(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 January 2002 (03.01.2002)

PCT

(10) International Publication Number
WO 02/00861 A2

(51) International Patent Classification[7]: C12N 9/00

(21) International Application Number: PCT/US01/20371

(22) International Filing Date: 26 June 2001 (26.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/214,174 26 June 2000 (26.06.2000) US

(71) Applicant *(for all designated States except US)*: **MILLENNIUM PHARMACEUTICALS, INC.** [US/US]; 75 Sydney Street, Cambridge, MA 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MEYERS, Rachel, A.** [US/US]; 115 Devonshire Road, Newton, MA 02468 (US). **WILLIAMSON, Mark** [US/US]; 15 Stonecrest Drive, Saugus, MA 01906 (US).

(74) Agents: **KRON, Eric, J.** et al.; Alston & Bird LLP, Bank of America Plaza, Suite 4000, 101 South Tryon Street, Charlotte, NC 28280-4000 (US).

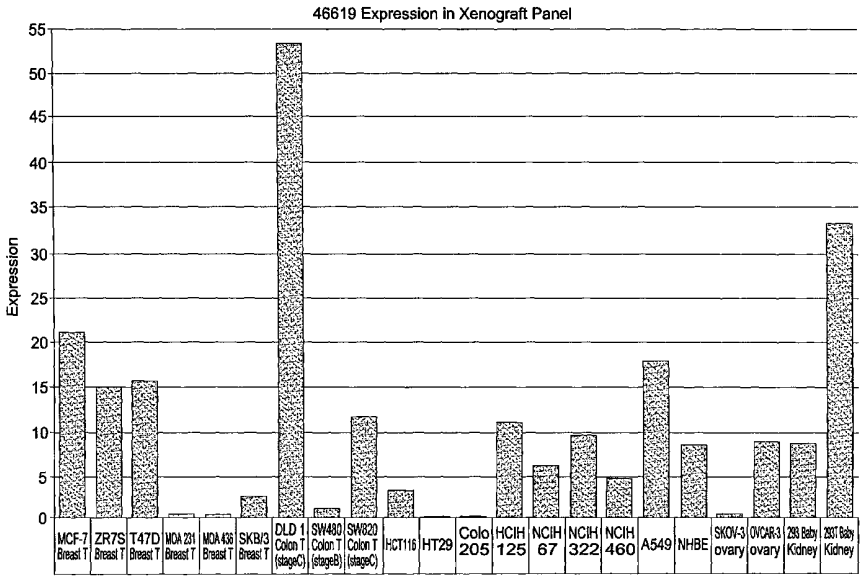
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, DZ, EC, EE, EE (utility model), ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

[Continued on next page]

(54) **Title:** 46619, A NOVEL HUMAN BETA-KETOACYL SYNTHASE AND USES THEREOF



(57) **Abstract:** The invention provides isolated and novel nucleic acids molecules, designated beta-ketoacyl synthase nucleic acid molecules, which encode novel beta-ketoacyl synthase polypeptides. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing beta-ketoacyl synthase nucleic acid molecules, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a beta-ketoacyl synthase gene has been introduced or disrupted. The invention still further provides isolated beta-ketoacyl synthase proteins, fusion proteins, antigenic peptides and anti-beta-ketoacyl synthase antibodies. Diagnostic methods utilizing compositions of the invention are also provided.

WO 02/00861 A2

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

46619, A NOVEL HUMAN BETA-KETOACYL SYNTHASE AND USES THEREOF

BACKGROUND OF THE INVENTION

Fatty acid biosynthesis in animals occurs in a multienzyme complex known as fatty acid synthase. The complex catalyzes the synthesis of the long-chain fatty acid palmitate from acetyl-CoA, malonyl-CoA, and NADPH (Wakil (1989) *Biochemistry 28*: 4523-4530). In addition to enzymes, the complex includes acyl carrier protein (ACP). The use of the multienzyme complex, together with the coenzyme attachment of intermediates greatly increases the efficiency of fatty acid synthesis. (Zubay (1988) "*Biochemistry*" 2d ed., Macmillan Publishing Company). There are seven reactions which are catalyzed by the fatty acid synthase complex. Beta-ketoacyl synthase is one of the enzymes involved in a condensation type reaction in fatty acid biosynthesis.

Animal fatty acid synthase (FAS) consists of two identical polypeptides each carrying six enzymes and an acyl carrier protein that are juxtaposed to form two centers for the synthesis of palmitic acid from acetyl- and malonyl-CoA. The six catalytic domains are clusterd in two regions separated by approximately 600 residues. The beta-ketoacyl synthase, malonyl/acetyl transferase, and dehydrase domains are located within the amino-terminal half of the polypeptide, whereas the enoyl reductase, beta-ketoacyl reductase and thioesterase are located in the carboxy-terminal half: the ACP domain is located between the beta-ketoacyl reductase and thioesterase domains (Rangan *et al.* (1998) *J. Biol. Chem.* 273(*52*): 34949-34953).

The order of the various FAS activities along the polypeptide have been established: beta-ketoacyl synthase; acetyl-CoA and malony-CoA transacylases; beta-hydroxyacyl dehydratase; enoyl reductase; beta-ketoacyl reductase; the site for the prosthetic group; 4'-phosphopantetheine (ACP); and the thioreductase (Jayakumar et al. (1996) *Proc. Natl. Acad. Sci. USA 93*: 14509-14514).

Human FAS has been used successfully as a prognostic factor in identifying patients with breast cancer who have a markedly worse prognosis (Kuhajda *et al.*

(1994) *Proc. Natl. Acad. Sci. 91*: 6379-6383). Also, breast tumors marked with high levels of FAS are four times more likely to recur metastasize than those not so marked. Also, investigators have found an association between high levels of FAS expression and a worsened prognosis in patients with adenocarcinoma of the prostrate (Shurbaji *et al.* (1992) *Am. J. Clin. Path. 97*:686-691) or colon (Redston *et al.* (1992) *Lab. Invest.* 66, 47A (abstract)).

Accordingly, beta-ketoacyl synthases are a major target for drug action and development. Thus, it is valuable to the field of pharmaceutical development to identify and characterize novel beta-ketoacyl synthases and tissues and disorders in which these enzymes are differentially expressed. The present invention advances the state of the art by providing novel human beta-ketoacyl synthase molecules and the uses thereof.

SUMMARY OF THE INVENTION

The present invention is based, in part, on the discovery of a novel human fatty acid synthase enzyme, referred to herein as "beta-ketoacyl synthase". The nucleotide sequence of a cDNA encoding beta-ketoacyl synthase is shown in SEQ.ID NO:1 (nucleotides 58-1437) and SEQ ID NO:3 (nucleotides 1-1380), and the amino acid sequence of a beta-ketoacyl synthase polypeptide is shown in SEQ ID NO:2.

Accordingly, in one aspect, the invention features a nucleic acid molecule which encodes a beta-ketoacyl synthase protein or polypeptide, e.g., a biologically active portion of the beta-ketoacyl synthase protein. In a preferred embodiment, the isolated nucleic acid molecule encodes a polypeptide having the amino acid sequence of SEQ ID NO:2. In other embodiments, the invention provides an isolated beta-ketoacyl synthase nucleic acid molecule having the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the sequence of the DNA insert of the plasmid deposited with ATCC Accession Number _____. In still other embodiments, the invention provides nucleic acid molecules that are substantially identical (e.g., naturally occurring allelic variants) to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the sequence of the DNA insert of the plasmid deposited

with ATCC Accession Number ____. In other embodiments, the invention provides a nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the sequence of the DNA insert of the plasmid deposited with ATCC Accession Number ____, wherein the nucleic acid encodes a full length beta-ketoacyl synthase protein or an active fragment thereof.

In a related aspect, the invention further provides nucleic acid constructs which include a beta-ketoacyl synthase nucleic acid molecule described herein. In certain embodiments, the nucleic acid molecules of the invention are operatively linked to native or heterologous regulatory sequences. Also included, are vectors and host cells containing the beta-ketoacyl synthase nucleic acid molecules of the invention e.g., vectors and host cells suitable for producing beta-ketoacyl synthase nucleic acid molecules and polypeptides.

In another related aspect, the invention provides nucleic acid fragments suitable as primers or hybridization probes for the detection of beta-ketoacyl synthase-encoding nucleic acids.

In still another related aspect, isolated nucleic acid molecules that are antisense to a beta-ketoacyl synthase encoding nucleic acid molecule are provided.

In another aspect, the invention features, beta-ketoacyl synthase polypeptides, and biologically active or antigenic fragments thereof that are useful, e.g., as reagents or targets in assays applicable to treatment and diagnosis of beta-ketoacyl synthase-mediated or -related disorders. In another embodiment, the invention provides beta-ketoacyl synthase polypeptides having a beta-ketoacyl synthase activity. Preferred polypeptides are beta-ketoacyl synthase proteins including at least one beta-ketoacyl synthase active site. In other embodiments, the invention provides beta-ketoacyl synthase polypeptides, e.g., a beta-ketoacyl synthase polypeptide having the amino acid sequence shown in SEQ ID NO:2; the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC Accession Number ___; an amino acid sequence that is substantially identical to the amino acid sequence shown in SEQ ID NO:2; or an amino acid sequence encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a

nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the sequence of the DNA insert of the plasmid deposited with ATCC Accession Number ____, wherein the nucleic acid encodes a full length beta-ketoacyl synthase protein or an active fragment thereof.

In a related aspect, the invention further provides nucleic acid constructs which include a beta-ketoacyl synthase nucleic acid molecule described herein.

In a related aspect, the invention provides beta-ketoacyl synthase polypeptides or fragments operatively linked to non-beta-ketoacyl synthase polypeptides to form fusion proteins.

In another aspect, the invention features antibodies and antigen-binding fragments thereof, that react with, or more preferably specifically bind beta-ketoacyl synthase polypeptides.

In another aspect, the invention provides methods of screening for compounds that modulate the expression or activity of the beta-ketoacyl synthase polypeptides or nucleic acids.

In still another aspect, the invention provides a process for modulating beta-ketoacyl synthase polypeptide or nucleic acid expression or activity, e.g. using the screened compounds. In certain embodiments, the methods involve treatment of conditions related to aberrant activity or expression of the beta-ketoacyl synthase polypeptides or nucleic acids, such as conditions involving aberrant or deficient cellular proliferation or differentiation.

The invention also provides assays for determining the activity of or the presence or absence of beta-ketoacyl synthase polypeptides or nucleic acid molecules in a biological sample, including for disease diagnosis.

In further aspect the invention provides assays for determining the presence or absence of a genetic alteration in a beta-ketoacyl synthase polypeptide or nucleic acid molecule, including for disease diagnosis.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B show a cDNA sequence (SEQ ID NO:1) and predicted amino acid sequence (SEQ ID NO:2) of human beta-ketoacyl synthase. The methionine-initiated open reading frame of human beta-ketoacyl synthase (without the 5' and 3' untranslated regions) extends from nucleotide position 58 to position 1437 of SEQ ID NO:1, including the terminal codon.

Figure 2 depicts a hydropathy plot of human beta-ketoacyl synthase. Relative hydrophobic residues are shown above the dashed horizontal line, and relative hydrophilic residues are below the dashed horizontal line. The cysteine residues (cys) and N glycosylation site (Ngly) are indicated by short vertical lines just below the hydropathy trace. The numbers corresponding to the amino acid sequence of human beta-ketoacyl synthase are indicated. Polypeptides of the invention include fragments which include: all or a part of a hydrophobic sequence (a sequence above the dashed line); or all or part of a hydrophilic fragment (a sequence below the dashed line). Other fragments include a cysteine residue or as N-glycosylation site.

Figure 3 shows expression of 46619 in various tissues and cells, as follows: Artery, Normal (Column 1); Aorta, Diseased (Column 2); Vein, Normal (Column 3); Coronary Smooth Muscle Cells (SMC) (Column 4); Human Umbilical Vein Endothelial Cells (HUVEC) (Column 5); Hemangioma (Column 6); Heart, Normal (Column 7); Heart, Congestive Heart Failure (CHF) (Column 8); Kidney (Column 9); Skeletal Muscle (Column 10); Adipose, Normal (Column 11); Pancreas (Column 12); Primary Osteoblasts (Column 13); Osteoclasts, Differentiated (Column 14); Skin, Normal (Column 15); Spinal Cord, Normal (Column 16); Brain Cortex, Normal (Column 17); Brain Hypothalamus, Normal (Column 18); Nerve (Column 19); Dorsal Root Ganglion (DRG) (Column 20); Breast, Normal (Column 21); Breast, Tumor (Column 22); Ovary, Normal (Column 23); Ovary, Tumor (Column 24); Prostate, Normal (Column 25); Prostate, Tumor (Column 26); Salivary Glands (Column 27);

Colon, Normal (Column 28); Colon, Tumor (Column 29); Lung, Normal (Column 30); Lung, Tumor (Column 31); Lung, Chronic Obstructive Pulmonary Disease (COPD) (Column 32); Colon, Inflammatory Bowel Disease (IBD) (Column 33); Liver, Normal (Column 34); Liver, Fibrosis (Column 35); Spleen, Normal (Column 36); Tonsil, Normal (Column 37); Lymph Node, Normal (Column 38); Small Intestine, Normal (Column 39); Macrophages (Column 40); Synovium (Column 41); Bone Marrow-Mononuclear Cells (BM-MNC) (Column 42); Activated Peripheral Blood Mononuclear Cells (PBMC) (Column 43); Neutrophils (Column 44); Megakaryocytes (Column 45); Erythroid (Column 46); Positive Control (Column 47).

Expression of 46619 was observed in the following cells and tissues: Coronary smooth muscle cells (column 4), human umbilical vein endothelial cells (column 5), normal heart (column 7), congestive heart failure tissue (column 8), kidney (column 9), skeletal muscle (column 10), pancreas (column 12), normal brain cortex (column 17), normal ovary (column 23), ovary, tumor (column 24), normal prostate (column 25), prostate, tumor (column 26), normal colon (column 28), colon, tumor (column 29), lung, tumor (column 31), inflammatory bowel disease colon tissue (column 33), normal liver (column 34), normal tonsil (column 37), megakaryocytes (column 45), erythroid (column 46), and positive control (column 47). In the prostate, colon, and lung, 46619 expression levels were found to be elevated in tumor tissue as compared with normal tissue. (See columns 25-26, 28-29, and 30-31.)

Expression levels of 46619 in various tissue and cell types were determined by quantitative RT-PCR (Reverse Transcriptase Polymerase Chain Reaction; Taqman® brand PCR kit, Applied Biosystems). The quantitative RT-PCR reactions were performed according to the kit manufacturer's instructions.

