

[0014] Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

[0015] Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X, and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1A. The nucleotide sequence of each deposited plasmid can readily be determined by sequencing the deposited plasmid in accordance with known methods.

[0016] The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular plasmid can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

[0017] Also provided in Table 1A is the name of the vector which contains the cDNA plasmid. Each vector is routinely used in the art. The following additional information is provided for convenience.

[0018] Vectors Lambda Zap (U.S. Pat. Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Pat. Nos. 5,128,256 and 5,286,636), Zap Express (U.S. Pat. Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., *Nucleic Acids Res.* 16:7583-7600 (1988); Altung-Mees, M. A. and Short, J. M., *Nucleic Acids Res.* 17:9494 (1989)) and pBK (Altung-Mees, M. A. et al., *Strategies* 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, Calif., 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Phagemid pBS may be excised from the Lambda Zap and Uni-Zap XR vectors, and phagemid pBK may be excised from the Zap Express vector. Both phagemids may be transformed into *E. coli* strain XL-1 Blue, also available from Stratagene Vectors pSport1, pCMV Sport 1.0, pCMV Sport 2.0 and pCMV Sport 3.0, were obtained from Life Technologies, Inc., P.O. Box 6009, Gaithersburg, Md. 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into *E. coli* strain DH10B, also available from Life Technologies. See, for instance, Gruber, C. E., et al., *Focus* 15:59 (1993). Vector lafmid BA (Bento Soares, Columbia University, New York, N.Y.) contains an ampicillin resistance gene and can be transformed into *E. coli* strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, Calif. 92008, contains an ampicillin resistance gene and may be transformed into *E. coli* strain DH10B, available from Life Technologies. See, for instance, Clark, J. M., *Nuc. Acids Res.* 16:9677-9686 (1988) and Mead, D. et al., *Bio/Technology* 9: (1991).

[0019] The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, and/or a deposited cDNA (cDNA Clone ID). The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include, but are not limited to, preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

[0020] Also provided in the present invention are allelic variants, orthologs, and/or species homologs. Procedures known in the art can be used to obtain full-length genes, allelic variants, splice variants, full-length coding portions, orthologs, and/or species homologs of genes corresponding to SEQ ID NO:X and SEQ ID NO:Y using information from the sequences disclosed herein or the clones deposited with the ATCC. For example, allelic variants and/or species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for allelic variants and/or the desired homologue.

[0021] The present invention provides a polynucleotide comprising, or alternatively consisting of, the nucleic acid sequence of SEQ ID NO:X and/or a cDNA contained in ATCC Deposit No. Z. The present invention also provides a polypeptide comprising, or alternatively, consisting of, the polypeptide sequence of SEQ ID NO:Y, a polypeptide encoded by SEQ ID NO:X, and/or a polypeptide encoded by a cDNA contained in ATCC deposit No. Z. Polynucleotides encoding a polypeptide comprising, or alternatively consisting of, the polypeptide sequence of SEQ ID NO:Y, a polypeptide encoded by SEQ ID NO:X and/or a polypeptide encoded by the cDNA contained in ATCC Deposit No. Z, are also encompassed by the invention. The present invention further encompasses a polynucleotide comprising, or alternatively consisting of, the complement of the nucleic acid sequence of SEQ ID NO:X, and/or the complement of the coding strand of the cDNA contained in ATCC Deposit No. Z.

[0022] Description of Table 1B (Comprised of Tables 1B.1 and 1B.2)

[0023] Table 1B.1 and Table 1B.2 summarize some of the polynucleotides encompassed by the invention (including cDNA clones related to the sequences (Clone ID:), contig sequences (contig identifier (Contig ID:)) and contig nucleotide sequence identifiers (SEQ ID NO:X)) and further summarizes certain characteristics of these polynucleotides and the polypeptides encoded thereby. The first column of Tables 1B.1 and 1B.2 provide the gene numbers in the application for each clone identifier. The second column of Tables 1B.1 and 1B.2 provide unique clone identifiers, "Clone ID:", for cDNA clones related to each contig sequence disclosed in Table 1A and/or Table 1B. The third column of Tables 1B.1 and 1B.2 provide unique contig identifiers, "Contig ID:" for each of the contig sequences disclosed in these tables. The fourth column of Tables 1B.1 and 1B.2 provide the sequence identifiers, "SEQ ID NO:X", for each of the contig sequences disclosed in Table 1A and/or 1B.

[0024] Table 1B.1

[0025] The fifth column of Table 1B.1, "ORF (From-To)", provides the location (i.e., nucleotide position numbers)