Transmissible familial Creutzfeldt–Jakob disease associated with five, seven, and eight extra octapeptide coding repeats in the PRNP gene

(amyloid precursor protein/prion protein)

LEV G. GOLDFARB*, PAUL BROWN*, W. RICHARD McCOMBIE†, DMITRY GOLDBABER‡, GARY D. SWERGOLD§, PETER R. WILLS‖, LARISA CERVENAKOVA‖, HENRY BARON**, C. J. GIBBS, JR.* AND D. CARLETON GAJDUSEK*

*Laboratory of Central Nervous System Studies, and †Section of Receptor Biochemistry and Molecular Biology, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892; ‡Department of Psychiatry, State University of New York, Stony Brook, NY; §Laboratory of Biochemistry, Division of Cancer Biology and Diagnosis, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892; ‖Department of Physics, University of Auckland, New Zealand; ‖Institute of Preventive and Clinical Medicine, Bratislava, Czechoslovakia; and **Searle Pharmaceuticals, Paris, France

Contributed by D. Carleton Gajdusek, September 6, 1991

ABSTRACT The PRNP gene, encoding the amyloid precursor protein that is centrally involved in Creutzfeldt–Jakob disease (CJD), has an unstable region of five variant tandem octapeptide coding repeats between codons 51 and 91. We screened a total of 535 individuals for the presence of extra repeats in this region, including patients with sporadic and familial forms of spongiform encephalopathy, members of their families, other neurological and non-neurological patients, and normal controls. We identified three CJD families (in each of which the proband’s disease was neurologically confirmed and experimentally transmitted to primates) that were heterozygous for alleles with 10, 12, or 13 repeats, some of which had “wobble” nucleotide substitutions. We also found one individual with 9 repeats and no nucleotide substitutions who had no evidence of neurological disease. These observations, together with data on published British patients with 11 and 14 repeats, strongly suggest that the occurrence of 10 or more octapeptide repeats in the encoded amyloid precursor protein predisposes to CJD.

Transmissible familial Creutzfeldt–Jakob disease (CJD) is a subacute mental and neurological disorder with a prominent dementia and movement abnormalities typically affecting middle-aged individuals and leading to death in 3–12 months after the onset of symptoms. The disease is randomly distributed in the world with an annual mortality rate of one per million people and in 5–10% of cases shows an autosomal dominant pattern of inheritance (1, 2). A human gene (PRNP) on the short arm of chromosome 20 (3) codes for the precursor of the amyloid protein (PrP, also known as “prion protein”) that accumulates in the brains of diseased individuals, and several different point and insertion mutations in this gene have been shown to be linked to familial cases of transmissible spongiform encephalopathy (4, 5).

The first genetic abnormality described in CJD was an insert of six extra 24-nucleotide repeats in the area coding for tandem octapeptides near the N-terminal region of the PrP molecule (6, 7). A variable number of these repeats (five or six) also occurs in the bovine PrP gene (8). On the basis of newly available data, we suggest that the occurrence of 10 or more octapeptide coding repeats with irregular point substitutions is a predisposing cause of familial CJD.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

The Tested Population. A total of 535 individuals were screened for extra repeats in the PRNP gene: 117 patients with spongiform encephalopathies (including 60 familial cases); 159 first-degree relatives of patients with familial spongiform encephalopathies; 39 patients with other neurological diseases; 16 patients with non-neurological diseases; and 204 normal controls.

The American family “Kell” has a father and son with CJD of whom we tested only the proband (son) for mutations in the PRNP gene. The American family “Ald” (9) has five affected members, of whom we tested the proband, her sister, and all five of her daughters from two marriages (two affected and three unaffected). The French family “Che” has nine patients with CJD (10–12), of whom we tested the proband and two unaffected at-risk daughters of his brother (who also died of CJD).

DNA Samples. DNA was extracted from the frozen brain tissue or frozen blood of the deceased patients and from fresh or frozen anticoagulated blood of the living CJD patients, unaffected family members, and control individuals.