Figure 4 shows expression of 46619 in various tissues and cells, as follows: Breast, Normal (Column 1); Breast, Normal (Column 2); Breast, Normal (Column 3); Breast, Tumor - Moderately Differentiated-Invasive Ductal Carcinoma (MD-IDC); (Column 4); Breast, Tumor (Column 5); Breast, Tumor: Poorly Differentiated-IDC Or Invasive Lobular Carcinoma (ILC) (Column 6); Breast, Tumor (IDC) (Column 7);

Breast, Tumor (ILC) (Column 8); Lymph Node (Breast Metasteses) (Column 9); Lung (Breast Metasteses) (Column 10); Ovary, Normal (Column 11); Ovary, Normal (Column 12); Ovary, Tumor (Column 13); Ovary, Tumor (Column 14); Ovary, Tumor (Column 15); Ovary, Tumor (Column 16); Ovary, Tumor (Column 17); Lung, Normal (Column 18); Lung, Normal (Column 19); Lung, Normal (Column 20); Lung, Tumor (Column 21); Lung, Tumor (Column 22); Lung, Tumor (Column 23); Lung, Tumor: Squamous Cell Carcinoma (SCC) (Column 24); Lung, Tumor (Column 25); Lung, Tumor (Column 26); Colon, Normal (Column 27); Colon, Normal (Column 28); Colon, Normal (Column 29); Colon, Tumor - MD (Column 30); Colon Tumor - MD (Column 31); Colon Tumor (Column 32); Colon Tumor MD-PD (Column 33); Colon-Liver Metasteses (Column 34); Colon-Liver Metasteses (Column 35); Liver Normal (Female) (Column 36); Cervix Squamous (Column 37); Cervix Squamous (Column 38); Human Umbilical Vein Endothelial Cells (HMVEC)-Arresting (Column 39); HMVEC-Proliferating (Column 40); Pooled Hemangiomas (Column 41); Colon Cancer Cell Line HCT116H22 - Normoxic (Column 42); Colon Cancer Cell Line HCT116H22 - Hypoxic (Column 43).

Expression of 46619 was observed in the following cells and tissues: Ovary, tumor (columns 15-16), lung, tumor (22, 24-25), normal colon and colon tumor (columns 29-33), normal liver (column 36), proliferating human umbilical vein endothelial cells (column 40), normoxic colon cancer (column 42), and hypoxic colon cancer (column 43). Expression levels were determined as described in Figure 3.

Figure 5 shows expression of 46619 in various tissues and cells, as follows: Breast, Tumor (Column 1); Breast, Tumor (Column 2); Breast, Tumor (Column 3); Breast, Tumor (Column 4); Breast Tumor (Column 5); Breast (Column 6); Colon, Tumor (Stage C) (Column 7); Colon, Tumor (Stage B) (Column 8); Colon, Tumor (Stage C) (Column 9); HCT116 Colon Cancer Cell Line (Column 10); HT29 Colon Cancer Cell Line (Column 11); Colo 205 Colon Cancer Cell Line (Column 12); NCIH125 Lung Cancer Cell Line (Column 13); NCIH67 Cell Line (Column 14); NCIH322 Lung Cancer Cell Line (Column 15); NCIH460 Lung Cancer Cell Line (Column 16); A549 Lung Cancer Cell Line (Column 17); Normal Human Broinchial

Epithelial (NHBE) (Column 18); SKOV-3 Ovary Cell Line (Column 19); OVCAR-3 Ovary Cell Line (Column 20); 293 Baby Kidney Cell Line (Column 21); 293T Baby Kidney Cell Line (Column 22). Expression levels were determined as described in Figure 3.

Figures 6A and 6B depict an alignment of the 46619 beta-ketoacyl-synthase with a consensus amino acid sequence derived from a hidden Markov model. The upper sequence is the consensus amino acid sequence (SEQ ID NO:4), while the lower amino acid sequence corresponds to amino acids 44 to 459 of SEQ ID NO:2.

Other features and advantages of the invention will be apparent from the following detailed description, and from the claims.

DETAILED DESCRIPTION OF THE INVENTION

Human beta-ketoacyl synthase

The human beta-ketoacyl synthase sequence (Figures 1Aand 1B; SEQ ID NO:1), which is approximately 1489 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1380 nucleotides (nucleotides 58-1437 of SEQ ID NO:1; nucleotides 1-1380 of SEQ ID NO:3) including the terminal codon. The coding sequence encodes a 459 amino acid protein (SEQ ID NO:2).

This protein form is approximately 459 amino acid residues in length. Human beta-ketoacyl synthase contains the following regions or other structural features: a predicted beta-ketoacyl synthase active site located at about amino acid residues 200 to 216 of SEQ ID NO:2; and MEMSAT predicted transmembrane domains which extend from about amino acid residue 44 to 63; 144 to 160; 230 to 248; and 386 to 404.

The beta-ketoacyl synthase protein also includes the following domains: N-glycosylation site at aa 451 to 454; protein kinase C phosphorylation sites at aa 13 to 15, aa 27 to 29, aa 193 to 195, aa 219 to 221, aa 378 to 380, aa 434 to 436; casein

kinase II phosphorylation sites at 108 to 111, 139 to 142, 256 to 259, 315 to 318, 416 to 419, tryrosine kinase phosphorylation sitess at aa 297 to 305; N-myristoylation sites at aa 68 to 73, aa 95 to 100, aa 147 to 152, aa 234 to 239, aa 323 to 328, aa 388 to 393, ribosomal protein S14 signature at aa 20 to 43; and a beta-ketoacyl synthase signature at aa 200 to 216.

A plasmid containing the nucleotide sequence encoding human beta-ketoacyl synthase was deposited with American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on _____ and assigned Accession Number _____. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

As used herein, the term "beta-ketoacyl synthase" refers to a protein or polypeptide which is capable of forming acetoacetyl-ACP. In this condensation reaction, the acetyl group is transferred from the cysteine-SH group to the malonyl group so that the aetyl group becomes the methyl terminal two-carbon unit of the new acetoacetyl group. Assays for beta-ketoacyl synthases have been previously described (Joshi *et al.* (1993) *Biochem. J. 296*:143-149) and (Kim *et al.* (1977) *Arch. Biochem. Biophys. 178*:475-485). *See also* Juayakumar *et al.* (1995) *Biochemistry* 92:8695-8699 and Stoops *et al.* (1981) *J. Biol. Chem.* 256:5128-5133.

Typically, beta-ketoacyl synthases play a role in diverse cellular processes. For example, the biosynthesized fatty acids are incorporated into the membrane lipids and can be used for long-term energy storage. In animals, liver is the primary tissue in which fatty acids are synthesized and their usual fate is to be used in the biosynthesis of triacylglycerols. Catabolism of triacylglycerols results in the release of fatty acids that are absorbed by various tissues. There they can be degraded to satisfy immediate energy needs or reesterified into phospholipids for membrane assembly.

A beta-ketoacyl synthase polypeptide can include a “beta-ketoacyl synthase active site” or regions homologous with a "beta-ketoacyl synthase active site. "

As the beta-ketoacyl synthase polypeptides of the invention may modulate beta-ketoacyl synthase-mediated activities, they may be useful for developing novel diagnostic and therapeutic agents for beta-ketoacyl synthase-mediated or related disorders, as described below.

As used herein, a “beta-ketoacyl synthase activity”, “biological activity of beta-ketoacyl synthase” or "functional activity of beta-ketoacyl synthase", refers to an activity exerted by a beta-ketoacyl synthase protein, polypeptide or nucleic acid molecule on e.g., a beta-ketoacyl synthase-responsive cell or on a beta-ketoacyl synthase substrate, e.g., acyl or malonyl groups or as determined *in vivo* or *in vitro*. In one embodiment, a beta-ketoacyl synthase activity is a direct activity, such as an association with a beta-ketoacyl synthase target molecule. A "target molecule" or "binding partner" is a molecule with which a beta-ketoacyl synthase protein binds or interacts in nature.

Accordingly, beta-ketoacyl synthase protein may be mediate various disorders, including cellular proliferative and/or differentiative disorders, brain disorders, heart disorders, blood vessel disorders, and platelet disorders.

Examples of cellular proliferative and/or differentiative disorders include cancer, e.g., carcinoma, sarcoma, metastatic disorders or hematopoietic neoplastic disorders, e.g., leukemias. A metastatic tumor can arise from a multitude of primary tumor types, including but not limited to those of prostate, colon, lung, breast and liver origin.

As used herein, the terms “cancer”, "hyperproliferative" and "neoplastic" refer to cells having the capacity for autonomous growth, i.e., an abnormal state or condition characterized by rapidly proliferating cell growth. Hyperproliferative and neoplastic disease states may be categorized as pathologic, i.e., characterizing or constituting a disease state, or may be categorized as non-pathologic, i.e., a deviation from normal but not associated with a disease state. The term is meant to include all types of cancerous growths or oncogenic processes, metastatic tissues or malignantly transformed cells, tissues, or organs, irrespective of histopathologic type or stage of invasiveness. "Pathologic hyperproliferative" cells occur in disease states

characterized by malignant tumor growth. Examples of non-pathologic hyperproliferative cells include proliferation of cells associated with wound repair.

The terms "cancer" or "neoplasms" include malignancies of the various organ systems, such as affecting lung, breast, thyroid, lymphoid, gastrointestinal, and genito-urinary tract, as well as adenocarcinomas which include malignancies such as most colon cancers, renal-cell carcinoma, prostate cancer and/or testicular tumors, non-small cell carcinoma of the lung, cancer of the small intestine and cancer of the esophagus.

The term "carcinoma" is art recognized and refers to malignancies of epithelial or endocrine tissues including respiratory system carcinomas, gastrointestinal system carcinomas, genitourinary system carcinomas, testicular carcinomas, breast carcinomas, prostatic carcinomas, endocrine system carcinomas, and melanomas. Exemplary carcinomas include those forming from tissue of the cervix, lung, prostate, breast, head and neck, colon and ovary. The term also includes carcinosarcomas, e.g., which include malignant tumors composed of carcinomatous and sarcomatous tissues. An "adenocarcinoma" refers to a carcinoma derived from glandular tissue or in which the tumor cells form recognizable glandular structures.

The term "sarcoma" is art recognized and refers to malignant tumors of mesenchymal derivation.

As noted above, expression of the beta-ketoacyl synthase of the invention is observed in erythroid, brain, heart, blood vessel, lung, liver, kidney, breast, prostate, skeletal muscle, pancreas, ovary, and colon tissues and cells. (See, e.g., Figures 3-5.)

Thus, the beta-ketoacyl synthase nucleic acid and protein of the invention can be used to treat and/or diagnose a variety of proliferative disorders. E.g., such disorders include hematopoietic neoplastic disorders. As used herein, the term "hematopoietic neoplastic disorders" includes diseases involving hyperplastic/neoplastic cells of hematopoietic origin, e.g., arising from myeloid, lymphoid or erythroid lineages, or precursor cells thereof. Preferably, the diseases arise from poorly differentiated acute leukemias, e.g., erythroblastic leukemia and acute megakaryoblastic leukemia. Additional exemplary myeloid disorders include, but are not limited to, acute promyeloid leukemia (APML), acute myelogenous

leukemia (AML) and chronic myelogenous leukemia (CML) (reviewed in Vaickus (1991) *Crit. Rev. in Oncol./Hemotol.* 11:267-97); lymphoid malignancies include, but are not limited to acute lymphoblastic leukemia (ALL) which includes B-lineage ALL and T-lineage ALL, chronic lymphocytic leukemia (CLL), prolymphocytic leukemia (PLL), hairy cell leukemia (HLL) and Waldenstrom's macroglobulinemia (WM). Additional forms of malignant lymphomas include, but are not limited to non-Hodgkin lymphoma and variants thereof, peripheral T cell lymphomas, adult T cell leukemia/lymphoma (ATL), cutaneous T-cell lymphoma (CTCL), large granular lymphocytic leukemia (LGF), Hodgkin's disease and Reed-Sternberg disease.

Disorders involving the brain include, but are not limited to, disorders involving neurons, and disorders involving glia, such as astrocytes, oligodendrocytes, ependymal cells, and microglia; cerebral edema, raised intracranial pressure and herniation, and hydrocephalus; malformations and developmental diseases, such as neural tube defects, forebrain anomalies, posterior fossa anomalies, and syringomyelia and hydromyelia; perinatal brain injury; cerebrovascular diseases, such as those related to hypoxia, ischemia, and infarction, including hypotension, hypoperfusion, and low-flow states--global cerebral ischemia and focal cerebral ischemia--infarction from obstruction of local blood supply, intracranial hemorrhage, including intracerebral (intraparenchymal) hemorrhage, subarachnoid hemorrhage and ruptured berry aneurysms, and vascular malformations, hypertensive cerebrovascular disease, including lacunar infarcts, slit hemorrhages, and hypertensive encephalopathy; infections, such as acute meningitis, including acute pyogenic (bacterial) meningitis and acute aseptic (viral) meningitis, acute focal suppurative infections, including brain abscess, subdural empyema, and extradural abscess, chronic bacterial meningoencephalitis, including tuberculosis and mycobacterioses, neurosyphilis, and neuroborreliosis (Lyme disease), viral meningoencephalitis, including arthropod-borne (Arbo) viral encephalitis, *Herpes simplex* virus Type 1, *Herpes simplex* virus Type 2, *Varicalla-zoster* virus (*Herpes zoster*), cytomegalovirus, poliomyelitis, rabies, and human immunodeficiency virus 1, including HIV-1 meningoencephalitis (subacute encephalitis), vacuolar myelopathy, AIDS-associated myopathy, peripheral neuropathy, and AIDS in children, progressive multifocal leukoencephalopathy,

subacute sclerosing panencephalitis, fungal meningoencephalitis, other infectious diseases of the nervous system; transmissible spongiform encephalopathies (prion diseases); demyelinating diseases, including multiple sclerosis, multiple sclerosis variants, acute disseminated encephalomyelitis and acute necrotizing hemorrhagic encephalomyelitis, and other diseases with demyelination; degenerative diseases, such as degenerative diseases affecting the cerebral cortex, including Alzheimer disease and Pick disease, degenerative diseases of basal ganglia and brain stem, including Parkinsonism, idiopathic Parkinson disease (paralysis agitans), progressive supranuclear palsy, corticobasal degenration, multiple system atrophy, including striatonigral degenration, Shy-Drager syndrome, and olivopontocerebellar atrophy, and Huntington disease; spinocerebellar degenerations, including spinocerebellar ataxias, including Friedreich ataxia, and ataxia-telanglectasia, degenerative diseases affecting motor neurons, including amyotrophic lateral sclerosis (motor neuron disease), bulbospinal atrophy (Kennedy syndrome), and spinal muscular atrophy; inborn errors of metabolism, such as leukodystrophies, including Krabbe disease, metachromatic leukodystrophy, adrenoleukodystrophy, Pelizaeus-Merzbacher disease, and Canavan disease, mitochondrial encephalomyopathies, including Leigh disease and other mitochondrial encephalomyopathies; toxic and acquired metabolic diseases, including vitamin deficiencies such as thiamine (vitamin B_1) deficiency and vitamin B_{12} deficiency, neurologic sequelae of metabolic disturbances, including hypoglycemia, hyperglycemia, and hepatic encephatopathy, toxic disorders, including carbon monoxide, methanol, ethanol, and radiation, including combined methotrexate and radiation-induced injury; tumors, such as gliomas, including astrocytoma, including fibrillary (diffuse) astrocytoma and glioblastoma multiforme, pilocytic astrocytoma, pleomorphic xanthoastrocytoma, and brain stem glioma, oligodendroglioma, and ependymoma and related paraventricular mass lesions, neuronal tumors, poorly differentiated neoplasms, including medulloblastoma, other parenchymal tumors, including primary brain lymphoma, germ cell tumors, and pineal parenchymal tumors, meningiomas, metastatic tumors, paraneoplastic syndromes, peripheral nerve sheath tumors, including schwannoma, neurofibroma, and malignant peripheral nerve sheath tumor (malignant schwannoma), and

neurocutaneous syndromes (phakomatoses), including neurofibromotosis, including Type 1 neurofibromatosis (NF1) and TYPE 2 neurofibromatosis (NF2), tuberous sclerosis, and Von Hippel-Lindau disease.

