Murine Model of Niemann-Pick C Disease: Mutation in a Cholesterol Homeostasis Gene

Stacie K. Loftus, Jill A. Morris, Eugene D. Carstea, Jessie Z. Gu, Christiano Cummings, Anthony Brown, Jane Ellison, Kousaku Ohno, Melissa A. Rosenfeld, Danilo A. Tagle, Peter G. Pentchev, William J. Pavan*

An integrated human-mouse positional candidate approach was used to identify the gene responsible for the phenotypes observed in a mouse model of Niemann-Pick type C (NP-C) disease. The predicted murine NPC1 protein has sequence homology to the putative transmembrane domains of the Hedgehog signaling molecule Patched, to the cholesterol-sensing regions of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase and SREBP cleavage-activating protein (SCAP), and to the NPC1 orthologs identified in human, the nematode *Caenorhabditis elegans*, and the yeast *Saccharomyces cerevisiae*. The mouse model may provide an important resource for studying the role of NPC1 in cholesterol homeostasis and neurodegeneration and for assessing the efficacy of new drugs for NP-C disease.

Niemann-Pick type C disease is an autosomal recessive, neurovisceral lipid storage disorder (OMIM number 257220). The most pronounced phenotypic cellular abnormality is an alteration of intracellular cholesterol homeostasis. In NP-C cells, cholesterol is sequestered in lysosomal compartments. In response to low density lipoprotein (LDL), the cells exhibit a delayed down-regulation of both the LDL receptor and of de novo cholesterol synthesis, as well as delayed up-regulation of cellular cholesterol esterification (*1*). These alterations are thought to reflect a defect in the Golgi-mediated efflux of unesterified cholesterol from lysosomes to the endoplasmic reticulum (ER) (*2*).

Two murine models share many of the clinical abnormalities observed in humans with NP-C, that is, increased levels of sphingomyelin and unesterified cholesterol in liver and spleen, and the presence of foamy macrophages, neuronal vacuoles, focal axonal swelling, and decreased Purkinje cell number (*3*). The defects in these two mouse strains, C57BLKS/J *spm* and BALB/c npc^{nih}, arose as spontaneous mutations, were determined allelic by crossbreeding, and have been independently localized to mouse chromosome 18 in a region syntenic to the human *NPC1* locus (*4*, *5*). Experiments with heterokaryon fusions and DNA-mediated complementation confirmed that the two mouse loci belong to the same complementation group as human *NPC1* (*6*) and indicate that the same gene is altered in the two mouse NP-C models (*spm* and npc^{nih}) and the orthologous gene is defective at the human NPC1 locus.

We have combined the candidate gene map from the human *NPC1* critical region with high-resolution linkage mapping and candidate gene analysis using the BALB/c npc^{nih} mouse model to identify the molecular defect responsible for the neurovisceral abnormalities in NP-C disease (Fig. 1). We integrated the murine and human genetic resources by using mouse linkage markers and mouse orthologs of two human expressed sequence tags (ESTs) located within the human *NPC1* critical region (*7*) to assemble a mouse bacterial artificial chromosome (BAC) contig (Fig. 1C). Partial cDNA clones for the two orthologous genes in mouse (GenBank accession numbers AA002656 and MW83C06) were identified in dbEST by BLAST (*8*) analysis with sequences from the respective human orthologous genes 190B6 and *Npc1* (*7*).

The expression patterns of AA002656 and MW83C06 (the latter now termed *Npc1*) (*9*) were examined by Northern blot analysis of RNA isolated from wild-type (WT) and mutant tissues (Fig. 2). Although both genes were expressed in all WT tissues examined, there was a marked reduction of *Npc1* mRNA levels in npc^{nih}/npc^{nih} liver and brain and *spm/spm* liver compared to WT tissues (Fig. 2B). We next compared the *Npc1* cDNA sequences from WT and affected animals to determine if the reduced expression was due to a defect