Amplification, Subcloning, and Sequencing of the PRNP Coding Region. For easier sequencing of the PRNP coding region, we modified the technique described earlier (13, 14) by amplifying two overlapping templates from this region. Both fragments were amplified from the genomic DNA by using the polymerase chain reaction (PCR; ref. 15) with Taq polymerase (Perkin–Elmer). The oligonucleotide primers used to produce the 5’ fragment, 5’-TACTGAGCGGC-CGCGCAGTCATTTATGGCGAACCTTG-3’ and 5’-TACTGAGTCGACAATGTATGATGGGCCTGCTCA-3’, were composed of 6 random nucleotides followed by Not I and Sal I restriction sites, respectively, and 21 and 22 nucleotides of the ends of the Northern fragment. To produce the 3’ fragment of the PRNP coding region we used oligonucleotides 5’-TACTGAGCGGCCCGCCAAATGAAGCAGACATG-GCTGTT-3’ and 5’-TACTGAGTCGACCCCTTCCATC-CCACATACGG-3’, composed in the same fashion. The PCR-amplified fragments were cleaved with Not I and Sal I restriction endonucleases and ligated into pGEM-5Zf(+) plasmid with T4 DNA ligase (Promega). The recombinant plasmid was then transfected into DH5α competent Escherichia coli cells (GIBCO/BRL), which were subsequently grown on LB agar with ampicillin (100 μg/ml). Colonies containing the recombinant plasmids were selected and se-

Abbreviations: CJD, Creutzfeldt–Jakob disease; GSS, Gerstmann–Sträussler–Scheinker syndrome; PrP, amyloid precursor protein.
Normal (regular) sequence

codon 51
CCT CAT GGC GGT GGC TGG GGG CAG
pro gly gly gly gly trp gly gin

CCT CAT GGT GGC TGG GGG CAG
pro his gly gly gly trp gly gin

CCT CAT GGT GGC TGG GGG CAG
pro his gly gly gly trp gly gin

CCC CAT GGT GGC TGG GGA CAG
pro his gly gly gly trp gly gin
codon 91

Abnormal (irregular) sequences

CCT CAT GGT GGC TGG GGA CAG
pro his gly gly gly trp gly gin

CCT CAT GGG GGT GGC TGG GGG CAG
pro his gly gly gly trp gly gin

CCC CAT GGT GGC TGG GGG CAG
pro his gly gly gly trp gly gin

Fig. 1. Nucleotide sequence of codons 51–91 in normal healthy people and irregular sequences in CJD patients.

sequenced in both directions by the dideoxy chain-termination method (16) using an automated sequencer (Applied Biosystems) (17). Complete sequencing of the PRNP coding region was performed with DNA from five CJD patients, one unaffected family member, and a control individual.

Visualization of the 5' Fragments. The 5' fragment of the PRNP gene from individuals with various numbers of extra repeats was PCR-amplified, as indicated in the previous section, and electrophoresed on a horizontal 3% NuSieve plus 1% SeaKem agarose gel (FMC). DNA was stained with ethidium bromide (0.5 µg/ml).

Experimental Transmission Studies. Healthy adult primates were inoculated intracerebrally with 0.1 ml of 5% or 10% saline suspensions of brain tissue from the neuropathologically verified probands in all three families. The tissue from the Kel family proband was inoculated into a squirrel monkey and a spider monkey. The tissue from the Ald family proband was inoculated into a chimpanzee, a squirrel monkey, and a capuchin monkey. The tissue from the Che family proband was inoculated into a chimpanzee and a stump-tailed monkey. All dying animals were sacrificed when in a terminal state of illness, and their brains were examined histologically for evidence of spongiform encephalopathy.

RESULTS

Normal Structure of the Repeat Region in the PRNP Gene.

In 99% of our collection of 535 genetically tested individuals, the repeat region started at codon 51 and extended to codon 91, consisting of the five repeating structural elements shown in Fig. 1. An initial 27-nucleotide sequence, designated R1 (11), with an additional glycine codon, and a CAT-to-CAG substitution at the second codon that converts glutamine to histidine, is followed by four 24-nucleotide sequences. The second and third repeats, designated R2 (11), are identical, and they differ from the fourth repeat element, designated R3, by substitutions in positions 3 (T to C) and 21 (G to A). The fifth repeat, R4, is similar to R2, with a substitution at position 24 (G to A).

Nine Repeats. One of the control patients (Hay) died at age 63 of advanced micronodular cirrhosis. This individual had no family history of neurological disease and no clinical or pathological signs of spongiform encephalopathy, and brain tissue did not transmit disease to experimental primates. The open reading frame of the PRNP gene was completely sequenced, and instead of the usual five octamer repeats, nine repeats were found (Table 1). No irregular nucleotide substitutions occurred anywhere in the sequence.

