

group consisting of SEQ ID NO: 1 through SEQ ID NO: 9395 or complements thereof.

[0009] The present invention further provides a substantially purified protein, peptide, or fragment thereof encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 9395 or complements thereof.

[0010] The present invention also provides a substantially purified nucleic acid molecule encoding an *Chlorella sarokiniana* protein homologue or fragment thereof, wherein the nucleic acid molecule comprises a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 9395.

[0011] The present invention also provides a transformed cell having a nucleic acid molecule which comprises: (A) an exogenous promoter region which functions in the cell to cause the production of a mRNA molecule; which is operably linked to (B) a structural nucleic acid molecule, wherein the structural nucleic acid molecule comprises a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 9395 or complements thereof; which is operably linked to (C) a 3' non-translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule.

[0012] The present invention also provides a plant cell, a mammalian cell, a bacterial cell, an insect cell, a fungal cell and an algal cell transformed with a nucleic acid molecule of the present invention.

[0013] The present invention also provides a computer readable medium having recorded thereon one or more of the nucleotide sequences depicted in SEQ ID NO: 1 through SEQ ID NO: 9395 or complements thereof.

DETAILED DESCRIPTION OF THE INVENTION

Agents of the Present Invention:

[0014] (a) Nucleic Acid Molecules

[0015] Agents of the present invention include substantially purified (or isolated) nucleic acid molecules and more specifically EST nucleic acid molecules or nucleic acid fragment molecules thereof. Fragment EST nucleic acid molecules may encode significant portion(s) of, or indeed most of, the EST nucleic acid molecule. Alternatively, the fragments may comprise smaller oligonucleotides (having from about 15 to about 250 nucleotide residues, and more preferably, about 15 to about 30 nucleotide residues).

[0016] In a preferred embodiment the nucleic acid molecules of the present invention are derived from a unicellular green alga and in an even more preferred embodiment the nucleic acid molecules of the present invention are derived from unicellular green algae belonging to the genus *Chlorella*. In a particularly preferred embodiment the nucleic acid molecules of the present invention are derived from *Chlorella sarokiniana*.

[0017] The term "nucleic acid molecule" or "nucleic acid" refers to a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. Nucleic acid molecules may also optionally contain

synthetic, non-natural or altered nucleotide bases that permit correct read through by a polymerase and do not alter expression of a polypeptide encoded by that nucleic acid molecule.

[0018] As used herein "a substantially purified nucleic acid" or "an isolated nucleic acid" refers to a nucleic acid that is no longer accompanied by some of materials with which it is associated in its natural state or to a nucleic acid the structure of which is not identical to that of any of naturally occurring nucleic acid. Examples of a substantially purified nucleic acid include: (1) DNAs which have the sequence of part of a naturally occurring genomic DNA molecules but are not flanked by two coding sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs; (2) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (3) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; (4) recombinant DNAs; and (5) synthetic DNAs. A substantially purified nucleic acid may also be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

[0019] It is also contemplated by the inventors that the substantially purified (or isolated) nucleic acids of the present invention also include known types of modifications, for example, labels which are known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog. Other known modifications include internucleotide modifications, for example, those with uncharged linkages (methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.) and with charged linkages (phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, proteins (including nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), those with intercalators (acridine, psoralen, etc.), those containing chelators (metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, and those with modified linkages.

[0020] It is understood that the agents of the present invention may be labeled with reagents that facilitate detection of the agent (e.g. fluorescent labels (Prober, et al., *Science* 238:336-340 (1987), Albarella et al., EP 144914), chemical labels (Sheldon et al., U.S. Pat. No. 4,582,789; Albarella et al., U.S. Pat. No. 4,563,417), modified bases (Miyoshi et al., EP 119448).

[0021] The term "nucleotide sequence" or "nucleic acid sequence" refers to both the sense and antisense strands of a nucleic acid as either individual single strands or in the duplex. It includes, but is not limited to, self-replicating plasmids, chromosomal sequences, and infectious polymers of DNA or RNA.

[0022] A "coding sequence", "structural nucleotide sequence" or "structural nucleic acid molecule" is a nucleotide sequence which is translated into a polypeptide, usually via mRNA, when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the 5'-terminus and a translation stop codon at the 3'-terminus. A coding sequence can include, but is not limited to, genomic DNA, cDNA, and recombinant nucleotide sequences.