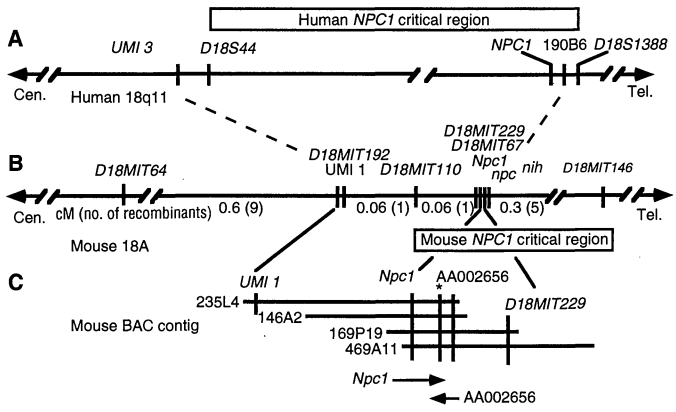

Fig. 1. Integration of murine and human genetic maps. (**A**) Summary of the 1-cM *NPC1* critical region on human chromosome 18q11 containing 190B6 (GenBank accession number hsu55986) and *NPC1* (GenBank accession number AF002020). Murine and human critical regions were integrated with one polymorphic locus (denoted *UMI 1* in mouse and *UMI 3* in human) and the murine orthologs of the two candidate human genes (AA002656 corresponding to 190B6; MW83C06/*Npc1* corresponding to *NPC1*) (*25*). (**B**) A total of 1552 meioses were used to generate a murine 0.36-cM *NPC1* critical region. Genetic distance and number of recombinants are noted between markers typed. (**C**) BAC clones were identified by hybridization of mouse BAC high-density filters (Research Genetics, Huntsville, AL) with *D18MIT229* and AA002656. The relative position and orientations of *Npc1* and AA002656 within the BAC contig were determined by PCR, partial restriction enzyme mapping, and hybridization with mouse EST 5′- and 3′-specific probes. 5′ probe of AA002656: amplified by PCR with primers AA002-ma (5′-GAGCTACACAATGGCAGTAG-3′) and AA002-mr (5′-CAGCTATCACAAACTTGTGC-3′); 5′ probe of *Npc1*: amplified by PCR with primers mp25-JF (5′-CTGTGTCCGAAATCCCACCTGC-3′) and mp25-3ASNB (5′-GTTAAATATCTGCTGCACCAGG-3′); 3′ overlap region between Npc1 and AA002656 (asterisk): PCR with primers AA002-mc (5′-GGACTTCAGAATGTTGTAGGC-3′) and AA002-md (5′-TTGGCTGCCCTGAGGTTCATC-3′). Arrows denote 5′ to 3′ direction of transcription. The chromosome location of AA002656 was independently confirmed by single-strand conformational polymorphism linkage analysis (primers AA002-mc and AA002-md, Alu I–digested).

S. K. Loftus, J. Ellison, W. J. Pavan, Laboratory of Genetic Disease Research, National Human Genome Research Institute, National Institutes of Health (NIH), Bethesda, MD 20892, USA.
J. A. Morris, Laboratory of Gene Transfer, National Human Genome Research Institute, and Developmental and Metabolic Neurology Branch, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, MD 20892, USA.
E. D. Carstea, C. Cummings, A. Brown, P. G. Pentchev, Developmental and Metabolic Neurology Branch, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, MD 20892, USA.
J. Z. Gu, M. A. Rosenfeld, D. A. Tagle, Laboratory of Gene Transfer, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA.
K. Ohno, Department of Neurobiology, School of Life Sciences, Tottori University Faculty of Medicine, Yonago, 683, Japan.

*To whom correspondence should be addressed. E-mail: bpavan@nhgri.nih.gov

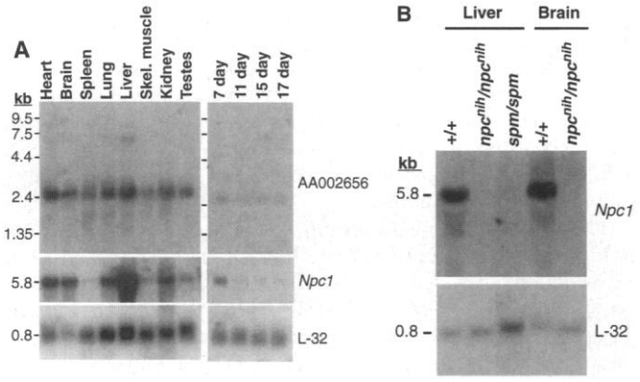

Fig. 2. Northern blot analysis of *Npc1* in (**A**) WT and (**B**) mutant mice. Polyadenlyated RNA (A) from adult tissue (left) and from 7- to 17-day mouse embryos (right). The probe for AA002656 (*26*) detects a 2.5-kb transcript, and the *Npc1* probe (*26*) detects a 5.8-kb transcript in all WT tissues examined; the AA002656 probe detects an additional 5.8-kb transcript in liver. (B) Liver and brain total RNA was isolated from adult (8-week-old) mice. Lanes: +/+, BALB/cStCrlfC3HfNctr $+^{npc}/+^{npc}$; npc^{nih}/npc^{nih}, BALB/cStCrlfC3HfNctr npc^{nih}/npc^{nih}; *spm/spm*, C57BLKS/J *spm/spm*. RNA loading was assessed by hybridization with L-32 (*27*). Molecular sizes are indicated on the left.