Disorders involving the heart, include but are not limited to, heart failure, including but not limited to, cardiac hypertrophy, left-sided heart failure, and right-sided heart failure; ischemic heart disease, including but not limited to angina pectoris, myocardial infarction, chronic ischemic heart disease, and sudden cardiac death; hypertensive heart disease, including but not limited to, systemic (left-sided) hypertensive heart disease and pulmonary (right-sided) hypertensive heart disease; valvular heart disease, including but not limited to, valvular degeneration caused by calcification, such as calcific aortic stenosis, calcification of a congenitally bicuspid aortic valve, and mitral annular calcification, and myxomatous degeneration of the mitral valve (mitral valve prolapse), rheumatic fever and rheumatic heart disease, infective endocarditis, and noninfected vegetations, such as nonbacterial thrombotic endocarditis and endocarditis of systemic lupus erythematosus (Libman-Sacks disease), carcinoid heart disease, and complications of artificial valves; myocardial disease, including but not limited to dilated cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy, and myocarditis; pericardial disease, including but not limited to, pericardial effusion and hemopericardium and pericarditis, including acute pericarditis and healed pericarditis, and rheumatoid heart disease; neoplastic heart disease, including but not limited to, primary cardiac tumors, such as myxoma, lipoma, papillary fibroelastoma, rhabdomyoma, and sarcoma, and cardiac effects of noncardiac neoplasms; congenital heart disease, including but not limited to, left-to-right shunts--late cyanosis, such as atrial septal defect, ventricular septal defect, patent ductus arteriosus, and atrioventricular septal defect, right-to-left shunts--early cyanosis, such as tetralogy of fallot, transposition of great arteries, truncus arteriosus, tricuspid atresia, and total anomalous pulmonary venous connection, obstructive congenital anomalies, such as coarctation of aorta, pulmonary stenosis and atresia, and aortic stenosis and atresia, and disorders involving cardiac transplantation.

Disorders involving blood vessels include, but are not limited to, responses of vascular cell walls to injury, such as endothelial dysfunction and endothelial activation and intimal thickening; vascular diseases including, but not limited to, congenital anomalies, such as arteriovenous fistula, atherosclerosis, and hypertensive vascular disease, such as hypertension; inflammatory disease--the vasculitides, such as giant cell (temporal) arteritis, Takayasu arteritis, polyarteritis nodosa (classic), Kawasaki syndrome (mucocutaneous lymph node syndrome), microscopic polyanglitis (microscopic polyarteritis, hypersensitivity or leukocytoclastic anglitis), Wegener granulomatosis, thromboanglitis obliterans (Buerger disease), vasculitis associated with other disorders, and infectious arteritis; Raynaud disease; aneurysms and dissection, such as abdominal aortic aneurysms, syphilitic (luetic) aneurysms, and aortic dissection (dissecting hematoma); disorders of veins and lymphatics, such as varicose veins, thrombophlebitis and phlebothrombosis, obstruction of superior vena cava (superior vena cava syndrome), obstruction of inferior vena cava (inferior vena cava syndrome), and lymphangitis and lymphedema; tumors, including benign tumors and tumor-like conditions, such as hemangioma, lymphangioma, glomus tumor (glomangioma), vascular ectasias, and bacillary angiomatosis, and intermediate-grade (borderline low-grade malignant) tumors, such as Kaposi sarcoma and hemangloendothelioma, and malignant tumors, such as angiosarcoma and hemangiopericytoma; and pathology of therapeutic interventions in vascular disease, such as balloon angioplasty and related techniques and vascular replacement, such as coronary artery bypass graft surgery.

Disorders involving the lung include, but are not limited to, congenital anomalies; atelectasis; diseases of vascular origin, such as pulmonary congestion and edema, including hemodynamic pulmonary edema and edema caused by microvascular injury, adult respiratory distress syndrome (diffuse alveolar damage), pulmonary embolism, hemorrhage, and infarction, and pulmonary hypertension and vascular sclerosis; chronic obstructive pulmonary disease, such as emphysema, chronic bronchitis, bronchial asthma, and bronchiectasis; diffuse interstitial (infiltrative, restrictive) diseases, such as pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity

pneumonitis, pulmonary eosinophilia (pulmonary infiltration with eosinophilia), *Bronchiolitis obliterans*-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, including Goodpasture syndrome, idiopathic pulmonary hemosiderosis and other hemorrhagic syndromes, pulmonary involvement in collagen vascular disorders, and pulmonary alveolar proteinosis; complications of therapies, such as drug-induced lung disease, radiation-induced lung disease, and lung transplantation; tumors, such as bronchogenic carcinoma, including paraneoplastic syndromes, bronchioloalveolar carcinoma, neuroendocrine tumors, such as bronchial carcinoid, miscellaneous tumors, and metastatic tumors; pathologies of the pleura, including inflammatory pleural effusions, noninflammatory pleural effusions, pneumothorax, and pleural tumors, including solitary fibrous tumors (pleural fibroma) and malignant mesothelioma.

Disorders involving the liver include, but are not limited to, hepatic injury; jaundice and cholestasis, such as bilirubin and bile formation; hepatic failure and cirrhosis, such as cirrhosis, portal hypertension, including ascites, portosystemic shunts, and splenomegaly; infectious disorders, such as viral hepatitis, including hepatitis A-E infection and infection by other hepatitis viruses, clinicopathologic syndromes, such as the carrier state, asymptomatic infection, acute viral hepatitis, chronic viral hepatitis, and fulminant hepatitis; autoimmune hepatitis; drug- and toxin-induced liver disease, such as alcoholic liver disease; inborn errors of metabolism and pediatric liver disease, such as hemochromatosis, Wilson disease, a_1-antitrypsin deficiency, and neonatal hepatitis; intrahepatic biliary tract disease, such as secondary biliary cirrhosis, primary biliary cirrhosis, primary sclerosing cholangitis, and anomalies of the biliary tree; circulatory disorders, such as impaired blood flow into the liver, including hepatic artery compromise and portal vein obstruction and thrombosis, impaired blood flow through the liver, including passive congestion and centrilobular necrosis and peliosis hepatis, hepatic vein outflow obstruction, including hepatic vein thrombosis (Budd-Chiari syndrome) and veno-occlusive disease; hepatic disease associated with pregnancy, such as preeclampsia and eclampsia, acute fatty liver of pregnancy, and intrehepatic cholestasis of pregnancy; hepatic complications of organ or bone marrow transplantation, such as drug toxicity after bone marrow

transplantation, graft-versus-host disease and liver rejection, and nonimmunologic damage to liver allografts; tumors and tumorous conditions, such as nodular hyperplasias, adenomas, and malignant tumors, including primary carcinoma of the liver and metastatic tumors.

Disorders involving the kidney include, but are not limited to, congenital anomalies including, but not limited to, cystic diseases of the kidney, that include but are not limited to, cystic renal dysplasia, autosomal dominant (adult) polycystic kidney disease, autosomal recessive (childhood) polycystic kidney disease, and cystic diseases of renal medulla, which include, but are not limited to, medullary sponge kidney, and nephronophthisis-uremic medullary cystic disease complex, acquired (dialysis-associated) cystic disease, such as simple cysts; glomerular diseases including pathologies of glomerular injury that include, but are not limited to, in situ immune complex deposition, that includes, but is not limited to, anti-GBM nephritis, Heymann nephritis, and antibodies against planted antigens, circulating immune complex nephritis, antibodies to glomerular cells, cell-mediated immunity in glomerulonephritis, activation of alternative complement pathway, epithelial cell injury, and pathologies involving mediators of glomerular injury including cellular and soluble mediators, acute glomerulonephritis, such as acute proliferative (poststreptococcal, postinfectious) glomerulonephritis, including but not limited to, poststreptococcal glomerulonephritis and nonstreptococcal acute glomerulonephritis, rapidly progressive (crescentic) glomerulonephritis, nephrotic syndrome, membranous glomerulonephritis (membranous nephropathy), minimal change disease (lipoid nephrosis), focal segmental glomerulosclerosis, membranoproliferative glomerulonephritis, IgA nephropathy (Berger disease), focal proliferative and necrotizing glomerulonephritis (focal glomerulonephritis), hereditary nephritis, including but not limited to, Alport syndrome and thin membrane disease (benign familial hematuria), chronic glomerulonephritis, glomerular lesions associated with systemic disease, including but not limited to, systemic lupus erythematosus, Henoch-Schönlein purpura, bacterial endocarditis, diabetic glomerulosclerosis, amyloidosis, fibrillary and immunotactoid glomerulonephritis, and other systemic disorders; diseases affecting tubules and interstitium, including acute tubular necrosis and tubulointerstitial nephritis, including

but not limited to, pyelonephritis and urinary tract infection, acute pyelonephritis, chronic pyelonephritis and reflux nephropathy, and tubulointerstitial nephritis induced by drugs and toxins, including but not limited to, acute drug-induced interstitial nephritis, analgesic abuse nephropathy, nephropathy associated with nonsteroidal anti-inflammatory drugs, and other tubulointerstitial diseases including, but not limited to, urate nephropathy, hypercalcemia and nephrocalcinosis, and multiple myeloma; diseases of blood vessels including benign nephrosclerosis, malignant hypertension and accelerated nephrosclerosis, renal artery stenosis, and thrombotic microangiopathies including, but not limited to, classic (childhood) hemolytic-uremic syndrome, adult hemolytic-uremic syndrome/thrombotic thrombocytopenic purpura, idiopathic HUS/TTP, and other vascular disorders including, but not limited to, atherosclerotic ischemic renal disease, atheroembolic renal disease, sickle cell disease nephropathy, diffuse cortical necrosis, and renal infarcts; urinary tract obstruction (obstructive uropathy); urolithiasis (renal calculi, stones); and tumors of the kidney including, but not limited to, benign tumors, such as renal papillary adenoma, renal fibroma or hamartoma (renomedullary interstitial cell tumor), angiomyolipoma, and oncocytoma, and malignant tumors, including renal cell carcinoma (hypernephroma, adenocarcinoma of kidney), which includes urothelial carcinomas of renal pelvis.

Disorders of the breast include, but are not limited to, disorders of development; inflammations, including but not limited to, acute mastitis, periductal mastitis, periductal mastitis (recurrent subareolar abscess, squamous metaplasia of lactiferous ducts), mammary duct ectasia, fat necrosis, granulomatous mastitis, and pathologies associated with silicone breast implants; fibrocystic changes; proliferative breast disease including, but not limited to, epithelial hyperplasia, sclerosing adenosis, and small duct papillomas; tumors including, but not limited to, stromal tumors such as fibroadenoma, phyllodes tumor, and sarcomas, and epithelial tumors such as large duct papilloma; carcinoma of the breast including in situ (noninvasive) carcinoma that includes ductal carcinoma in situ (including Paget's disease) and lobular carcinoma in situ, and invasive (infiltrating) carcinoma including, but not limited to, invasive ductal carcinoma, no special type, invasive lobular carcinoma, medullary carcinoma, colloid (mucinous) carcinoma, tubular carcinoma, and invasive papillary carcinoma, and miscellaneous malignant neoplasms.

Disorders in the male breast include, but are not limited to, gynecomastia and carcinoma.

Disorders involving the prostate include, but are not limited to, inflammations, benign enlargement, for example, nodular hyperplasia (benign prostatic hypertrophy or hyperplasia), and tumors such as carcinoma.

Disorders involving the skeletal muscle include tumors such as rhabdomyosarcoma.

Disorders involving the pancreas include those of the exocrine pancreas such as congenital anomalies, including but not limited to, ectopic pancreas; pancreatitis, including but not limited to, acute pancreatitis; cysts, including but not limited to, pseudocysts; tumors, including but not limited to, cystic tumors and carcinoma of the pancreas; and disorders of the endocrine pancreas such as, diabetes mellitus; islet cell tumors, including but not limited to, insulinomas, gastrinomas, and other rare islet cell tumors.

Disorders involving the ovary include, for example, polycystic ovarian disease, Stein-leventhal syndrome, Pseudomyxoma peritonei and stromal hyperthecosis; ovarian tumors such as, tumors of coelomic epithelium, serous tumors, mucinous tumors, endometeriod tumors, clear cell adenocarcinoma, cystadenofibroma, brenner tumor, surface epithelial tumors; germ cell tumors such as mature (benign) teratomas, monodermal teratomas, immature malignant teratomas, dysgerminoma, endodermal sinus tumor, choriocarcinoma; sex cord-stomal tumors such as, granulosa-theca cell tumors, thecoma-fibromas, androblastomas, hill cell tumors, and gonadoblastoma; and metastatic tumors such as Krukenberg tumors.

Disorders involving the colon include, but are not limited to, congenital anomalies, such as atresia and stenosis, Meckel diverticulum, congenital aganglionic megacolon-Hirschsprung disease; enterocolitis, such as diarrhea and dysentery, infectious enterocolitis, including viral gastroenteritis, bacterial enterocolitis, necrotizing enterocolitis, antibiotic-associated colitis (pseudomembranous colitis), and collagenous and lymphocytic colitis, miscellaneous intestinal inflammatory disorders, including parasites and protozoa, acquired immunodeficiency syndrome, transplantation, drug-induced intestinal injury, radiation enterocolitis, neutropenic

colitis (typhlitis), and diversion colitis; idiopathic inflammatory bowel disease, such as Crohn disease and ulcerative colitis; tumors of the colon, such as non-neoplastic polyps, adenomas, familial syndromes, colorectal carcinogenesis, colorectal carcinoma, and carcinoid tumors.

The beta-ketoacyl synthase protein, fragments thereof, and derivatives and other variants of the sequence in SEQ ID NO:2 are collectively referred to as "polypeptides or proteins of the invention" or " beta-ketoacyl synthase polypeptides or proteins". Nucleic acid molecules encoding such polypeptides or proteins are collectively referred to as "nucleic acids of the invention" or "beta-ketoacyl synthase nucleic acids." Beta-ketoacyl synthase molecules refer to beta-ketoacyl synthase nucleic acids, polypeptides, and antibodies.

As used herein, the term "nucleic acid molecule" includes DNA molecules (e.g., a cDNA or genomic DNA) and RNA molecules (e.g., an mRNA) and analogs of the DNA or RNA generated, e.g., by the use of nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

The term "isolated or purified nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and/or 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of 5' and/or 3' nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

As used herein, the term "hybridizes under stringent conditions" describes conditions for hybridization and washing. Stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Aqueous and nonaqueous methods are described in that reference and either can be used. A preferred, example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50° C. Another example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 55°C. A further example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 60°C. Preferably, stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 65°C. Particularly preferred stringency conditions (and the conditions that should be used if the practitioner is uncertain about what conditions should be applied to determine if a molecule is within a hybridization limitation of the invention) are 0.5M Sodium Phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2X SSC, 1% SDS at 65°C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1 corresponds to a naturally-occurring nucleic acid molecule.