Ten Repeats. Ten repeats were present in the proband of the American Kel family. The sixth repeat had an irregular single-nucleotide change from A to G at the 21st position, which we designate R3g because the start resembles R3. This substitution does not change the encoded amino acid.

Twelve Repeats. The affected members of the American family Ald showed 12 repeats, identical in all three tested patients. Two of the repeats were irregular: the 3rd repeat had a T-to-C substitution at the 9th position (R2c), and the 10th repeat had an A-to-G substitution at the 21st position, identical to that of the Kel family (R3g). Neither substitution changes the encoded amino acid. Four unaffected members of the family had a normal five-repeat pattern on both alleles.

Thirteen Repeats. The proband and his presently healthy niece in the French Che family showed an identical pattern of 13 repeats. The 12th repeat (which is designated R2a) had a G-to-A substitution at position 21. The other niece was normal.

Fig. 2 shows the PCR-amplified 5' fragment of the PRNP coding region in individuals with 5, 10, 12, and 13 repeats.

Phenotype Expression. Individuals with 5 and 9 repeats showed no signs of CJD. The occurrence of 10, 12, and 13 repeats was associated with CJD or a mixed picture of CJD and GSS, for which the clinical and pathological features are summarized in Table 2.

Experimental Transmission. Inoculated brain suspension from the proband of the Kel family transmitted spongiform encephalopathy to a squirrel monkey (43-month incubation period, 3-week illness) and a spider monkey (31-month incubation period, 2-month illness). Brain suspension from the proband of the Ald family transmitted spongiform encephalopathy to a chimpanzee (13-month incubation period, 1-month illness), a capuchin monkey (35-month incubation period, 3-month illness), and a squirrel monkey (24-month incubation period, 2-week illness). Brain suspension from the proband of the Che family transmitted spongiform encephalopathy to a chimpanzee (13-month incubation period, 4-month illness). A stump-tailed monkey inoculated at the same time died of a non-neurological illness after an observation period of 11 years; microscopic examination showed that the brain was normal.

Table 1. Octapeptide coding repeats in the PRNP gene found in normal or non-CJD individuals and in familial CJD or Gerstmann–Sträussler–Scheinker syndrome (GSS) patients

<table>
<thead>
<tr>
<th>Number of repeats</th>
<th>Family</th>
<th>Medical condition</th>
<th>Order of repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>—</td>
<td>Normal</td>
<td>R1, R2, R2, R3, R4</td>
</tr>
<tr>
<td>9</td>
<td>Hay</td>
<td>Cirrhosis</td>
<td>R1, R2, R2, R3, R2, R3, R2, R3, R4</td>
</tr>
<tr>
<td>10</td>
<td>Kel</td>
<td>CJD</td>
<td>R1, R2, R2, R3, R3, R3g, R2, R2, R3, R4</td>
</tr>
<tr>
<td>11</td>
<td>A</td>
<td>CJD</td>
<td>R1, R2, R2, R3, R2, R3, R3g, R2, R2, R3, R4</td>
</tr>
<tr>
<td>12</td>
<td>Ald</td>
<td>CJD</td>
<td>R1, R2, R2c, R3, R3, R2, R3, R2, R2, R3g, R3, R4</td>
</tr>
<tr>
<td>13</td>
<td>Che</td>
<td>CJD/GSS</td>
<td>R1, R2, R2, R3, R2, R2, R2, R2, R2, R2, R2a, R4</td>
</tr>
</tbody>
</table>
The range of susceptible primate species, incubation periods, and durations of illness in these inoculated animals did not differ from those seen after inoculation of tissue from sporadic cases of CJD.

**DISCUSSION**

In contrast to sporadic CJD, in which no disease-specific mutations in the PRNP gene have yet been identified, familial forms of spongiform encephalopathy are being linked to a growing variety of point and insertion mutations in this gene. Most of these mutations appear to influence the clinical expression of disease, particularly with respect to the age at onset and the duration of illness, but also to some extent the clinical presentation and evolution of symptoms. For example, point mutations in codons 102 (18–20), 117 (21–23), and 198 (24) are associated with the distinctive clinical and pathological features of GSS, whereas point mutations in codons 178 and 200 are associated with illnesses that in most respects are typical of sporadic CJD, except for their comparatively early onset and long duration, and in the case of the codon 178 mutation, the absence of periodic electroencephalographic activity (25, 26).