in the *Npc1* gene or if it was a secondary event resulting from the NP-C phenotypes. Because both npc^{nih} and *spm* are isogenic mutations (arising and maintained on inbred genetic backgrounds), any genetically linked genomic alteration identified between affected and WT control mice is most likely causative of the disorder. Sequence analysis of the cDNA clones from BALB/c npc^{nih}/npc^{nih} mouse liver and brain RNA revealed that 44 base pairs (bp) of WT sequence were replaced with 24 bp of previously unidentified sequence (*10*). This alteration caused a frameshift, resulting in premature truncation of the putative open reading frame (ORF) (Fig. 3, A and B).

The putative mutation in *Npc1* was subsequently confirmed by isolation of the corresponding genomic region from npc^{nih}/npc^{nih} affected and WT mice (GenBank accession number AF003348) (Fig. 3C). Sequence comparison of the genomic region identified an 824-bp insertion of retrotransposon-like sequences from the mammalian apparent LTR-retrotransposon (MaLR) family (*11*–*13*) (Fig. 3, C and D). The insert does not contain a full-length MaLR, but two distinct regions could be identified. The initial 458 bp had 81% identity to internal sequences of the human endogenous retroviral-like element, HERV-L (*12*), and the terminal 370 bp were identical to the 3′ terminus of a mouse transcript (MT) retrotransposon-like sequence (*13*). Consistent with previous observations that MaLR transposition events are prone to rearrangements at integration sites (*13*), a comparison of WT and npc^{nih}/npc^{nih} mutant *Npc1* intronic sequences revealed that in addition to the inserted sequences, 703 bp of WT sequences were deleted (Fig. 3).

Our results demonstrate that the NP-C phenotypes observed in BALB/c npc^{nih}/npc^{nih} mice result from a mutation of the *Npc1* gene and, together with work by Carstea *et al.* (*7*), they establish that this gene is responsible for NP-C disease in humans. Analysis of the *Npc1* cDNA predicts an ORF of 1278 amino acids that encodes an NH_2-terminal putative signal peptide sequence, a domain that is unique to the *NPC1* orthologs (in mouse, human, the nematode *Caenorhabditis elegans*, and the yeast *Saccharomyces cerevisiae*), and 13 putative transmembrane (TM) domains that include a potential sterol-sensing domain (SSD) (Fig. 4A). The NPC domain consists of 112 amino acids and is marked by eight cysteine residues with conserved spacing between all *NPC1* orthologs analyzed (Fig. 4B). There is extensive sequence similarity between the murine and human *NPC1* in both the ORF and a portion of the 3′ untranslated region. In addition, both murine and human orthologs of AA002656 and *Npc1* are transcribed in opposite orientations, overlap for 284 bp of the 3′ end, and within this region exhibit 86% nucleic acid identity (Fig. 2A).

The most pronounced cellular phenotypes in NP-C disease are an accumulation of unesterified cholesterol in the perinuclear lysosomal compartments and a delayed response in sterol-responsive pathways to exogenous cholesterol caused by disruption of intracellular cholesterol trafficking (*14*). It has been postulated that one pathway of cholesterol trafficking from the lysosome and plasma membrane to the ER is modulated by a Golgi complex–mediated process (*2*) and that disruption of this pathway may be responsible for the NP-C phenotype. Analysis of the putative signal peptide and TM domains of NPC1 (*15*) supports the notion that NPC1 is localized in the Golgi, ER, or the plasma membrane. Additionally, NPC1 contains a COOH-terminal dileucine motif, which has been implicated as a signal for endocytosis from the plasma membrane and lysosomal targeting (*16*).