As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a beta-ketoacyl synthase protein, preferably a mammalian beta-ketoacyl synthase protein, and can further include non-coding regulatory sequences, and introns.

An "isolated" or "purified" polypeptide or protein is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free from chemical precursors or other

chemicals when chemically synthesized. In one embodiment, the language "substantially free" means preparation of beta-ketoacyl synthaseprotein having less than about 30%, 20%, 10% and more preferably 5% (by dry weight), of non-beta-ketoacyl synthase protein (also referred to herein as a "contaminating protein"), or of chemical precursors or non-beta-ketoacyl synthase chemicals. When the beta-ketoacyl synthase protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation. The invention includes isolated or purified preparations of at least 0.01, 0.1, 1.0, and 10 milligrams in dry weight.

A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of beta-ketoacyl synthase (e.g., the sequence of SEQ ID NO:2 or the amino acid sequence encoded by the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____) without abolishing or more preferably, without substantially altering a biological activity, whereas an "essential" amino acid residue results in such a change. For example, amino acid residues that are conserved among the polypeptides of the present invention, e.g., those present in the fatty acid synthase domain, are predicted to be particularly unamenable to alteration.

A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a beta-ketoacyl synthase protein is preferably replaced with another amino acid residue from the same

side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a beta-ketoacyl synthase coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for beta-ketoacyl synthase biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

As used herein, a "biologically active portion" of a beta-ketoacyl synthase protein includes a fragment of a beta-ketoacyl synthase protein which participates in an interaction between a beta-ketoacyl synthase molecule and a non-beta-ketoacyl synthase molecule. Biologically active portions of a beta-ketoacyl synthase protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the beta-ketoacyl synthase protein, e.g., the amino acid sequence shown in SEQ ID NO:2, which include less amino acids than the full length beta-ketoacyl synthase proteins, and exhibit at least one activity of a beta-ketoacyl synthase protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the beta-ketoacyl synthase protein, e.g., fatty acid synthase activity. A biologically active portion of a beta-ketoacyl synthase protein can be a polypeptide which is, for example, 10, 25, 50, 100, 200 or more amino acids in length.

Alternatively, a fragment of a polypeptide of the present invention comprises an amino acid sequence consisting of amino acid residues 1-20, 20-40, 40-60, 60-80, 80-100, 100-120, 120-140, 140-160, 160-180, 180-200, 200-220, 220-240, 240-260, 260-280, 280-300, 300-320, 320-340, 340-360, 360-380, 380-400, 400-420, 420-440, or 440-459 of SEQ ID NO:2. Biologically active portions of a beta-ketoacyl synthase protein can be used as targets for developing agents which modulate a beta-ketoacyl synthase mediated activity, e.g., fatty acid synthase activity.

Calculations of homology or sequence identity between sequences (the terms are used interchangeably herein) are performed as follows.

To determine the percent identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes

(e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, 90%, 100% of the length of the reference sequence (e.g., when aligning a second sequence to the beta-ketoacyl synthase amino acid sequence of SEQ ID NO:2 having 138 amino acid residues, at least 184, preferably at least 230, more preferably at least 278, even more preferably at least 321, and even more preferably at least 367, 413 or 459 amino acid residues are aligned. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A particularly preferred set of parameters (and the one that should be used if the

practitioner is uncertain about what parameters should be applied to determine if a molecule is within a sequence identity or homology limitation of the invention) is using a Blossum 62 scoring matrix with a gap open penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

The percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences described herein can be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al., (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to beta-ketoacyl synthase nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to beta-ketoacyl synthase protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov.

"Misexpression or aberrant expression", as used herein, refers to a non-wild type pattern of gene expression, at the RNA or protein level. It includes: expression at non-wild type levels, i.e., over or under expression; a pattern of expression that differs from wild type in terms of the time or stage at which the gene is expressed, e.g., increased or decreased expression (as compared with wild type) at a predetermined developmental period or stage; a pattern of expression that differs from wild type in terms of decreased expression (as compared with wild type) in a predetermined cell type or tissue type; a pattern of expression that differs from wild type in terms of the splicing size, amino acid sequence, post-transitional modification,

or biological activity of the expressed polypeptide; a pattern of expression that differs from wild type in terms of the effect of an environmental stimulus or extracellular stimulus on expression of the gene, e.g., a pattern of increased or decreased expression (as compared with wild type) in the presence of an increase or decrease in the strength of the stimulus.

"Subject", as used herein, can refer to a mammal, e.g., a human, or to an experimental or animal or disease model. The subject can also be a non-human animal, e.g., a horse, cow, goat, or other domestic animal.

A "purified preparation of cells", as used herein, refers to, in the case of plant or animal cells, an in vitro preparation of cells and not an entire intact plant or animal. In the case of cultured cells or microbial cells, it consists of a preparation of at least 10% and more preferably 50% of the subject cells.

Various aspects of the invention are described in further detail below.

Isolated Nucleic Acid Molecules

In one aspect, the invention provides, an isolated or purified, nucleic acid molecule that encodes a beta-ketoacyl synthase polypeptide described herein, e.g., a full length beta-ketoacyl synthase protein or a fragment thereof, e.g., a biologically active portion of beta-ketoacyl synthase protein. Also included is a nucleic acid fragment suitable for use as a hybridization probe, which can be used, e.g., to a identify nucleic acid molecule encoding a polypeptide of the invention, beta-ketoacyl synthase mRNA, and fragments suitable for use as primers, e.g., PCR primers for the amplification or mutation of nucleic acid molecules.

In one embodiment, an isolated nucleic acid molecule of the invention includes the nucleotide sequence shown in SEQ ID NO:1, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences. In one embodiment, the nucleic acid molecule includes sequences encoding the human beta-ketoacyl synthase protein (i.e., "the coding region", from nucleotides of SEQ ID NO:1, not including the terminal codon), as well as 5' untranslated sequences (nucleotides of SEQ ID NO:1). Alternatively, the nucleic acid molecule can include only the coding region of SEQ ID

NO:1 (e.g., nucleotides of SEQ ID NO:1). and e.g., no flanking sequences which normally accompany the subject sequence. In another embodiment, the nucleic acid molecule encodes a sequence corresponding to the mature protein of SEQ ID NO:2.

In another embodiment, an isolated nucleic acid molecule of the invention includes a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences. In other embodiments, the nucleic acid molecule of the invention is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby forming a stable duplex.

In one embodiment, an isolated nucleic acid molecule of the present invention includes a nucleotide sequence which is at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more homologous to the nucleotide sequence shown in SEQ ID NO:1 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In the case of an isolated nucleic acid molecule which is longer than or equivalent in length to the reference sequence, e.g., SEQ ID NO:1, the comparison is made with the full length of the reference sequence. Where the isolated nucleic acid molecule is shorter than the reference sequence, e.g., shorter than SEQ ID NO:1, the comparison is made to a segment of the reference sequence of the same length (excluding any loop required by the homology calculation).

Beta-Ketoacyl Synthase Nucleic Acid Fragments

A nucleic acid molecule of the invention can include only a portion of the nucleic acid sequence of SEQ ID NO:1 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. For example, such a nucleic acid molecule can include a fragment which can be used as a probe or

primer or a fragment encoding a portion of a beta-ketoacyl synthase protein, e.g., an immunogenic or biologically active portion of a beta-ketoacyl synthase protein. A fragment can comprise: nucleotides 187-1434 of SEQ ID NO:1, which encodes a beta-ketoacyl synthase domain of human beta-ketoacyl synthase. The nucleotide sequence determined from the cloning of the beta-ketoacyl synthase gene allows for the generation of probes and primers designed for use in identifying and/or cloning other beta-ketoacyl synthase family members, or fragments thereof, as well as beta-ketoacyl synthase homologues, or fragments thereof, from other species.

In another embodiment, a nucleic acid includes a nucleotide sequence that includes part, or all, of the coding region and extends into either (or both) the 5' or 3' noncoding region. Other embodiments include a fragment which includes a nucleotide sequence encoding an amino acid fragment described herein. Nucleic acid fragments can encode a specific domain or site described herein or fragments thereof, particularly fragments thereof which are at least 150 amino acids in length. Fragments also include nucleic acid sequences corresponding to specific amino acid sequences described above or fragments thereof. Nucleic acid fragments should not to be construed as encompassing those fragments that may have been disclosed prior to the invention.

A nucleic acid fragment can include a sequence corresponding to a domain, region, or functional site described herein. A nucleic acid fragment can also include one or more domain, region, or functional site described herein. Thus, for example, the nucleic acid fragment can include a beta-ketoacyl synthase. In a preferred embodiment the fragment is at least, 50, 100, 200, 300, 400, 500, 600, 700, 900, 1000, 1100, 1200, 1300, 1400, or 1489 base pairs in length.

Alternatively, a nucleic acid molecule that is a fragment of a beta-ketoacyl synthase nucleotide sequence of the present invention comprises a nucleotide sequence consisting of nucleotides 1-100, 100-200, 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900, 900-1000, 100-1100, 1100-1200, 1200-1300, 1300-1400, or 1400-1489 of SEQ ID NO:1 or nucleotides 1-100, 100-200, 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900, 900-1000, 100-1100, 1100-1200, 1200-1300, 1300-1380 of SEQ ID NO:3.

Beta-ketoacyl synthase probes and primers are provided. Typically a probe/primer is an isolated or purified oligonucleotide. The oligonucleotide typically includes a region of nucleotide sequence that hybridizes under stringent conditions to at least about 7, 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 consecutive nucleotides of a sense or antisense sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

In a preferred embodiment the nucleic acid is a probe which is at least 5 or 10, and less than 200, more preferably less than 100, or less than 50, base pairs in length. It should be identical, or differ by 1, or less than in 5 or 10 bases, from a sequence disclosed herein. If alignment is needed for this comparison the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.

A probe or primer can be derived from the sense or anti-sense strand of a nucleic acid which encodes a beta-ketoacyl synthase domain.

In another embodiment a set of primers is provided, e.g., primers suitable for use in a PCR, which can be used to amplify a selected region of a beta-ketoacyl synthase sequence, e.g., a region described herein. The primers should be at least 5, 10, or 50 base pairs in length and less than 100, or less than 200, base pairs in length. The primers should be identical, or differs by one base from a sequence disclosed herein or from a naturally occurring variant. E.g., primers suitable for amplifying all or a portion of any of the following regions are provided: a beta-ketoacyl synthase active site (e.g., amino acid residues 200 to 216 of SEQ ID NO:2).

A nucleic acid fragment can encode an epitope bearing region of a polypeptide described herein.

A nucleic acid fragment encoding a "biologically active portion of a beta-ketoacyl synthase polypeptide" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1 or the nucleotide sequence of the DNA insert of

the plasmid deposited with ATCC as Accession Number _____, which encodes a polypeptide having a beta-ketoacyl synthase biological activity (e.g., the biological activities of the beta-ketoacyl synthase proteins as described herein), expressing the encoded portion of the beta-ketoacyl synthase protein (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the beta-ketoacyl synthase protein. For example, a nucleic acid fragment encoding a biologically active portion of includes a beta-ketoacyl synthase domain (e.g., about amino acid residues 44-459 of SEQ ID NO:2). A nucleic acid fragment encoding a biologically active portion of a beta-ketoacyl synthase polypeptide, may comprise a nucleotide sequence which is greater than 300 -1400 or more nucleotides in length.

In preferred embodiments, nucleic acid fragments include a nucleotide sequence which is about 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, or 1489 nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

Beta-Ketoacyl Synthase Nucleic Acid Variants

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. Such differences can be due to degeneracy of the genetic code (and result in a nucleic acid which encodes the same beta-ketoacyl synthase proteins as those encoded by the nucleotide sequence disclosed herein.

Nucleic acids of the invention can be chosen for having codons, which are preferred, or non preferred, for a particular expression system. E.g., the nucleic acid can be one in which at least one colon, at preferably at least 10%, or 20% of the codons has been altered such that the sequence is optimized for expression in E. coli, yeast, human, insect, or CHO cells.

Nucleic acid variants can be naturally occurring, such as allelic variants (same locus), homologs (different locus), and orthologs (different organism) or can be non-naturally occurring. Non-naturally occurring variants can be made by mutagenesis

techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product).

In a preferred embodiment, the nucleic acid differs from that of SEQ ID NO:1 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, e.g., as follows: by at least one but less than 10, 20, 30, or 40 nucleotides; at least one but less than 1%, 5%, 10% or 20% of the in the subject nucleic acid. If necessary for this analysis the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.

Orthologs, homologs, and allelic variants can be identified using methods known in the art. These variants comprise a nucleotide sequence encoding a polypeptide that is at least about 45%, 55%, 65%, preferably about 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:2. Variants also include polypeptides encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____, or polypeptides encoded by a nucleic acid molecule that hybridizes to the nucleic acid molecule of SEQ ID NO:1, or a complement thereof, under stringent conditions. In another embodiment, a variant of an isolated polypeptide of the present invention differs, by at least 1, but less than 5, 10, 20, 50, or 100 amino acid residues from the sequence shown in SEQ ID NO:2. If alignment is needed for this comparison the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences. Such variants generally retain the functional activity of the beta-ketoacyl synthase proteins of the invention. Variants include polypeptides that differ in amino acid sequence due to natural allelic variation or mutagenesis. Nucleic acid molecules corresponding to orthologs, homologs, and allelic variants of the beta-ketoacyl synthase cDNAs of the invention can further be isolated by mapping to the same chromosome or locus as the beta-ketoacyl synthase gene. Preferred variants include those that are correlated with fatty acid synthase activity.

Allelic variants of beta-ketoacyl synthase, e.g., human beta-ketoacyl synthase, include both functional and non-functional proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the beta-ketoacyl synthase protein within a population that maintain the ability to couple acyl and malonyl groups to yield beta-keto acyl derivatives. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2, or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein. Non-functional allelic variants are naturally-occurring amino acid sequence variants of the beta-ketoacyl synthase, e.g., human beta-ketoacyl synthase, protein within a population that do not have the ability to attach an acyl chain to a lipid precursor. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion, or premature truncation of the amino acid sequence of SEQ ID NO:2, or a substitution, insertion, or deletion in critical residues or critical regions of the protein.

Moreover, nucleic acid molecules encoding other beta-ketoacyl synthase family members and, thus, which have a nucleotide sequence which differs from the beta-ketoacyl synthase sequences of SEQ ID NO:1 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention.

Antisense Nucleic Acid Molecules, Ribozymes and Modified Beta-Ketoacyl Synthase Nucleic Acid Molecules

In another aspect, the invention features, an isolated nucleic acid molecule which is antisense to beta-ketoacyl synthase. An "antisense" nucleic acid can include a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. The antisense nucleic acid can be complementary to an entire beta-ketoacyl synthase coding strand, or to only a portion thereof. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding beta-ketoacyl synthase (e.g., the 5' and 3' untranslated regions).