Almost all of the familial patients with insertion mutations that we describe in this report had exceptionally prolonged illnesses (average of 7 years) beginning at a comparatively early age (average of 38 years) and characterized by progressive dementia accompanied by cerebellar and other neurological signs. Affected members in the English kindred with 11 repeats, first reported by Owen et al. (6, 7), and additional patients with 11 and 14 repeats subsequently reported by the same group (27-29) have also experienced prolonged illnesses (5–12 years) beginning at a very early age (22–28 years).

The genetic mechanism for the generation of extra repeats is probably unequal crossover. Fig. 3 illustrates a hypothetical scheme for generating a variable number of repeats by such a mechanism. The similarity of the repeats allows limited frameshifting in the process of duplication; recombination can happen at any site. The result of the first event could be a 7-repeat allele, which has not so far been observed. The result of the second event could be the 9-repeat allele that was observed in the control patient Hay, and further crossover events could result in alleles with 10 or more repetitions. The irregular elements R3g (Kel and Ald families) and R2a (Che family) could have been generated by a recombination inside the R1, R2, and R3 repeats. Thus, all the observed repetitions, including the 11-repeat insert described in the English family (7), can be accounted for by unequal crossover and recombination, except for the R2c repeat in the Ald family, which may represent an independent point mutation.

Table 2. Summary of the clinical, neuropathological, and transmissible features of three families with 10, 12, or 13 octapeptide coding repeats

<table>
<thead>
<tr>
<th>Family</th>
<th>Sex/age at onset</th>
<th>Duration of illness</th>
<th>Clinical evolution</th>
<th>Electroencephalogram</th>
<th>Neuropathology</th>
<th>Transmission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kel</td>
<td>M/31 (proband)</td>
<td>15 yr</td>
<td>Progressive dementia, abnormal behavior, cerebellar signs, tremor, rigidity, myoclonus</td>
<td>Diffuse slowing</td>
<td>Spongiosis, gliosis, neuronal loss</td>
<td>Squirrel and spider monkeys</td>
</tr>
<tr>
<td></td>
<td>M/45</td>
<td>5 yr</td>
<td>Progressive dementia, mood change, disorientation, cerebellar signs, tremor, hyper-reflexia</td>
<td>Not recorded</td>
<td>Spongiosis, gliosis, neuronal loss</td>
<td>Not inoculated</td>
</tr>
<tr>
<td>Ald</td>
<td>F/31 (proband)</td>
<td>11 yr</td>
<td>Mood change, indifference, confusion, aphasia, cerebellar signs, invol. movements, rigidity</td>
<td>Diffuse slowing</td>
<td>Spongiosis, gliosis, neuronal loss</td>
<td>Chimpanzee; squirrel and capuchin monkeys</td>
</tr>
<tr>
<td></td>
<td>F/23</td>
<td>10 yr</td>
<td>Abnormal behavior, cerebellar signs, mutism, rigidity, dementia, myoclonus</td>
<td>Diffuse slowing</td>
<td>Mild gliosis and neuronal loss (no spongiosis)</td>
<td>Not inoculated</td>
</tr>
<tr>
<td>Che</td>
<td>M/35 (proband)</td>
<td>&gt;13 yr (alive)</td>
<td>Clumsiness, dysphasia, cerebellar signs, euphoria, dementia, myoclonus</td>
<td>Slow-wave burst suppression</td>
<td>No biopsy or autopsy (patient still alive)</td>
<td>Not inoculated</td>
</tr>
<tr>
<td></td>
<td>M/55</td>
<td>1 yr</td>
<td>Abnormal behavior, cerebellar signs, mutism, pyramidal signs, hyper-reflexia, myoclonus</td>
<td>Periodic 1- to 2-Hz triphasic complexes</td>
<td>Spongiosis, gliosis, neuronal loss, multicentric (GSS) cerebellar plaques</td>
<td>Chimpanzee</td>
</tr>
<tr>
<td></td>
<td>M/43</td>
<td>5 yr</td>
<td>Abnormal behavior, cerebellar signs, mutism, rigidity, myoclonus</td>
<td>Not recorded</td>
<td>Spongiosis, gliosis, neuronal loss, multicentric (GSS) cerebellar plaques</td>
<td>Not inoculated</td>
</tr>
<tr>
<td></td>
<td>M/48</td>
<td>3 yr</td>
<td>Intellectual slowing, tremor, hypertonus, cerebellar signs, myoclonus</td>
<td>Slow-wave burst suppression</td>
<td>No autopsy</td>
<td>Not inoculated</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intellectual slowing, decreased vision, tremor, cerebellar signs, seizures, myoclonus</td>
<td>Diffuse slowing</td>
<td>No autopsy</td>
<td>Not inoculated</td>
</tr>
</tbody>
</table>
Two normal alleles crossing over with a shift:

\[
\begin{array}{cccc}
R1 & R2 & R2 & R3 \\
R2 & R3 & R4 & R4 \\
\end{array}
\]