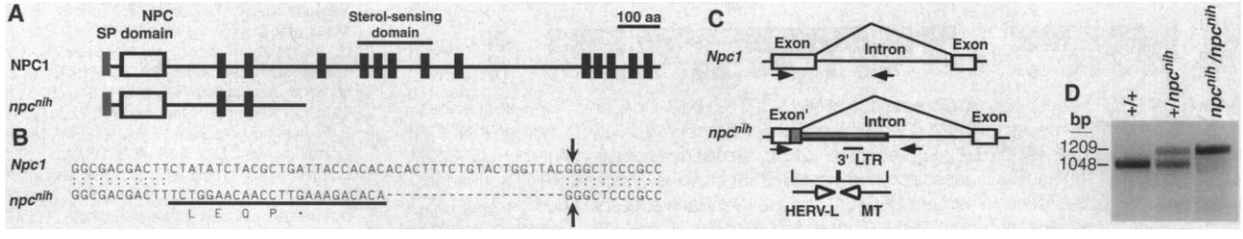

Fig. 3. Comparison of *Npc1* WT and mutant alleles. (**A**) Schematic representation of NPC1 predicted protein in WT and npc^{nih}/npc^{nih} mutant mice (*15*). The npc^{nih} mutation results in protein truncation before the SSD and 11 of the 13 TM domains (vertical black bars). The putative signal peptide sequence is denoted by SP and a gray vertical bar. (**B** and **C**) Alignment of *Npc1* cDNA and genomic sequences from WT ($+^{npc}$) and npc^{nih}/npc^{nih} mutant (npc^{nih}) mice spanning the mutated region. In the mutant cDNA (B), the 44 bp of WT sequence (nucleotides 1634 to 1638) are deleted and replaced by 24 bp from a MaLR (underline) resulting in a stop codon (asterisk) and premature truncation of NPC1. Arrowheads designate splice junctions. In the mutant genomic DNA (C), 703 bp of WT sequence (44 bp of exon and 659 bp of intron) are replaced with 824 bp of MaLR. The orientation of the MaLR sequences (HERV-L and MT) located within the insertion (gray bar) are indicated by the open arrows, and the 3′ long-terminal repeat (LTR) sequence is indicated by the horizontal line. Splicing from the WT exons and from within a newly introduced splice site located within the inserted MaLR sequence (denoted Exon′) is indicated by lines above each allele. Solid arrows denote primers used for genotype analysis (*28*). (**D**) Genotype analysis of npc^{nih} allele demonstrates a size alteration caused by the MaLR insertion or deletion in the *Npc1* locus. Lanes: +/+, BALB/cStCrlfC3HfNctr $+^{nih}/+^{nih}$; $+/npc^{nih}$, BALB/cStCrlfC3HfNctr $+^{nih}/npc^{nih}$; and npc^{nih}/npc^{nih}, BALB/cStCrlfC3HfNctr npc^{nih}/npc^{nih}. PCR products were resolved by 1.2% agarose gel electrophoresis; molecular sizes are indicated on the left.

Together, these observations suggest that NPC1 is associated with one or more membranous structures implicated in intracellular cholesterol homeostasis.

The prediction of an SSD in NPC1 is based on homology to two other proteins with an SSD: 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase and SREBP cleavage-activating protein (SCAP) (Fig. 4A). The SSD in HMG-CoA reductase targets degradation of the enzyme in response to intracellular sterol levels (*17*). Transcription of the gene encoding HMG-CoA reductase is also regulated by the membrane-bound transcription factors SREBP1 and SREBP2, whose cleavage and release from the membrane is in turn regulated by the sterol-responsive protein, SCAP. An $Asp^{443} \rightarrow Asn$ mutation in the SSD of SCAP blocks responsiveness of SCAP to intracelluar cholesterol levels (*18*). This amino acid is conserved in the putative SSD of mouse, human, and *C. elegans* NPC1 (Fig. 4). The putative SSD suggests that NPC1 function may involve direct interactions with sterol moieties. This is consistent with the abnormalities in cellular cholesterol homeostasis observed in individuals with NP-C disease.

There is also extensive amino acid homology between the putative TM and SSD domains of NPC1 and the Hedgehog signaling protein Patched (PTC) (*19*) (Figs. 3A and 4). An SSD has not been previously detected in PTC, nor has PTC been implicated in cholesterol homeostasis. However, several links have been made between PTC signaling, neuronal development, and cholesterol homeostasis (*20, 21*). The secreted signaling molecule, Sonic Hedgehog (SHH), contains a covalently attached cholesterol moiety (*21*) and has been shown to biochemically interact with PTC (*22*). In addition, mutations in SHH (*23*) or mutations that produce cholesterol deficiencies (*24*) result in abnormal development and function of the central nervous system including holoprosencephaly (failure of the forebrain to divide into hemispheres). The structural similarities between NPC1 and PTC raise the possibility that NPC1 could also interact with protein-sterol complexes (*21*) that are required for normal neuronal development. Alternatively, alterations in cellular cholesterol homeostasis in utero could indirectly alter the function of proteins, such as SHH, that require a cholesterol adduct for normal neuronal development or function. While neurodevelopmental anomalies such as holoprosencephaly have not been observed in npc^{nih}/npc^{nih} mice, there is a significant deviation from the expected ratio of npc^{nih}/npc^{nih} mice obtained from intercrosses (*4*). Biochemical and genetic analyses of the NPC1 protein with mice, *C. elegans*, and *S. cerevisiae* may allow a better understanding of NPC1 function and assessment of pharmacological interventions.