An antisense nucleic acid can be designed such that it is complementary to the entire coding region of beta-ketoacyl synthase mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of beta-ketoacyl synthase mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of beta-ketoacyl synthase mRNA, e.g., between the -10 and +10 regions of the target gene nucleotide sequence of interest. An antisense oligonucleotide can be, for example, about 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or more nucleotides in length.

An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. The antisense nucleic acid also can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject (e.g., by direct injection at a tissue site), or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a beta-ketoacyl synthase protein to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient

intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier et al., (1987) *Nucleic Acids. Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al., (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al., (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. A ribozyme having specificity for a beta-ketoacyl synthase-encoding nucleic acid can include one or more sequences complementary to the nucleotide sequence of a beta-ketoacyl synthase cDNA disclosed herein (i.e., SEQ ID NO:1 and a sequence having known catalytic sequence responsible for mRNA cleavage (see U.S. Pat. No. 5,093,246 or Haselhoff and Gerlach, (1988) *Nature* 334:585-591). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a beta-ketoacyl synthase-encoding mRNA. See, e.g., Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742. Alternatively, beta-ketoacyl synthase mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Beta-ketoacyl synthase gene expression can be inhibited by targeting nucleotide sequences complementary o the regulatory region of the beta-ketoacyl synthase (e.g., the beta-ketoacyl synthase promoter and/or enhancers) to form triple helical structures that prevent transcription of the beta-ketoacyl synthase gene in target cells. See generally, Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene et al., (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14(12):807-15. The potential sequences that can be targeted for triple helix formation

can be increased by creating a so-called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

The invention also provides detectably labeled oligonucleotide primer and probe molecules. Typically, such labels are chemiluminescent, fluorescent, radioactive, or colorimetric.

A beta-ketoacyl synthase nucleic acid molecule can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup et al., (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acid" or "PNA" refers to a nucleic acid mimic, e.g., a DNA mimic, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of a PNA can allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al., (1996) *supra*; Perry-O'Keefe et al., *Proc. Natl. Acad. Sci.* 93: 14670-675.

PNAs of beta-ketoacyl synthase nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of beta-ketoacyl synthase nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (e.g., S1 nucleases (Hyrup (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup et al., (1996) *supra*; Perry-O'Keefe *supra*).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating

transport across the cell membrane (see, e.g., Letsinger et al., (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre et al., (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, e.g., Krol et al., (1988) *Bio-Techniques* 6:958-976) or intercalating agents. (See, e.g., Zon, (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

The invention also includes molecular beacon oligonucleotide primer and probe molecules having at least one region which is complementary to a beta-ketoacyl synthase nucleic acid of the invention, two complementary regions one having a fluorophore and one a quencher such that the molecular beacon is useful for quantitating the presence of the beta-ketoacyl synthase nucleic acid of the invention in a sample. Molecular beacon nucleic acids are described, for example, in Lizardi et al., U.S. Patent No. 5,854,033; Nazarenko et al., U.S. Patent No. 5,866,336, and Livak et al., U.S. Patent 5,876,930.

Isolated Beta-Ketoacyl Synthase Polypeptides

In another aspect, the invention features, an isolated beta-ketoacyl synthase protein, or fragment, e.g., a biologically active portion, for use as immunogens or antigens to raise or test (or more generally to bind) anti-beta-ketoacyl synthase antibodies. Beta-ketoacyl synthase protein can be isolated from cells or tissue sources using standard protein purification techniques. Beta-ketoacyl synthase protein or fragments thereof can be produced by recombinant DNA techniques or synthesized chemically.

Polypeptides of the invention include those which arise as a result of the existence of multiple genes, alternative transcription events, alternative RNA splicing events, and alternative translational and postranslational events. The polypeptide can be expressed in systems, e.g., cultured cells, which result in substantially the same postranslational modifications present when expressed the polypeptide is expressed

in a native cell, or in systems which result in the alteration or omission of postranslational modifications, e.g., glycosylation or cleavage, present when expressed in a native cell.

In a preferred embodiment, a beta-ketoacyl synthase polypeptide has one or more of the following characteristics:

(i) it is involved in a condensation reaction between acyl and malonyl groups to yield beta-ketoacyl derivatives;

(ii) it has a molecular weight, e.g., a deduced molecular weight, amino acid composition or other physical characteristic of the polypeptide of SEQ ID NO:2;

(iii) it has an overall sequence similarity of at least 50%, preferably at least 60%, more preferably at least 70, 80, 90, or 95%, with a polypeptide of SEQ ID NO:2;

(iv) it has a beta-ketoacyl synthase active site which preferably has an overall sequence similarity of about 70%, 80%, 90% or 95% with amino acid residues 200-216 of SEQ ID NO:2;

In a preferred embodiment the beta-ketoacyl synthase protein, or fragment thereof, differs from the corresponding sequence in SEQ ID NO:2. In one embodiment it differs by at least one but by less than 15, 10 or 5 amino acid residues. In another it differs from the corresponding sequence in SEQ ID NO:2 by at least one residue but less than 20%, 15%, 10% or 5% of the residues in it differ from the corresponding sequence in SEQ ID NO:2. (If this comparison requires alignment the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.) The differences are, preferably, differences or changes at a non-essential residue or a conservative substitution. In a preferred embodiment the differences are not in the fatty acid synthase domain. In another preferred embodiment one or more differences are in non-active site residues, e.g. outside of the fatty acid synthase domain.

Other embodiments include a protein that contain one or more changes in amino acid sequence, e.g., a change in an amino acid residue which is not essential for activity. Such beta-ketoacyl synthase proteins differ in amino acid sequence from SEQ ID NO:2, yet retain biological activity.

In one embodiment, a biologically active portion of a beta-ketoacyl synthase protein includes an fatty acid synthase domain. In another embodiment, a biologically active portion of a beta-ketoacyl synthase protein includes a domain. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native beta-ketoacyl synthase protein.

In a preferred embodiment, the beta-ketoacyl synthase protein has an amino acid sequence shown in SEQ ID NO:2. In other embodiments, the beta-ketoacyl synthase protein is substantially identical to SEQ ID NO:2. In yet another embodiment, the beta-ketoacyl synthase protein is substantially identical to SEQ ID NO:2 and retains the functional activity of the protein of SEQ ID NO:2, as described in detail above. Accordingly, in another embodiment, the beta-ketoacyl synthase protein is a protein which includes an amino acid sequence at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more identical to SEQ ID NO:2.

Beta-Ketoacyl Synthase Chimeric or Fusion Proteins

In another aspect, the invention provides beta-ketoacyl synthase chimeric or fusion proteins. As used herein, a beta-ketoacyl synthase "chimeric protein" or "fusion protein" includes a beta-ketoacyl synthase polypeptide linked to a non-beta-ketoacyl synthase polypeptide. A "non-beta-ketoacyl synthase polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the beta-ketoacyl synthase protein, e.g., a protein which is different from the beta-ketoacyl synthase protein and which is derived from the same or a different organism. The beta-ketoacyl synthase polypeptide of the fusion protein can correspond to all or a portion e.g., a fragment described herein of a beta-ketoacyl synthase amino acid sequence. In a preferred embodiment, a beta-ketoacyl synthase fusion protein includes at least one (or two) biologically active portion of a beta-ketoacyl synthase protein. The non-beta-ketoacyl synthase polypeptide can be fused to the N-terminus or C-terminus of the beta-ketoacyl synthase polypeptide.

The fusion protein can include a moiety which has a high affinity for a ligand. For example, the fusion protein can be a GST-beta-ketoacyl synthase fusion protein in

which the beta-ketoacyl synthase sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant beta-ketoacyl synthase. Alternatively, the fusion protein can be a beta-ketoacyl synthase protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of beta-ketoacyl synthase can be increased through use of a heterologous signal sequence.

Fusion proteins can include all or a part of a serum protein, e.g., an IgG constant region, or human serum albumin.

The beta-ketoacyl synthase fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The beta-ketoacyl synthase fusion proteins can be used to affect the bioavailability of a beta-ketoacyl synthase substrate. Beta-ketoacyl synthase fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a beta-ketoacyl synthase protein; (ii) mis-regulation of the beta-ketoacyl synthase gene; and (iii) aberrant post-translational modification of a beta-ketoacyl synthase protein.

Moreover, the beta-ketoacyl synthase-fusion proteins of the invention can be used as immunogens to produce anti-beta-ketoacyl synthase antibodies in a subject, to purify beta-ketoacyl synthase ligands and in screening assays to identify molecules which inhibit the interaction of beta-ketoacyl synthase with a beta-ketoacyl synthase substrate.

Expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A beta-ketoacyl synthase-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the beta-ketoacyl synthase protein.

Variants of Beta-Ketoacyl Synthase Proteins

In another aspect, the invention also features a variant of a beta-ketoacyl synthase polypeptide, e.g., which functions as an agonist (mimetics) or as an antagonist. Variants of the beta-ketoacyl synthase proteins can be generated by mutagenesis, e.g., discrete point mutation, the insertion or deletion of sequences or the

truncation of a beta-ketoacyl synthase protein. An agonist of the beta-ketoacyl synthase proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a beta-ketoacyl synthase protein. An antagonist of a beta-ketoacyl synthase protein can inhibit one or more of the activities of the naturally occurring form of the beta-ketoacyl synthase protein by, for example, competitively modulating a beta-ketoacyl synthase-mediated activity of a beta-ketoacyl synthase protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. Preferably, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the beta-ketoacyl synthase protein.

Variants of a beta-ketoacyl synthase protein can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of a beta-ketoacyl synthase protein for agonist or antagonist activity.

Libraries of fragments e.g., N terminal, C terminal, or internal fragments, of a beta-ketoacyl synthase protein coding sequence can be used to generate a variegated population of fragments for screening and subsequent selection of variants of a beta-ketoacyl synthase protein.

Variants in which a cysteine residues is added or deleted or in which a residue which is glycosylated is added or deleted are particularly preferred.

Methods for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify beta-ketoacyl synthase variants (Arkin and Yourvan, (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave et al., (1993) *Protein Engineering* 6(3):327-331).

Cell based assays can be exploited to analyze a variegated beta-ketoacyl synthase library. For example, a library of expression vectors can be transfected into a cell line, e.g., a cell line, which ordinarily responds to beta-ketoacyl synthase in a substrate-dependent manner. The transfected cells are then contacted with beta-

ketoacyl synthase and the effect of the expression of the mutant on signaling by the beta-ketoacyl synthase substrate can be detected, e.g., by measuring fatty acid synthase activity. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the beta-ketoacyl synthase substrate, and the individual clones further characterized.

In another aspect, the invention features a method of making a beta-ketoacyl synthase polypeptide, e.g., a peptide having a non-wild type activity, e.g., an antagonist, agonist, or super agonist of a naturally occurring beta-ketoacyl synthase polypeptide, e.g., a naturally occurring beta-ketoacyl synthase polypeptide. The method includes: altering the sequence of a beta-ketoacyl synthase polypeptide, e.g., altering the sequence, e.g., by substitution or deletion of one or more residues of a non-conserved region, a domain or residue disclosed herein, and testing the altered polypeptide for the desired activity.

In another aspect, the invention features a method of making a fragment or analog of a beta-ketoacyl synthase polypeptide a biological activity of a naturally occurring beta-ketoacyl synthase polypeptide. The method includes: altering the sequence, e.g., by substitution or deletion of one or more residues, of a beta-ketoacyl synthase polypeptide, e.g., altering the sequence of a non-conserved region, or a domain or residue described herein, and testing the altered polypeptide for the desired activity.

Anti-Beta-Ketoacyl Synthase Antibodies

In another aspect, the invention provides an anti-beta-ketoacyl synthase antibody. The term "antibody" as used herein refers to an immunoglobulin molecule or immunologically active portion thereof, i.e., an antigen-binding portion. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin.

The antibody can be a polyclonal, monoclonal, recombinant, e.g., a chimeric or humanized, fully human, non-human, e.g., murine, or single chain antibody. In a

preferred embodiment it has effector function and can fix complement. The antibody can be coupled to a toxin or imaging agent.

A full-length beta-ketoacyl synthase protein or, antigenic peptide fragment of beta-ketoacyl synthase can be used as an immunogen or can be used to identify anti-beta-ketoacyl synthase antibodies made with other immunogens, e.g., cells, membrane preparations, and the like. The antigenic peptide of beta-ketoacyl synthase should include at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2 and encompasses an epitope of beta-ketoacyl synthase. Preferably, the antigenic peptide includes at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

In a preferred embodiment the antibody fails to bind an Fc receptor, e.g. it is a type which does not support Fc receptor binding or has been modified, e.g., by deletion or other mutation, such that is does not have a functional Fc receptor binding region.

Preferred epitopes encompassed by the antigenic peptide are regions of beta-ketoacyl synthase are located on the surface of the protein, e.g., hydrophilic regions, as well as regions with high antigenicity. For example, an Emini surface probability analysis of the human beta-ketoacyl synthase protein sequence can be used to indicate the regions that have a particularly high probability of being localized to the surface of the beta-ketoacyl synthase protein and are thus likely to constitute surface residues useful for targeting antibody production.

In a preferred embodiment the antibody binds an epitope on any domain or region on beta-ketoacyl synthase proteins described herein.

Chimeric, humanized, but most preferably, completely human antibodies are desirable for applications which include repeated administration, e.g., therapeutic treatment (and some diagnostic applications) of human patients.

The anti-beta-ketoacyl synthase antibody can be a single chain antibody. A single-chain antibody (scFV) may be engineered (see, for example, Colcher et al., June 30, 1999, *Ann. NY Acad. Sci.,* 880:263-80; and Reiter, Feb. 1996, *Clin. Cancer Res.* 2(2):245-52). The single chain antibody can be dimerized or multimerized to

generate multivalent antibodies having specificities for different epitopes of the same target beta-ketoacyl synthase protein.

An anti-beta-ketoacyl synthase antibody (e.g., monoclonal antibody) can be used to isolate beta-ketoacyl synthase by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, an anti-beta-ketoacyl synthase antibody can be used to detect beta-ketoacyl synthase protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the protein. Anti-beta-ketoacyl synthase antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance (i.e., antibody labeling). Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

Recombinant Expression Vectors, Host Cells and Genetically Engineered Cells

In another aspect, the invention includes, vectors, preferably expression vectors, containing a nucleic acid encoding a polypeptide described herein. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked and can include a plasmid, cosmid or viral vector. The vector can be capable of autonomous replication or it can integrate into a host DNA. Viral vectors include, e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses.

A vector can include a beta-ketoacyl synthase nucleic acid in a form suitable for expression of the nucleic acid in a host cell. Preferably the recombinant expression vector includes one or more regulatory sequences operatively linked to the nucleic acid sequence to be expressed. The term "regulatory sequence" includes promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence, as well as tissue-specific regulatory and/or inducible sequences. The design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or polypeptides, including fusion proteins or polypeptides, encoded by nucleic acids as described herein (e.g., beta-ketoacyl synthase proteins, mutant forms of beta-ketoacyl synthase proteins, fusion proteins, and the like).

It is further recognized that the nucleic acid sequences of the invention can be altered to contain codons, which are preferred, or non preferred, for a particular expression system. For example, the nucleic acid can be one in which at least one altered codon, and preferably at least 10%, or 20% of the codons have been altered such that the sequence is optimized for expression in *E. coli*, yeast, human, insect, or CHO cells. Methods for determining such codon usage are well known in the art.