The resulting 7-repeat allele crossing over with a normal allele:

\[
\begin{array}{cccc}
R1 & R2 & R2 & R3 \\
R2 & R3 & R3 & R4 \\
\end{array}
\]

The resulting 9-repeat allele observed in a control individual crossing over with a normal allele, but recombination happens inside R3 and R1 repeats:

\[
\begin{array}{cccc}
R1 & R2 & R2 & R3 \\
R2 & R3 & R3 & R3 \\
R3 & R4 & R4 & R4 \\
\end{array}
\]

or, in detail,

\[
\begin{array}{l}
\text{CCC CAT} \quad \text{GTT GGT GGC} \\
\text{CCT CAG GGC GGT GGC} \\
\end{array}\]

resulting in an irregular element

\[
\begin{array}{l}
\text{CCC CAT} \quad \text{GTT GGT GGC} \\
\text{GGA GGG} \\
\text{R1} \\
\end{array}\]

and a 10-repeat allele as observed in patient Ke:

\[
\begin{array}{l}
R1 \quad R2 \quad R2 \quad R3 \quad R3 \quad R3 \quad R3 \quad R3 \quad R4 \\
\end{array}\]

**Fig. 3.** Likely mechanisms of generation of extra repeats.

A number of structural proteins, including collagen (30), silk fibroin (31), keratin (32), and elastin (33, 34), contain large numbers of repeats of from 2 to 6 amino acids, none of which closely resembles the PrP repeat region. However, there are three other proteins containing consecutive repeats of a heptapeptide or octapeptide whose structure and function may give some insight into variations in the CJD amyloid precursor structure. The maize ribosomal protein S18 has a 7-fold heptapeptide repeat at its N terminus that is thought to have some role in RNA binding (35). The DNA-dependent RNA polymerase II contains many repeats of a heptapeptide in its C-terminal domain, where a regular spacing of proline residues appears to be important in maintaining its unusual secondary structure. Although some of the repeats may be functionally redundant (their number varies), it has been suggested that variations in this tandem repeat structure could act to regulate its enzymatic activity (36). Ice-nucleation proteins also contain very large numbers of an octapeptide, and each octapeptide seems to contribute more or less individually to the nucleation process (37).

A substantial body of evidence links alterations in the PrP structure with the occurrence of spongiform encephalopathy in humans and animals, and insertions of variable numbers of amino acids into human PrP adds to existing models of conformational structure and their relation to disease and transmissibility. The repeat region of the precursor molecule may be involved directly in a templating mechanism whereby the normal protein is converted posttranslationally into an enantiomorphic form of unusual conformation (38). The presence of extra repeats might be expected to increase the probability of spontaneous transition to the abnormal enantiomer and thereafter provide an increased concentration of template for the conversion of further molecules.

The cis–trans isomerization of regularly spaced proline residues could play a role in this process, or, if production of the unusual amyloidogenic form of the precursor required
cotranslational ribosomal frameshifting events in the octapeptide region (39), the extra repeats might increase the likelihood of their spontaneous occurrence. Moreover, adjacent 24-base units within the region of the mRNA encoding amyloid octapeptide repeats are able to fold into extraordinary pseudoknot structures (40). Extra repeats might increase the capability of the mRNA to form these pseudoknots.

Whatever the mechanism by which extra repeats exert their effect, the presence of 10 or more such elements is associated with a spectrum of dementing illnesses that, while usually clinically recognizable as CJD, differ from the typical disease picture in their early age at onset and very slow progression (only one patient had an illness lasting less than a year). They may be characterized by the classic pathology of either CJD or GSS, or, exceptionally, by a nearly normal microscopic appearance. Experimentally, they share the same transmissibility characteristics as other familial and sporadic forms of spongiform encephalopathy.