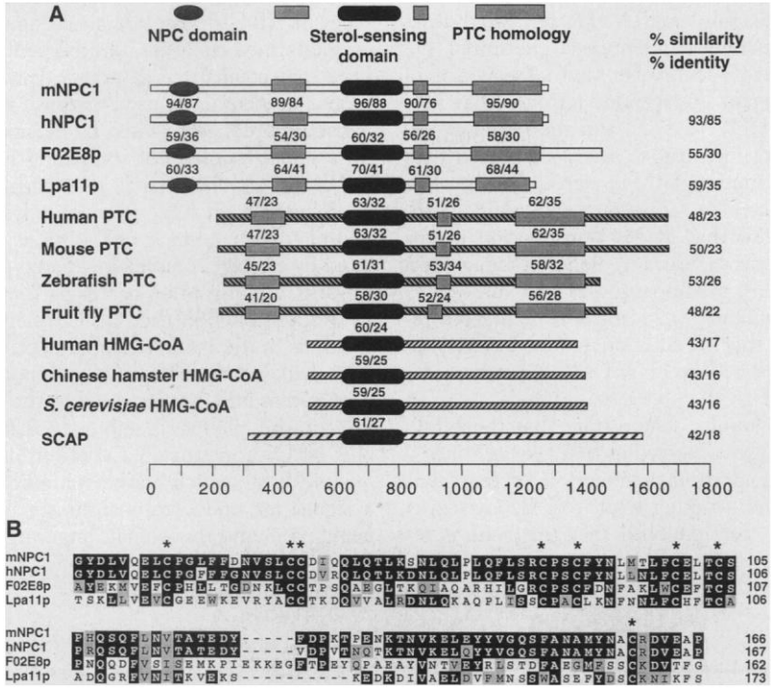

Fig. 4. Analysis of NPC1 sequences. (**A**) Comparison of NPC1, PTC, HMG-CoA, and SCAP sequences. NPC1 sequences are from mouse (mNPC1; GenBank accession number AF003348), human (hNPC1; GenBank accession number AF002020), *C. elegans* (F02E8p; GenBank accession number U53340), and *S. cerevisiae* (Lpa11p; GenBank accession number U33335) proteins. Patched (PTC) sequences are from human (GenBank accession number U59464), mouse (GenBank accession number U46155), zebrafish (GenBank accession number X98883), and fruit fly PTC (GenBank accession number M28999) proteins. HMG-CoA sequences are from human (GenBank accession number M11058), Chinese hamster (GenBank accession number L00183), and *S. cerevisiae* (GenBank accession number M22002) proteins. The SCAP sequence is from the Chinese hamster (SCAP; GenBank accession number U67060) protein. Blocks of homologies were identified with the program MACAW (*29*). Pairwise comparisons in relation to mNPC1 by use of entire protein sequences (values shown at right) or within NPC, SSD, or PTC domains (values shown above each domain) were performed with GCG software package. (**B**) Sequence comparisons within the NPC domain. Residues that are identical in at least two of the four sequences are highlighted in black. Similar residues are shaded in gray. The NPC domain represents the most conserved region outside of the TM domains and SSDs. Conserved cysteines are indicated by asterisks. Abbreviations for the amino acids are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