The recombinant expression vectors of the invention can be designed for expression of beta-ketoacyl synthase proteins in prokaryotic or eukaryotic cells. For example, polypeptides of the invention can be expressed in E. coli, insect cells (e.g., using baculovirus expression vectors), yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in E. coli with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein.

Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith *et al.*, (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be used in beta-ketoacyl synthase activity assays, (e.g., direct assays or competitive assays described in detail below), or to generate antibodies specific for beta-ketoacyl synthase proteins. In a preferred embodiment, a fusion protein expressed in a retroviral expression vector of the present invention can be used to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (e.g., six (6) weeks).

To maximize recombinant protein expression in *E. coli* is to express the protein in host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman (1990), *Gene Expression Technology: Methods in Enzymology* 185:119-128, Academic Press, San Diego, California). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada et al., (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

The beta-ketoacyl synthase expression vector can be a yeast expression vector, a vector for expression in insect cells, e.g., a baculovirus expression vector or a vector suitable for expression in mammalian cells.

When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al., (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton, (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore, (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji et al., (1983) *Cell* 33:729-740; Queen and Baltimore, (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle, (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund et al., (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example, the murine hox promoters (Kessel and Gruss, (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman, (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. Regulatory sequences (e.g., viral promoters and/or enhancers) operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the constitutive, tissue specific or cell type specific expression of antisense RNA in a variety of cell types. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus. For a discussion of the regulation of gene expression using antisense genes see Weintraub et al., Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, (1986),Vol. 1(1).

Another aspect the invention provides a host cell which includes a nucleic acid molecule described herein, e.g., a beta-ketoacyl synthase nucleic acid molecule within

a recombinant expression vector or a beta-ketoacyl synthase nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. Such terms refer not only to the particular subject cell but rather also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a beta-ketoacyl synthase protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into host cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation

A host cell of the invention can be used to produce (i.e., express) a beta-ketoacyl synthase protein. Accordingly, the invention further provides methods for producing a beta-ketoacyl synthase protein using the host cells of the invention. In one embodiment, the method includes culturing the host cell of the invention (into which a recombinant expression vector encoding a beta-ketoacyl synthase protein has been introduced) in a suitable medium such that a beta-ketoacyl synthase protein is produced. In another embodiment, the method further includes isolating a beta-ketoacyl synthase protein from the medium or the host cell.

In another aspect, the invention features, a cell or purified preparation of cells which include a beta-ketoacyl synthase transgene, or which otherwise misexpress beta-ketoacyl synthase. The cell preparation can consist of human or non-human cells, e.g., rodent cells, e.g., mouse or rat cells, rabbit cells, or pig cells. In preferred embodiments, the cell or cells include a beta-ketoacyl synthase transgene, e.g., a

heterologous form of a beta-ketoacyl synthase, e.g., a gene derived from humans (in the case of a non-human cell). The beta-ketoacyl synthase transgene can be misexpressed, e.g., overexpressed or underexpressed. In other preferred embodiments, the cell or cells include a gene which misexpress an endogenous beta-ketoacyl synthase, e.g., a gene the expression of which is disrupted, e.g., a knockout. Such cells can serve as a model for studying disorders which are related to mutated or mis-expressed beta-ketoacyl synthase alleles or for use in drug screening.

In another aspect, the invention features, a human cell, e.g., a hematopoietic stem cell, transformed with nucleic acid which encodes a subject beta-ketoacyl synthase polypeptide.

Also provided are cells or a purified preparation thereof, e.g., human cells, in which an endogenous beta-ketoacyl synthase is under the control of a regulatory sequence that does not normally control the expression of the endogenous beta-ketoacyl synthase gene. The expression characteristics of an endogenous gene within a cell, e.g., a cell line or microorganism, can be modified by inserting a heterologous DNA regulatory element into the genome of the cell such that the inserted regulatory element is operably linked to the endogenous beta-ketoacyl synthase gene. For example, an endogenous beta-ketoacyl synthase gene, e.g., a gene which is "transcriptionally silent," e.g., not normally expressed, or expressed only at very low levels, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell. Techniques such as targeted homologous recombinations, can be used to insert the heterologous DNA as described in, e.g., Chappel, US 5,272,071; WO 91/06667, published on May 16, 1991.

Transgenic Animals

The invention provides non-human transgenic animals. Such animals are useful for studying the function and/or activity of a beta-ketoacyl synthase protein and for identifying and/or evaluating modulators of beta-ketoacyl synthase activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of

the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA or a rearrangement, e.g., a deletion of endogenous chromosomal DNA, which preferably is integrated into or occurs in the genome of the cells of a transgenic animal. A transgene can direct the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal, other transgenes, e.g., a knockout, reduce expression. Thus, a transgenic animal can be one in which an endogenous beta-ketoacyl synthase gene has been altered by, e.g., by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a transgene of the invention to direct expression of a beta-ketoacyl synthase protein to particular cells. A transgenic founder animal can be identified based upon the presence of a beta-ketoacyl synthase transgene in its genome and/or expression of beta-ketoacyl synthase mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a beta-ketoacyl synthase protein can further be bred to other transgenic animals carrying other transgenes.

Beta-ketoacyl synthase proteins or polypeptides can be expressed in transgenic animals or plants, e.g., a nucleic acid encoding the protein or polypeptide can be introduced into the genome of an animal. In preferred embodiments the nucleic acid is placed under the control of a tissue specific promoter, e.g., a milk or egg specific promoter, and recovered from the milk or eggs produced by the animal. Suitable animals are mice, pigs, cows, goats, and sheep.

The invention also includes a population of cells from a transgenic animal, as discussed herein.

Uses

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (e.g., therapeutic and prophylactic). The isolated nucleic acid molecules of the invention can be used, for example, to express a beta-ketoacyl synthase protein (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect a beta-ketoacyl synthase mRNA (e.g., in a biological sample) or a genetic alteration in a beta-ketoacyl synthase gene, and to modulate beta-ketoacyl synthase activity, as described further below. The beta-ketoacyl synthase proteins can be used to treat disorders characterized by insufficient or excessive production of a beta-ketoacyl synthase substrate or production of beta-ketoacyl synthase inhibitors. In addition, the beta-ketoacyl synthase proteins can be used to screen for naturally occurring beta-ketoacyl synthase substrates, to screen for drugs or compounds which modulate beta-ketoacyl synthase activity, as well as to treat disorders characterized by insufficient or excessive production of beta-ketoacyl synthase protein or production of beta-ketoacyl synthase protein forms which have decreased, aberrant or unwanted activity compared to beta-ketoacyl synthase wild-type protein. Such disorders include those characterized by aberrant signaling or aberrant, e.g., hyperproliferative, cell growth. Moreover, the anti-beta-ketoacyl synthase antibodies of the invention can be used to detect and isolate beta-ketoacyl synthase proteins, regulate the bioavailability of beta-ketoacyl synthase proteins, and modulate beta-ketoacyl synthase activity.

A method of evaluating a compound for the ability to interact with, e.g., bind, a subject beta-ketoacyl synthase polypeptide is provided. The method includes: contacting the compound with the subject beta-ketoacyl synthase polypeptide; and evaluating ability of the compound to interact with, e.g., to bind or form a complex with the subject beta-ketoacyl synthase polypeptide. This method can be performed in vitro, e.g., in a cell free system, or *in vivo*, e.g., in a two-hybrid interaction trap assay. This method can be used to identify naturally occurring molecules which

interact with subject beta-ketoacyl synthase polypeptide. It can also be used to find natural or synthetic inhibitors of subject beta-ketoacyl synthase polypeptide. Screening methods are discussed in more detail below.

Screening Assays:

The invention provides methods (also referred to herein as "screening assays") for identifying modulators, i.e., candidate or test compounds or agents (e.g., proteins, peptides, peptidomimetics, peptoids, small molecules or other drugs) which bind to beta-ketoacyl synthase proteins, have a stimulatory or inhibitory effect on, for example, beta-ketoacyl synthase expression or beta-ketoacyl synthase activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of a beta-ketoacyl synthase substrate. Compounds thus identified can be used to modulate the activity of target gene products (e.g., beta-ketoacyl synthase genes) in a therapeutic protocol, to elaborate the biological function of the target gene product, or to identify compounds that disrupt normal target gene interactions.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a beta-ketoacyl synthase protein or polypeptide or a biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a beta-ketoacyl synthase protein or polypeptide or a biologically active portion thereof.

The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries [libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive] (see, e.g., Zuckermann et al., *J. Med. Chem.* 1994, 37: 2678-85); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library and peptoid library approaches are limited to peptide libraries, while the other four approaches are applicable to peptide,

non-peptide oligomer or small molecule libraries of compounds (Lam (1997), *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb et al., (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann et al., (1994). *J. Med. Chem.* 37:2678; Cho et al., (1993) *Science* 261:1303; Carrell et al., (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell et al., (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop et al., (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (e.g., Houghten, (1992) *Biotechniques* 13:412-421), or on beads (Lam, (1991) *Nature* 354:82-84), chips (Fodor, (1993) *Nature* 364:555-556), bacteria or spores (Ladner, United States Patent No. 5,223,409), plasmids (Cull et al., (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or on phage (Scott and Smith, (1990) *Science* 249:386-390); (Devlin, (1990) *Science* 249:404-406); (Cwirla et al., (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici, (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a beta-ketoacyl synthase protein or biologically active portion thereof is contacted with a test compound, and the ability of the test compound to modulate beta-ketoacyl synthase activity is determined. Determining the ability of the test compound to modulate beta-ketoacyl synthase activity can be accomplished by monitoring, for example, the formation of beta-keto acyl derivatives (Wakil *et al.* (1983) *Annu. Rev. Biochem. 52*: 537-579). The cell, for example, can be of mammalian origin, e.g., human. Cell homogenates, or fractions, preferably membrane containing fractions, can also be tested.

The ability of the test compound to modulate beta-ketoacyl synthase binding to a compound, e.g., a beta-ketoacyl synthase substrate, or to bind to beta-ketoacyl synthase can also be evaluated. This can be accomplished, for example, by coupling the compound, e.g., the substrate, with a radioisotope or enzymatic label such that binding of the compound, e.g., the substrate, to beta-ketoacyl synthase can be determined by detecting the labeled compound, e.g., substrate, in a complex. Alternatively, beta-ketoacyl synthase could be coupled with a radioisotope or

enzymatic label to monitor the ability of a test compound to modulate beta-ketoacyl synthase binding to a beta-ketoacyl synthase substrate in a complex. For example, compounds (e.g., beta-ketoacyl synthase substrates) can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

The ability of a compound (e.g., a beta-ketoacyl synthase substrate) to interact with beta-ketoacyl synthase with or without the labeling of any of the interactants can be evaluated. For example, a microphysiometer can be used to detect the interaction of a compound with beta-ketoacyl synthase without the labeling of either the compound or the beta-ketoacyl synthase. McConnell et al., (1992) *Science* 257:1906-1912. As used herein, a "microphysiometer" (e.g., Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and beta-ketoacyl synthase.

In yet another embodiment, a cell-free assay is provided in which a beta-ketoacyl synthase protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the beta-ketoacyl synthase protein or biologically active portion thereof is evaluated. Preferred biologically active portions of the beta-ketoacyl synthase proteins to be used in assays of the present invention include fragments which participate in interactions with non-beta-ketoacyl synthase molecules, e.g., fragments with high surface probability scores.

Soluble and/or membrane-bound forms of isolated proteins (e.g., beta-ketoacyl synthase proteins or biologically active portions thereof) can be used in the cell-free assays of the invention. When membrane-bound forms of the protein are used, it may be desirable to utilize a solubilizing agent. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide,

Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)$_n$, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate.

Cell-free assays involve preparing a reaction mixture of the target gene protein and the test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus forming a complex that can be removed and/or detected.

In one embodiment, assays are performed where the ability of an agent to block fatty acid synthase activity within a cell is evaluated.

The interaction between two molecules can also be detected, e.g., using fluorescence energy transfer (FET) (see, for example, Lakowicz *et al.*, U.S. Patent No. 5,631,169; Stavrianopoulos, et al., U.S. Patent No. 4,868,103). A fluorophore label on the first, 'donor' molecule is selected such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second, 'acceptor' molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the 'donor' protein molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the 'acceptor' molecule label may be differentiated from that of the 'donor'. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, the spatial relationship between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the 'acceptor' molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (e.g., using a fluorimeter).

In another embodiment, determining the ability of the beta-ketoacyl synthase protein to bind to a target molecule can be accomplished using real-time Biomolecular Interaction Analysis (BIA) (see, e.g., Sjolander and Urbaniczky (1991), *Anal. Chem.* 63:2338-2345 and Szabo et al., (1995) *Curr. Opin. Struct. Biol.* 5:699-705). "Surface plasmon resonance" or "BIA" detects biospecific interactions in real time, without labeling any of the interactants (e.g., BIAcore). Changes in the mass at the binding

surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

In one embodiment, the target gene product or the test substance is anchored onto a solid phase. The target gene product/test compound complexes anchored on the solid phase can be detected at the end of the reaction. Preferably, the target gene product can be anchored onto a solid surface, and the test compound, (which is not anchored), can be labeled, either directly or indirectly, with detectable labels discussed herein.

It may be desirable to immobilize either beta-ketoacyl synthase, an anti-beta-ketoacyl synthase antibody or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a beta-ketoacyl synthase protein, or interaction of a beta-ketoacyl synthase protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/beta-ketoacyl synthase fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or beta-ketoacyl synthase protein, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of beta-ketoacyl synthase binding or activity determined using standard techniques.

Other techniques for immobilizing either a beta-ketoacyl synthase protein or a target molecule on matrices include using conjugation of biotin and streptavidin. Biotinylated beta-ketoacyl synthase protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical).

In order to conduct the assay, the non-immobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (e.g., by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the immobilized component (the antibody, in turn, can be directly labeled or indirectly labeled with, e.g., a labeled anti-Ig antibody).

In one embodiment, this assay is performed utilizing antibodies reactive with beta-ketoacyl synthase protein or target molecules but which do not interfere with binding of the beta-ketoacyl synthase protein to its target molecule. Such antibodies can be derivatized to the wells of the plate, and unbound target or beta-ketoacyl synthase protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the beta-ketoacyl synthase protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the beta-ketoacyl synthase protein or target molecule.

Alternatively, cell free assays can be conducted in a liquid phase. In such an assay, the reaction products are separated from unreacted components, by any of a number of standard techniques, including but not limited to: differential centrifugation (see, for example, Rivas and Minton, *Trends Biochem Sci* 1993 Aug;18(8):284-7);

chromatography (gel filtration chromatography, ion-exchange chromatography); electrophoresis (see, e.g., Ausubel et al., eds. Current Protocols in Molecular Biology 1999, J. Wiley: New York.); and immunoprecipitation (see, for example, Ausubel et al., eds. Current Protocols in Molecular Biology 1999, J. Wiley: New York). Such resins and chromatographic techniques are known to one skilled in the art (see, e.g., Heegaard, *J Mol. Recognit.* 1998 Winter;11(1-6):141-8; Hage and Tweed, *J. Chromatogr. B Biomed. Sci. Appl.* 1997 Oct 10;699(1-2):499-525). Further, fluorescence energy transfer may also be conveniently utilized, as described herein, to detect binding without further purification of the complex from solution.

