showed, at least once, three alleles for one of the three markers 4A, 9A, and 9B. An efficiency of 85% was achieved using a combination of six poly(AC) markers (13). Overall, taking into account dosage differences between fragments, 96% (125 of 130) were informative for two or more STRs, and none was noninformative for the combination of the three markers 4A, 9A, and 9B.

Differences in allele size allow separation on a non-denaturing gel even for the shortest STR 4A alleles, which differ by only 2 bp (Fig. 2). In addition, the absence of stuttering artifactual bands, often seen after amplification of poly(AC) repeats, makes the reading easy, particularly when dosage intensities are to be interpreted (Figs. 2–4). Amplification results for poly(AC) loci D17S122 (3) and D17S1357 (13) are more difficult to interpret. The three STRs were tested successfully in a family with one child carrying four copies of the PMP22 region and illustrating all possible allele combinations (Fig. 6).

The PCR primers were selected as close as possible to

Fig. 4. Fully informative transmission of STR 9A in a family with CMT1A disease.
The affected father (I:1) transmits the diseased alleles to his first son (II:1) and the healthy allele to his second son (II:2). From the transmission study, it follows that fragments of 172 and 167 bp are located on the same paternal chromosome.

Fig. 5. Profiles from CMT1A patients (patients 1, 3, 4, and 5) and a healthy control (patient 2) for STR 9B.
Patients 4 and 5 show three fragments, whereas the profiles of patients 1 and 3 show only two fragments but with different dosage. Deduced allele composition is given at the bottom.
the repeated sequences. The presently defined optimal PCR amplification conditions are different for each STR and do not allow multiplex amplification. Overlap in fragment length occurs between loci 4A and 9B and does not allow loading on the same slot of a polyacrylamide gel. However, when locus-specific fluorescent primers are used, 4A and 9B can be readily separated and simultaneously recorded by the detection system of automated electrophoresis equipment.

We tried to evaluate this new polymorphic system for diagnostic testing of HNPP. HNPP patients are hemizygous for the region duplicated in CMT1A patients. When samples are available from two or more family generations, a transmission study with RFLPs, microsatellites, or with the three STRs presented here is most often informative. When only the patient is sampled, the methods become uninformative, and the use of a method such as the characterization of a deletion junction fragment is required to separate homozygotes and hemizygotes. In our series, most of the patients noninformative for the three STRs were revealed to be hemizygotes by the characterization of a deletion junction fragment with probe pNEA102. In a group of 213 patients (without deletion or duplication), we found 210 persons who showed two alleles for at least one of the three STRs and 3 persons who were noninformative (1.4%). This result is compatible with the 1.3% noninformative patients expected when absence of haplotype association between the three loci \(0.20 \times 0.21 \times 0.31\) is assumed. As a conclusion, HNPP can be tested with the three STRs, and undeleted patients should be detected with a probability >98%.
In conclusion, we present a novel reliable diagnostic method to detect the CMT1A duplication that combines the informativeness of highly polymorphic markers located in the center of the duplication with the simplicity and speed of PCR methods using STRs, without stuttering and interpretation problems. Both laboratories involved in the development of this new technique have chosen to adopt it as their only diagnostic approach.

We thank A. Sarcey and C. Thouny for excellent technical help, and S. Gamon for redaction of the manuscript. Financial support was obtained from the Association Française contre les Myopathies (AFM) and the Hospices Civils de Lyon (HCL). The laboratories of Molecular Neurogenetics and of Medical Genetics are members of the European CMT-Consortium sponsored by Biomed Contracts CT961614 and CT960055.

References