REFERENCES AND NOTES

1. P. G. Pentchev, M. T. Vanier, K. Suzuki, M. C. Patterson, in *The Metabolic and Molecular Basis of Inherited Diseases*, C. R. Scriver, A. L. Beaudet, W. S. Sly, D. Valle, Eds. (McGraw-Hill, Philadelphia, PA, 1995), pp. 2625–2639.
2. E. B. Neufeld *et al.*, *J. Biol. Chem.* **271**, 21604 (1996); E. J. Blanchette-Mackie *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8022 (1988); R. A. Coxey, P. G. Pentchev, G. Campbell, E. J. Blanchette-Mackie, *J. Lipid. Res.* **34**, 1165 (1993).
3. Y. Higashi, P. G. Pentchev, S. Murayama, K. Suzuki, in *Neurology in Brain Research*, F. Ikuta, Ed. (Elsevier, Amsterdam, 1991), pp. 85–102.
4. R. P. Erickson *et al.*, *Mamm. Genome* **8**, 355 (1997).
5. C. Morris, H. Bhuvaneswaran, H. Shio, S. Fowler, *Am. J. Pathol.* **140**, 140 (1982); S. Miyawaki, S. Mitsuoka, T. Sakiyama, T. Kitagawa, *J. Hered.* **73**, 257 (1982); Y. Sakai, S. Miyawaki, A. Shimizu, S. Ohno, T. Watanabe, *Genetics* **29**, 103 (1991); T. Yamamoto, K. Iwasawa, T. Tokoro, Y. Eto, K. Maekawa, *No To Hattatsu* **26**, 318 (1994).
6. S. Akaboshi, T. Yano, S. Miyawaki, K. Ohno, K. Takeshita, *Hum. Genet.* **99**, 350 (1997); J. Z. Gu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7378 (1997).
7. E. D. Carstea *et al.*, *Science* **277**, 228 (1997).
8. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **215**, 403 (1990).
9. Extension of the mouse *Npc1* cDNA sequence (GenBank accession number AF003348) was obtained by three sequential rapid amplification of cDNA ends (RACE) reactions from Marathon-Ready brain and liver libraries (Clontech). Section 1 was obtained by 3′ RACE polymerase chain reaction (PCR) amplification with primer mp25-1F (5′-CGTTCTGTCATTCAGTGTTGCGG-3′) with sequence from clone MW83C06. Sections 2 and 3 were obtained in sequential 5′ RACE reactions: section 2, mp25-2R (5′-CAGTCAAAGTAGTCATCGATCCAG-3′) followed by nested primer mp25-4R (5′-CTCAGCTGCGTTAAATATCTGC-3′); section 3, mp25-ZR (5′-CAACACAAGCCACGGGAACACC-3′) followed by nested primer mp25-LR (5′-GACCAGAGCTCTACAGGATTG-3′). The complete sequence of murine clone AA002656 has been submitted to GenBank (accession number AF003352).