In a preferred embodiment, the assay includes contacting the beta-ketoacyl synthase protein or biologically active portion thereof with a known compound which binds beta-ketoacyl synthase to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a beta-ketoacyl synthase protein, wherein determining the ability of the test compound to interact with a beta-ketoacyl synthase protein includes determining the ability of the test compound to preferentially bind to beta-ketoacyl synthase or biologically active portion thereof, or to modulate the activity of a target molecule, as compared to the known compound.

The target gene products of the invention can, *in vivo*, interact with one or more cellular or extracellular macromolecules, such as proteins. For the purposes of this discussion, such cellular and extracellular macromolecules are referred to herein as "binding partners." Compounds that disrupt such interactions can be useful in regulating the activity of the target gene product. Such compounds can include, but are not limited to molecules such as antibodies, peptides, and small molecules. The preferred target genes/products for use in this embodiment are the beta-ketoacyl synthase genes herein identified. In an alternative embodiment, the invention provides methods for determining the ability of the test compound to modulate the activity of a beta-ketoacyl synthase protein through modulation of the activity of a downstream effector of a beta-ketoacyl synthase target molecule. For example, the activity of the effector molecule on an appropriate target can be determined, or the

binding of the effector to an appropriate target can be determined, as previously described.

To identify compounds that interfere with the interaction between the target gene product and its cellular or extracellular binding partner(s), e.g., a substrate, a reaction mixture containing the target gene product and the binding partner is prepared, under conditions and for a time sufficient, to allow the two products to form complex. In order to test an inhibitory agent, the reaction mixture is provided in the presence and absence of the test compound. The test compound can be initially included in the reaction mixture, or can be added at a time subsequent to the addition of the target gene and its cellular or extracellular binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the target gene product and the cellular or extracellular binding partner is then detected. The formation of a complex in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the target gene product and the interactive binding partner. Additionally, complex formation within reaction mixtures containing the test compound and normal target gene product can also be compared to complex formation within reaction mixtures containing the test compound and mutant target gene product. This comparison can be important in those cases wherein it is desirable to identify compounds that disrupt interactions of mutant but not normal target gene products.

These assays can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the target gene product or the binding partner onto a solid phase, and detecting complexes anchored on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the target gene products and the binding partners, e.g., by competition, can be identified by conducting the reaction in the presence of the test substance. Alternatively, test compounds that disrupt preformed complexes, e.g., compounds with higher binding constants that displace

one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are briefly described below.

In a heterogeneous assay system, either the target gene product or the interactive cellular or extracellular binding partner, is anchored onto a solid surface (e.g., a microtiter plate), while the non-anchored species is labeled, either directly or indirectly. The anchored species can be immobilized by non-covalent or covalent attachments. Alternatively, an immobilized antibody specific for the species to be anchored can be used to anchor the species to the solid surface.

In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (e.g., by washing) and any complexes formed will remain immobilized on the solid surface. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, can be directly labeled or indirectly labeled with, e.g., a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds that inhibit complex formation or that disrupt preformed complexes can be detected.

Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds that inhibit complex or that disrupt preformed complexes can be identified.

In an alternate embodiment of the invention, a homogeneous assay can be used. For example, a preformed complex of the target gene product and the interactive cellular or extracellular binding partner product is prepared in that either

the target gene products or their binding partners are labeled, but the signal generated by the label is quenched due to complex formation (see, e.g., U.S. Patent No. 4,109,496 that utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances that disrupt target gene product-binding partner interaction can be identified.

In yet another aspect, the beta-ketoacyl synthase proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos et al., (1993) *Cell* 72:223-232; Madura et al., (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al., (1993) *Biotechniques* 14:920-924; Iwabuchi et al., (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with beta-ketoacyl synthase ("beta-ketoacyl synthase-binding proteins" or "beta-ketoacyl synthase-bp") and are involved in beta-ketoacyl synthase activity. Such beta-ketoacyl synthase-bps can be activators or inhibitors of signals by the beta-ketoacyl synthase proteins or beta-ketoacyl synthase targets as, for example, downstream elements of a beta-ketoacyl synthase-mediated signaling pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a beta-ketoacyl synthase protein is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. (Alternatively the: beta-ketoacyl synthase protein can be the fused to the activator domain.) If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a beta-ketoacyl synthase-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription

factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the beta-ketoacyl synthase protein.

In another embodiment, modulators of beta-ketoacyl synthase expression are identified. For example, a cell or cell free mixture is contacted with a candidate compound and the expression of beta-ketoacyl synthase mRNA or protein evaluated relative to the level of expression of beta-ketoacyl synthase mRNA or protein in the absence of the candidate compound. When expression of beta-ketoacyl synthase mRNA or protein is greater in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of beta-ketoacyl synthase mRNA or protein expression. Alternatively, when expression of beta-ketoacyl synthase mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of beta-ketoacyl synthase mRNA or protein expression. The level of beta-ketoacyl synthase mRNA or protein expression can be determined by methods described herein for detecting beta-ketoacyl synthase mRNA or protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a beta-ketoacyl synthase protein can be confirmed *in vivo*, e.g., in an animal.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein (e.g., a beta-ketoacyl synthase modulating agent, an antisense beta-ketoacyl synthase nucleic acid molecule, a beta-ketoacyl synthase-specific antibody, or a beta-ketoacyl synthase-binding partner) in an appropriate animal model to determine the efficacy, toxicity, side effects, or mechanism of action, of treatment with such an agent. Furthermore, novel agents identified by the above-described screening assays can be used for treatments as described herein.

Detection Assays

Portions or fragments of the nucleic acid sequences identified herein can be used as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome e.g., to locate gene regions associated with genetic disease or to associate beta-ketoacyl synthase with a disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

Chromosome Mapping

The beta-ketoacyl synthase nucleotide sequences or portions thereof can be used to map the location of the beta-ketoacyl synthase genes on a chromosome. This process is called chromosome mapping. Chromosome mapping is useful in correlating the beta-ketoacyl synthase sequences with genes associated with disease.

Briefly, beta-ketoacyl synthase genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the beta-ketoacyl synthase nucleotide sequences. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the beta-ketoacyl synthase sequences will yield an amplified fragment.

A panel of somatic cell hybrids in which each cell line contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, can allow easy mapping of individual genes to specific human chromosomes. (D'Eustachio et al., (1983) *Science* 220:919-924).

Other mapping strategies e.g., in situ hybridization (described in Fan et al., (1990) *Proc. Natl. Acad. Sci. USA*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries can be used to map beta-ketoacyl synthase to a chromosomal location.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in

one step. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma et al., Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland et al., (1987) *Nature,* 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the beta-ketoacyl synthase gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

Tissue Typing

Beta-ketoacyl synthase sequences can be used to identify individuals from biological samples using, e.g., restriction fragment length polymorphism (RFLP). In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, the fragments separated, e.g., in a Southern blot, and probed to yield bands for identification. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can also be used to determine the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the beta-ketoacyl synthase nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it. Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences.

Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 can provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those of nucleotides 58-1437 in SEQ ID NO:1 or nucleotides 1-1380 of SEQ ID NO:3 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from beta-ketoacyl synthase nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using

the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

Use of Partial Beta-Ketoacyl Synthase Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e. another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 (e.g., fragments derived from the noncoding regions of SEQ ID NO:1 having a length of at least 20 bases, preferably at least 30 bases) are particularly appropriate for this use.

The beta-ketoacyl synthase nucleotide sequences described herein can further be used to provide polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, e.g., a tissue containing fatty acid synthase activity. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such beta-ketoacyl synthase probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, e.g., beta-ketoacyl synthase primers or probes can be used to screen tissue culture for contamination (i.e. screen for the presence of a mixture of different types of cells in a culture).

Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual.

Generally, the invention provides, a method of determining if a subject is at risk for a disorder related to a lesion in or the misexpression of a gene which encodes beta-ketoacyl synthase.

Such disorders include, e.g., a disorder associated with the misexpression of beta-ketoacyl synthase, or fatty acid biosynthesis or lipid metabolism related disorder.

The method includes one or more of the following:

detecting, in a tissue of the subject, the presence or absence of a mutation which affects the expression of the beta-ketoacyl synthase gene, or detecting the presence or absence of a mutation in a region which controls the expression of the gene, e.g., a mutation in the 5' control region;

detecting, in a tissue of the subject, the presence or absence of a mutation which alters the structure of the beta-ketoacyl synthase gene;

detecting, in a tissue of the subject, the misexpression of the beta-ketoacyl synthase gene, at the mRNA level, e.g., detecting a non-wild type level of a mRNA ;

detecting, in a tissue of the subject, the misexpression of the gene, at the protein level, e.g., detecting a non-wild type level of a beta-ketoacyl synthase polypeptide.

In preferred embodiments the method includes: ascertaining the existence of at least one of: a deletion of one or more nucleotides from the beta-ketoacyl synthase gene; an insertion of one or more nucleotides into the gene, a point mutation, e.g., a substitution of one or more nucleotides of the gene, a gross chromosomal rearrangement of the gene, e.g., a translocation, inversion, or deletion.

For example, detecting the genetic lesion can include: (i) providing a probe/primer including an oligonucleotide containing a region of nucleotide sequence which hybridizes to a sense or antisense sequence from SEQ ID NO:1 naturally occurring mutants thereof or 5' or 3' flanking sequences naturally associated with the beta-ketoacyl synthase gene; (ii) exposing the probe/primer to nucleic acid of the

tissue; and detecting, by hybridization, e.g., *in situ* hybridization, of the probe/primer to the nucleic acid, the presence or absence of the genetic lesion.

In preferred embodiments detecting the misexpression includes ascertaining the existence of at least one of: an alteration in the level of a messenger RNA transcript of the beta-ketoacyl synthase gene; the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; or a non-wild type level of beta-ketoacyl synthase.

Methods of the invention can be used prenatally or to determine if a subject's offspring will be at risk for a disorder.

In preferred embodiments the method includes determining the structure of a beta-ketoacyl synthase gene, an abnormal structure being indicative of risk for the disorder.

In preferred embodiments the method includes contacting a sample form the subject with an antibody to the beta-ketoacyl synthase protein or a nucleic acid, which hybridizes specifically with the gene. These and other embodiments are discussed below.

Diagnostic and Prognostic Assays

The presence, level, or absence of beta-ketoacyl synthase protein or nucleic acid in a biological sample can be evaluated by obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting beta-ketoacyl synthase protein or nucleic acid (e.g., mRNA, genomic DNA) that encodes beta-ketoacyl synthase protein such that the presence of beta-ketoacyl synthase protein or nucleic acid is detected in the biological sample. The term "biological sample" includes tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. A preferred biological sample is serum. The level of expression of the beta-ketoacyl synthase gene can be measured in a number of ways, including, but not limited to: measuring the mRNA encoded by the beta-ketoacyl synthase genes; measuring the amount of protein encoded by the beta-ketoacyl synthase genes; or measuring the activity of the protein encoded by the beta-ketoacyl synthase genes.

The level of mRNA corresponding to the beta-ketoacyl synthase gene in a cell can be determined both by *in situ* and by *in vitro* formats.

The isolated mRNA can be used in hybridization or amplification assays that include, but are not limited to, Southern or Northern analyses, polymerase chain reaction analyses and probe arrays. One preferred diagnostic method for the detection of mRNA levels involves contacting the isolated mRNA with a nucleic acid molecule (probe) that can hybridize to the mRNA encoded by the gene being detected. The nucleic acid probe can be, for example, a full-length beta-ketoacyl synthase nucleic acid, such as the nucleic acid of SEQ ID NO:1, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, such as nucleotides 58-1437 of SEQ ID NO:1 or nucleotides 1-1380 of SEQ ID NO:3, or an oligonucleotide of at least 7, 15, 30, 50, 100, 250, 500, or 1000 nucleotides in length and sufficient to specifically hybridize under stringent conditions to beta-ketoacyl synthase mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays are described herein.

In one format, mRNA (or cDNA) is immobilized on a surface and contacted with the probes, for example by running the isolated mRNA on an agarose gel and transferring the mRNA from the gel to a membrane, such as nitrocellulose. In an alternative format, the probes are immobilized on a surface and the mRNA (or cDNA) is contacted with the probes, for example, in a two-dimensional gene chip array. A skilled artisan can adapt known mRNA detection methods for use in detecting the level of mRNA encoded by the beta-ketoacyl synthase genes.

The level of mRNA in a sample that is encoded by one of beta-ketoacyl synthase can be evaluated with nucleic acid amplification, e.g., by rtPCR (Mullis, 1987, U.S. Patent No. 4,683,202), ligase chain reaction (Barany, 1991, *Proc. Natl. Acad. Sci. USA* 88:189-193), self sustained sequence replication (Guatelli et al., 1990, *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh et al., 1989, *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi et al., 1988, *Bio/Technology* 6:1197), rolling circle replication (Lizardi et al., U.S. Patent No. 5,854,033) or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques known in the art. As

used herein, amplification primers are defined as being a pair of nucleic acid molecules that can anneal to 5' or 3' regions of a gene (plus and minus strands, respectively, or vice-versa) and contain a short region in between. In general, amplification primers are from about 10 to 30 nucleotides in length and flank a region from about 50 to 200 nucleotides in length. Under appropriate conditions and with appropriate reagents, such primers permit the amplification of a nucleic acid molecule comprising the nucleotide sequence flanked by the primers.

For *in situ* methods, a cell or tissue sample can be prepared/processed and immobilized on a support, typically a glass slide, and then contacted with a probe that can hybridize to mRNA that encodes the beta-ketoacyl synthase gene being analyzed.

In another embodiment, the methods further contacting a control sample with a compound or agent capable of detecting beta-ketoacyl synthase mRNA, or genomic DNA, and comparing the presence of beta-ketoacyl synthase mRNA or genomic DNA in the control sample with the presence of beta-ketoacyl synthase mRNA or genomic DNA in the test sample.

A variety of methods can be used to determine the level of protein encoded by beta-ketoacyl synthase. In general, these methods include contacting an agent that selectively binds to the protein, such as an antibody with a sample, to evaluate the level of protein in the sample. In a preferred embodiment, the antibody bears a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with a detectable substance. Examples of detectable substances are provided herein.

The detection methods can be used to detect beta-ketoacyl synthase protein in a biological sample *in vitro* as well as *in vivo*. *In vitro* techniques for detection of beta-ketoacyl synthase protein include enzyme linked immunosorbent assays (ELISAs), immunoprecipitations, immunofluorescence, enzyme immunoassay (EIA), radioimmunoassay (RIA), and Western blot analysis. *In vivo* techniques for detection

of beta-ketoacyl synthase protein include introducing into a subject a labeled anti-beta-ketoacyl synthase antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In another embodiment, the methods further include contacting the control sample with a compound or agent capable of detecting beta-ketoacyl synthase protein, and comparing the presence of beta-ketoacyl synthase protein in the control sample with the presence of beta-ketoacyl synthase protein in the test sample.

