

nucleic acid structure. A nucleic acid molecule is said to be the “complement” of another nucleic acid molecule if they exhibit complete complementarity. As used herein, molecules are said to exhibit “complete complementarity” when every nucleotide of one of the molecules is complementary to a nucleotide of the other. Two molecules are said to be “minimally complementary” if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under at least conventional “low-stringency” conditions. Similarly, the molecules are said to be “complementary” if they can hybridize to one another with sufficient stability to permit them to remain annealed to one another under conventional “high-stringency” conditions. Conventional stringency conditions are described by Sambrook, et al., In: *Molecular Cloning, A Laboratory Manual, 2nd Edition*, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1989), and by Haymes, et al. In: *Nucleic Acid Hybridization, A Practical Approach*, IRL Press, Washington, D.C. (1985). Departures from complete complementarity are therefore permissible, as long as such departures do not completely preclude the capacity of the molecules to form a double-stranded structure. Thus, in order for an EST nucleic acid molecule or fragment EST nucleic acid molecule to serve as a primer or probe it need only be sufficiently complementary in sequence to be able to form a stable double-stranded structure under the particular solvent and salt concentrations employed.

**[0040]** Appropriate stringency conditions which promote DNA hybridization are, for example, 6.0x sodium chloride/sodium citrate (SSC) at about 45° C., followed by a wash of 2.0xSSC at 50° C., are known to those skilled in the art or can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0xSSC at 50° C. to a high stringency of about 0.2xSSC at 50° C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22° C., to high stringency conditions at about 65° C. Both temperature and salt may be varied, or either the temperature or the salt concentration may be held constant while the other variable is changed.

**[0041]** In a preferred embodiment, a nucleic acid of the present invention will specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NO: 1 through SEQ ID NO: 9112 or complements thereof under moderately stringent conditions, for example at about 2.0xSSC and about 65° C.

**[0042]** In a particularly preferred embodiment, a nucleic acid of the present invention will include those nucleic acid molecules that specifically hybridize to one or more of the nucleic acid molecules set forth in SEQ ID NO:1 through SEQ ID NO: 9112 or complements thereof under high stringency conditions.

**[0043]** In one aspect of the present invention, the nucleic acid molecules of the present invention have one or more of the nucleotide sequences set forth in SEQ ID NO: 1 through to SEQ ID NO:9112 or complements thereof. In another aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 90% sequence identity with one or more of the nucleotide sequences set forth in SEQ ID NO: 1 through to SEQ ID NO:9112 or complements thereof. In a further aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 95%

sequence identity with one or more of the nucleotide sequences set forth in SEQ ID NO: 1 through to SEQ ID NO:9112 or complements thereof. In a more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 98% sequence identity with one or more of the nucleotide sequences set forth in SEQ ID NO: 1 through to SEQ ID NO:9112 or complements thereof. In an even more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention share between 100% and 99% sequence identity with one or more of the sequences set forth in SEQ ID NO: 1 through to SEQ ID NO:9112 or complements thereof. In a further, even more preferred aspect of the present invention, one or more of the nucleic acid molecules of the present invention exhibit 100% sequence identity with one or more nucleic acid molecules present within the cDNA libraries designated LIB 149, LIB 150, LIB3026, and LIB3373 (Monsanto Company, St. Louis, Mo., United States of America).

**[0044]** “Percentage of sequence identity” is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the nucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

**[0045]** Polypeptides which are “substantially similar” share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. “Conservative amino acid substitutions” refer to substitutions of one or more amino acids in a native amino acid sequence with another amino acid(s) having similar side chains, resulting in a silent change. Conserved substitutes for an amino acid within a native amino acid sequence can be selected from other members of the group to which the naturally occurring amino acid belongs. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acid substitution groups are: valine-leucine, valine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, aspartic acid-glutamic acid, and asparagine-glutamine.

**[0046]** Optimal alignment of sequences for comparison can use any means to analyze sequence identity (homology) known in the art, e.g., by the progressive alignment method of termed “PILEUP” (Morrison, Mol. Biol. Evol. 14:428-441 (1997), as an example of the use of PILEUP); by the local homology algorithm of Smith & Waterman (Adv. Appl. Math. 2: 482 (1981)); by the homology alignment algorithm of