10. *Npc1* cDNA from BALB/cStCrlfC3HfNctr $+^{npc}/+^{npc}$, BALB/cStCrlfC3HfNctr npc^{nih}/npc^{nih}, and C57BLKS/J *spm/spm* strains was isolated with primers mnpc1-3′ARev (5′-GAGAACAAGCTCTA-ATGAG-3′) and mp25-2R for reverse transcription of liver RNA and primer pairs mp25-1F/mnpc1-3′BREV (5′-GCCTACAACATTCTGAAGTCC-3′) and mp25-JF/mp25-3ASNB for PCR amplification in two segments. Three clones were analyzed for each strain by complete coverage sequencing.
11. Genomic intronic sequences of BALB/cStCrlf-C3HfNctr $+^{npc\text{-}nih}/+^{npc\text{-}nih}$ (GenBank accession number AF003349), BALB/cStCrlfC3HfNctr npc^{nih}/npc^{nih} (GenBank accession number AF003350), MaLR insertion event (GenBank accession number AF003351).
12. A. Cordonnier, J. F. Casella, T. Heidmann, *J. Virol.* **69**, 5890 (1995).
13. U. A. O. Heilein, R. Lange-Stablitzky, H. Schaal, W. Wille, *Nucleic Acids Res.* **14**, 6403 (1986); R. G. Kelly, *Genomics* **24**, 509 (1994); F. A. Smit, *Nucleic Acids Res.* **21**, 1863 (1993).
14. P. G. Pentchev, E. J. Blanchette-Mackie, E. A. Dawidowicz, *Trends Cell Biol.* **4**, 365 (1994).
15. The putative ORF begins at the first methionine of the longest ORF in the cDNA sequence. PSORT was used to predict signal peptide sequence {D. J. McGeogh, *Virus Res.* **3**, 271 (1985), modified by K. Nakai and M. Kanehisa [*Proteins Struct. Funct. Genet.* **11**, 95 (1991)]}. TM domains were predicted with PSORT [P. Klein, M. Kanehisha C. DeLisi, *Biochim. Biophys. Acta* **815**, 468 (1985)] and TMbase [K. Hoffman and W. Stoffel, *Biol. Chem. Hoppe-Seyler* **374**, 166 (1993)]. Putative TM domains 7, 10, and 11 predicted in the human NPC1 protein were not predicted by means of identical analyses with mouse NPC1. One of these TM domains is in the SSD region (*8*).
16. F. Letourneur and R. D. Klausner, *Cell* **69**, 1143 (1992); S. Ogata and M. Fukuda, *J. Biol. Chem.* **269**, 5210 (1994); W. Hunziker and H. J. Geuze, *BioEssays* **18**, 379 (1996).
17. D. G. Skalnik, H. Narita, C. Kent, R. D. Simoni, *J. Biol. Chem.* **263**, 6836 (1988); G. Gil, J. R. Faust, D. J. Chin, J. L. Goldstein, M. S. Brown, *Cell* **41**, 249 (1985).
18. X. Hua, A. Nohturfft, J. L. Goldstein, M. S. Brown, *Cell* **87**, 415 (1996).
19. R. L. Johnson *et al.*, *Science* **272**, 1668 (1996); L. V. Goodrich, R. L. Johnson, L. Milenkovic, J. A. McMahon, M. P. Scott, *Genes Dev.* **10**, 301 (1996).
20. J. Herz, T. E. Willnow, R. Farese Jr., *Nature Genet.* **15**, 123 (1997).
21. J. A. Porter, K. E. Young, P. A. Beachy, *Science* **274**, 255 (1996); J. A. Porter *et al.*, *Cell* **86**, 21 (1996); J. A. Porter *et al.*, *Nature* **374**, 363 (1995).
22. D. M. Stone *et al.*, *Nature* **384**, 129 (1996).
23. C. Chiang *et al.*, *ibid.* **383**, 407 (1996).
24. G. E. Homanics *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 2389 (1993); T. E. Willnow *et al.*, *ibid.* **93**, 8460 (1996); G. Salen *et al.*, *J. Lipid Res.* **37**, 1169 (1996).
25. Polymorphic carriers were generated by crossing BALB/cStCrlfC3HfNctr $npc^{nih}/+^{nih}$ animals to inbred (C57BL/6J) and WT-derived (CAST/Ei and MOLF/Ei) strains (Jackson Laboratory). F_1-carriers were identified by progeny testing offspring [P. G. Pentchev, R. O. Brady, A. E. Gal, S. R. Hibbert, *Biochim. Biophys. Acta* **488**, 312 (1977)]. An initial 736 meioses were typed [W. J. Pavan, S. Mac, M. Cheng, S. M. Tilghman, *Genome Res.* **5**, 29 (1995)] with markers *Umi 1* [A. J. Griffith *et al.*, *Mamm. Genome* **7**, 417 (1996)] *D18MIT64*, *D18MIT110*, and *D18MIT146* [http://www-genome.wi.mit.edu] to establish flanking markers *D18MIT64* and *D18MIT146*. For the remaining 816 meioses, only animals recombinant between flanking loci were genotyped with remaining loci and test-crossed to determine carrier status. Recombinant mice were from F_1-intercrosses, backcrosses to $npc^{nih}/+^{nih}$ carrier mice, and crosses with nonrecombinant F_2-carriers.
26. *Npc1* expression was detected by hybridization of a 1.7-kb PCR amplification product generated with primers 5′JF (5′-CTGTGTCCGAAATCCCACCTGC-3′) and 3′ASNB (5′-GTTAAATATCTGCTGCAC-CAGG-3′), representing nucleotides 1161 to 2901. AA002656 expression was detected with the excised 1-kb clone insert.
27. K. P. Dudov and R. P. Perry, *Cell* **37**, 457 (1984).
28. PCR primers mp25-8F (GGTGCTGGACAGCCAAG-TA) and mp25-INTR3 (5′-GATGGTCTGTTCTCCC-ATG-3′) flanking the insertion event were used to amplify the WT and mutant allele in Fig. 3D.
29. G. D. Schuler, S. F. Altschul, D. J. Lipman, *Proteins* **9**, 180 (1991).

30. We thank E. Neufeld, H. Varmus, D. Lipman, A. Wynshaw-Boris, L. Biesecker, R. Nussbaum, M. Meisler, and J. Trent for discussions and D. Leja for graphics. Supported by the Ara Parseghian Medical Research Foundation. Animal care was in accordance with NIH guidelines. This paper is dedicated to the memory of Michael Parseghian.

8 May 1997; accepted 5 June 1997

Abnormal Lignin in a Loblolly Pine Mutant

John Ralph,* John J. MacKay, Ronald D. Hatfield, David M. O'Malley, Ross W. Whetten, Ronald R. Sederoff

Novel lignin is formed in a mutant loblolly pine (*Pinus taeda* L.) severely depleted in cinnamyl alcohol dehydrogenase (E.C. 1.1.1.195), which converts coniferaldehyde to coniferyl alcohol, the primary lignin precursor in pines. Dihydroconiferyl alcohol, a monomer not normally associated with the lignin biosynthetic pathway, is the major component of the mutant's lignin, accounting for ~30 percent (versus ~3 percent in normal pine) of the units. The level of aldehydes, including new 2-methoxybenzaldehydes, is also increased. The mutant pines grew normally indicating that, even within a species, extensive variations in lignin composition need not disrupt the essential functions of lignin.

Lignins are complex phenolic plant polymers essential for mechanical support, defense, and water transport in vascular terrestrial plants (*1*, *2*). They are usually derived from three hydroxycinnamyl alcohol precursors **2a** through **c** in varying proportions (Fig. 1). In gymnosperms—for example, pine and other conifers—lignin is polymerized from only two of the three monomers, *p*-coumaryl alcohol **2a** and coniferyl alcohol **2b**, with coniferyl alcohol being predominant (~90%). *p*-Coumaryl alcohol–derived subunit levels are increased in compression wood, which forms during mechanical or gravitational stress and in wood knots (*3*). In woody angiosperms, lignin is derived from coniferyl alcohol **2b** and sinapyl alcohol **2c** in roughly equal proportions. Precursors and derivatives of hydroxycinnamyl alcohols also contribute to the lignin structure. For example, acetylated monolignols (hydroxycinnamyl acetates) have been implicated in kenaf (*Hibiscus cannabinus*) (*4*) and woody angiosperms (*5*), and *p*-coumarate esters are found in all grass lignins, implicating hydroxycinnamyl *p*-coumarates as precursors (*6*, *7*). Low levels (~5%) of cinnamaldehydes and benzaldehydes are present in all isolated lignins and are responsible for the bright crimson staining of lignified tissues by phloroglucinol-HCl (*8*).

Removal of lignin from wood and plant fibers is the basis of chemical pulping to produce diverse pulp and paper products. Genetic engineering of the lignin biosynthetic pathway to lower lignin concentration or construct lignins more amenable to extraction is an active area of current research (*9*). However, several mutations have been identified and characterized that affect the lignin biosynthetic pathway (*10*). In maize (*Zea mays*) and related grasses, mutants characterized by a brown midrib (*bm* or *bmr*) have modified lignin (*11*). The *bm* phenotype can result from changes affecting cinnamyl alcohol dehydrogenase (CAD) (for example, *bm1* of maize) (*12*, *13*), O-methyl transferase (OMT) (for example, *bm3* of maize) (*12*, *13*), or both CAD and OMT (for example, *bmr6* of sorghum, *Sorghum bicolor*) (*14*). Mutations in two other maize genes also result in brown midrib phenotypes, but the products of these genes remain unknown. A mutation in the gene encoding ferulate-5-hydroxylase has been identified in *Arabidopsis thaliana*, but it does not result in a brown midrib phenotype (*15*). No lignin mutants have been previously identified in woody plants.

CAD catalyzes the last step of the lignin precursor biosynthetic pathway (Fig. 1), reduction of hydroxycinnamaldehydes **1** to hydroxycinnamyl alcohols **2** (the conventional lignin monomers or monolignols) (*16*). A reduction in CAD activity might

J. Ralph, U.S. Dairy Forage Research Center, U.S. Department of Agriculture (USDA)–Agricultural Research Service (ARS), Madison, WI 53706–1108, and Department of Forestry, University of Wisconsin, Madison, WI 53706–1598, USA.
J. J. MacKay, Department of Genetics, North Carolina State University (NCSU), Raleigh, NC 27695–7614, USA.
R. D. Hatfield, U.S. Dairy Forage Research Center, USDA-ARS, Madison, WI 53706–1108, USA.
D. M. O'Malley and R. W. Whetten, Department of Forestry, NCSU, Raleigh, NC 27695–8008, USA.
R. R. Sederoff, Department of Genetics, NCSU, Raleigh, NC 27695–7614 and Department of Forestry, NCSU, Raleigh, NC 27695–8008, USA.

*To whom correspondence should be addressed at U.S. Dairy Forage Research Center, USDA-ARS, 1925 Linden Drive West, Madison, WI 53706–1108, USA. E-mail: jralph@facstaff.wisc.edu