The invention also includes kits for detecting the presence of beta-ketoacyl synthase in a biological sample. For example, the kit can include a compound or agent capable of detecting beta-ketoacyl synthase protein or mRNA in a biological sample; and a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect beta-ketoacyl synthase protein or nucleic acid.

For antibody-based kits, the kit can include: (1) a first antibody (e.g., attached to a solid support) which binds to a polypeptide corresponding to a marker of the invention; and, optionally, (2) a second, different antibody which binds to either the polypeptide or the first antibody and is conjugated to a detectable agent.

For oligonucleotide-based kits, the kit can include: (1) an oligonucleotide, e.g., a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding a polypeptide corresponding to a marker of the invention or (2) a pair of primers useful for amplifying a nucleic acid molecule corresponding to a marker of the invention. The kit can also includes a buffering agent, a preservative, or a protein-stabilizing agent. The kit can also includes components necessary for detecting the detectable agent (e.g., an enzyme or a substrate). The kit can also contain a control sample or a series of control samples which can be assayed and compared to the test sample contained. Each component of the kit can be enclosed within an individual container and all of the various containers can be within a single package, along with instructions for interpreting the results of the assays performed using the kit.

The diagnostic methods described herein can identify subjects having, or at risk of developing, a disease or disorder associated with misexpressed or aberrant or

unwanted beta-ketoacyl synthase expression or activity. As used herein, the term "unwanted" includes an unwanted phenomenon involved in a biological response such as pain or deregulated cell proliferation.

In one embodiment, a disease or disorder associated with aberrant or unwanted beta-ketoacyl synthase expression or activity is identified. A test sample is obtained from a subject and beta-ketoacyl synthase protein or nucleic acid (e.g., mRNA or genomic DNA) is evaluated, wherein the level, e.g., the presence or absence, of beta-ketoacyl synthase protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted beta-ketoacyl synthase expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest, including a biological fluid (e.g., serum), cell sample, or tissue.

The prognostic assays described herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted beta-ketoacyl synthase expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a cellular growth related disorder.

The methods of the invention can also be used to detect genetic alterations in a beta-ketoacyl synthase gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in beta-ketoacyl synthase protein activity or nucleic acid expression, such as a cellular growth related disorder. In preferred embodiments, the methods include detecting, in a sample from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a beta-ketoacyl synthase-protein, or the mis-expression of the beta-ketoacyl synthase gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a beta-ketoacyl synthase gene; 2) an addition of one or more nucleotides to a beta-ketoacyl synthase gene; 3) a substitution of one or more nucleotides of a beta-ketoacyl synthase gene, 4) a chromosomal rearrangement of a

beta-ketoacyl synthase gene; 5) an alteration in the level of a messenger RNA transcript of a beta-ketoacyl synthase gene, 6) aberrant modification of a beta-ketoacyl synthase gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a beta-ketoacyl synthase gene, 8) a non-wild type level of a beta-ketoacyl synthase-protein, 9) allelic loss of a beta-ketoacyl synthase gene, and 10) inappropriate post-translational modification of a beta-ketoacyl synthase-protein.

An alteration can be detected without a probe/primer in a polymerase chain reaction, such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR), the latter of which can be particularly useful for detecting point mutations in the beta-ketoacyl synthase-gene. This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (e.g., genomic, mRNA or both) from the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a beta-ketoacyl synthase gene under conditions such that hybridization and amplification of the beta-ketoacyl synthase-gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli et al., (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh et al., (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi et al., (1988) *Bio-Technology* 6:1197), or other nucleic acid amplification methods, followed by the detection of the amplified molecules using techniques known to those of skill in the art.

In another embodiment, mutations in a beta-ketoacyl synthase gene from a sample cell can be identified by detecting alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined, e.g., by gel electrophoresis and compared. Differences in fragment

length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in beta-ketoacyl synthase can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, two-dimensional arrays, e.g., chip based arrays. Such arrays include a plurality of addresses, each of which is positionally distinguishable from the other. A different probe is located at each address of the plurality. The arrays can have a high density of addresses, e.g., can contain hundreds or thousands of oligonucleotides probes (Cronin et al., (1996) *Human Mutation* 7: 244-255; Kozal et al., (1996) *Nature Medicine* 2:753-759). For example, genetic mutations in beta-ketoacyl synthase can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin et al., *supra.* Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the beta-ketoacyl synthase gene and detect mutations by comparing the sequence of the sample beta-ketoacyl synthase with the corresponding wild-type (control) sequence. Automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry.

Other methods for detecting mutations in the beta-ketoacyl synthase gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers et al., (1985)

Science 230:1242; Cotton et al., (1988) *Proc. Natl. Acad. Sci. USA* 85:4397; Saleeba et al., (1992) *Methods Enzymol.* 217:286-295).

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in beta-ketoacyl synthase cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu et al., (1994) *Carcinogenesis* 15:1657-1662; U.S. Patent No. 5,459,039).

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in beta-ketoacyl synthase genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita et al., (1989) *Proc. Natl. Acad. Sci. USA*: 86:2766, see also Cotton, (1993) *Mutat. Res.* 285:125-144; and Hayashi, (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control beta-ketoacyl synthase nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen et al., (1991) *Trends Genet.* 7:5).

In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers et al., (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the

mobility of control and sample DNA (Rosenbaum and Reissner, (1987) *Biophys. Chem.* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension (Saiki et al., (1986) *Nature* 324:163); Saiki et al., (1989) *Proc. Natl. Acad. Sci. USA* 86:6230).

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs et al., (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner, (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini et al., (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany, (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.,* in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a beta-ketoacyl synthase gene.

Use of Beta-Ketoacyl Synthase Molecules as Surrogate Markers

The beta-ketoacyl synthase molecules of the invention are also useful as markers of disorders or disease states, as markers for precursors of disease states, as markers for predisposition of disease states, as markers of drug activity, or as markers of the pharmacogenomic profile of a subject. Using the methods described herein, the

presence, absence and/or quantity of the beta-ketoacyl synthase molecules of the invention may be detected, and may be correlated with one or more biological states *in vivo*. For example, the beta-ketoacyl synthase molecules of the invention may serve as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states. As used herein, a "surrogate marker" is an objective biochemical marker which correlates with the absence or presence of a disease or disorder, or with the progression of a disease or disorder (*e.g.*, with the presence or absence of a tumor). The presence or quantity of such markers is independent of the disease. Therefore, these markers may serve to indicate whether a particular course of treatment is effective in lessening a disease state or disorder. Surrogate markers are of particular use when the presence or extent of a disease state or disorder is difficult to assess through standard methodologies (*e.g.*, early stage tumors), or when an assessment of disease progression is desired before a potentially dangerous clinical endpoint is reached (*e.g.*, an assessment of cardiovascular disease may be made using cholesterol levels as a surrogate marker, and an analysis of HIV infection may be made using HIV RNA levels as a surrogate marker, well in advance of the undesirable clinical outcomes of myocardial infarction or fully-developed AIDS). Examples of the use of surrogate markers in the art include: Koomen *et al.* (2000) *J. Mass. Spectrom.* 35: 258-264; and James (1994) *AIDS Treatment News Archive* 209.

The beta-ketoacyl synthase molecules of the invention are also useful as pharmacodynamic markers. As used herein, a "pharmacodynamic marker" is an objective biochemical marker which correlates specifically with drug effects. The presence or quantity of a pharmacodynamic marker is not related to the disease state or disorder for which the drug is being administered; therefore, the presence or quantity of the marker is indicative of the presence or activity of the drug in a subject. For example, a pharmacodynamic marker may be indicative of the concentration of the drug in a biological tissue, in that the marker is either expressed or transcribed or not expressed or transcribed in that tissue in relationship to the level of the drug. In this fashion, the distribution or uptake of the drug may be monitored by the pharmacodynamic marker. Similarly, the presence or quantity of the

pharmacodynamic marker may be related to the presence or quantity of the metabolic product of a drug, such that the presence or quantity of the marker is indicative of the relative breakdown rate of the drug *in vivo*. Pharmacodynamic markers are of particular use in increasing the sensitivity of detection of drug effects, particularly when the drug is administered in low doses. Since even a small amount of a drug may be sufficient to activate multiple rounds of marker (e.g., a beta-ketoacyl synthase marker) transcription or expression, the amplified marker may be in a quantity which is more readily detectable than the drug itself. Also, the marker may be more easily detected due to the nature of the marker itself; for example, using the methods described herein, anti-beta-ketoacyl synthase antibodies may be employed in an immune-based detection system for a beta-ketoacyl synthase protein marker, or beta-ketoacyl synthase-specific radiolabeled probes may be used to detect a beta-ketoacyl synthase mRNA marker. Furthermore, the use of a pharmacodynamic marker may offer mechanism-based prediction of risk due to drug treatment beyond the range of possible direct observations. Examples of the use of pharmacodynamic markers in the art include: Matsuda *et al.* US 6,033,862; Hattis *et al.* (1991) *Env. Health Perspect.* 90: 229-238; Schentag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S21-S24; and Nicolau (1999) *Am, J. Health-Syst. Pharm.* 56 Suppl. 3: S16-S20.

The beta-ketoacyl synthase molecules of the invention are also useful as pharmacogenomic markers. As used herein, a "pharmacogenomic marker" is an objective biochemical marker which correlates with a specific clinical drug response or susceptibility in a subject (see, e.g., McLeod *et al.* (1999) *Eur. J. Cancer* 35(12): 1650-1652). The presence or quantity of the pharmacogenomic marker is related to the predicted response of the subject to a specific drug or class of drugs prior to administration of the drug. By assessing the presence or quantity of one or more pharmacogenomic markers in a subject, a drug therapy which is most appropriate for the subject, or which is predicted to have a greater degree of success, may be selected. For example, based on the presence or quantity of RNA, or protein (*e.g.*, beta-ketoacyl synthase protein or RNA) for specific tumor markers in a subject, a drug or course of treatment may be selected that is optimized for the treatment of the specific tumor likely to be present in the subject. Similarly, the presence or absence of a

specific sequence mutation in beta-ketoacyl synthase DNA may correlate beta-ketoacyl synthase drug response. The use of pharmacogenomic markers therefore permits the application of the most appropriate treatment for each subject without having to administer the therapy.

Pharmaceutical Compositions

The nucleic acid and polypeptides, fragments thereof, as well as anti-beta-ketoacyl synthase antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions. Such compositions typically include the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" includes solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the

extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It should be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules, e.g., gelatin capsules. Oral compositions can also be prepared using a fluid carrier for use

as a mouthwash. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as

pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.,* for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds which exhibit high therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

As defined herein, a therapeutically effective amount of protein or polypeptide (i.e., an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The protein or polypeptide can be administered one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

For antibodies, the preferred dosage is 0.1 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg). If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is usually appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (e.g., into the brain). A method for lipidation of antibodies is described by Cruikshank et al., ((1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193).

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics (e.g., peptoids), amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e,. including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having

a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (e.g., about 1microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1microgram per kilogram to about 50 micrograms per kilogram. It is furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. When one or more of these small molecules is to be administered to an animal (e.g., a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

An antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-

dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, .alpha.-interferon, .beta.-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophase colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see e.g., Chen et al., (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g., retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant or unwanted beta-ketoacyl synthase expression or activity. With regards to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how a patient's genes determine his or her response to a drug (e.g., a patient's "drug response phenotype", or "drug response genotype".) Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the beta-ketoacyl synthase molecules of the present invention or beta-ketoacyl synthase modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

"Treatment" is herein defined as the application or administration of a therapeutic agent to a patient, or application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disease, a symptom of disease or a predisposition toward a disease, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect the disease, the symptoms of disease or the predisposition toward disease. A "therapeutic agent" includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides. (See, above.)

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant or unwanted beta-ketoacyl synthase expression or activity, by administering to the subject a beta-ketoacyl synthase or an agent which modulates beta-ketoacyl synthase expression or at least one beta-ketoacyl

synthase activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted beta-ketoacyl synthase expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the beta-ketoacyl synthase aberrance, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of beta-ketoacyl synthase aberrance, for example, a beta-ketoacyl synthase, beta-ketoacyl synthase agonist or beta-ketoacyl synthase antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

It is possible that some beta-ketoacyl synthase disorders can be caused, at least in part, by an abnormal level of gene product, or by the presence of a gene product exhibiting abnormal activity. As such, the reduction in the level and/or activity of such gene products would bring about the amelioration of disorder symptoms.

As discussed, successful treatment of beta-ketoacyl synthase disorders can be brought about by techniques that serve to inhibit the expression or activity of target gene products. For example, compounds, e.g., an agent identified using an assays described above, that proves to exhibit negative modulatory activity, can be used in accordance with the invention to prevent and/or ameliorate symptoms of beta-ketoacyl synthase disorders. Such molecules can include, but are not limited to peptides, phosphopeptides, small organic or inorganic molecules, or antibodies (including, for example, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and FAb, $F(ab')_2$ and FAb expression library fragments, scFV molecules, and epitope-binding fragments thereof).

Further, antisense and ribozyme molecules that inhibit expression of the target gene can also be used in accordance with the invention to reduce the level of target gene expression, thus effectively reducing the level of target gene activity. Still further, triple helix molecules can be utilized in reducing the level of target gene activity. Antisense, ribozyme and triple helix molecules are discussed above.

It is possible that the use of antisense, ribozyme, and/or triple helix molecules to reduce or inhibit mutant gene expression can also reduce or inhibit the transcription

(triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles, such that the concentration of normal target gene product present can be lower than is necessary for a normal phenotype. In such cases, nucleic acid molecules that encode and express target gene polypeptides exhibiting normal target gene activity can be introduced into cells via gene therapy method. Alternatively, in instances in that the target gene encodes an extracellular protein, it can be preferable to co-administer normal target gene protein into the cell or tissue in order to maintain the requisite level of cellular or tissue target gene activity.

Another method by which nucleic acid molecules may be utilized in treating or preventing a disease characterized by beta-ketoacyl synthase expression is through the use of aptamer molecules specific for beta-ketoacyl synthase protein. Aptamers are nucleic acid molecules having a tertiary structure which permits them to specifically bind to protein ligands (see, e.g., Osborne, et al., *Curr. Opin. Chem. Biol.* 1997, 1(1): 5-9; and Patel, *Curr. Opin. Chem. Biol.* 1997 Jun;1(1):32-46). Since nucleic acid molecules may in many cases be more conveniently introduced into target cells than therapeutic protein molecules may be, aptamers offer a method by which beta-ketoacyl synthase protein activity may be specifically decreased without the introduction of drugs or other molecules which may have pluripotent effects.

Antibodies can be generated that are both specific for target gene product and that reduce target gene product activity. Such antibodies may, therefore, by administered in instances whereby negative modulatory techniques are appropriate for the treatment of beta-ketoacyl synthase disorders. For a description of antibodies, see the Antibody section above.

In circumstances wherein injection of an animal or a human subject with a beta-ketoacyl synthase protein or epitope for stimulating antibody production is harmful to the subject, it is possible to generate an immune response against beta-ketoacyl synthase through the use of anti-idiotypic antibodies (see, for example, Herlyn, *Ann. Med.* 1999;31(1):66-78; and Bhattacharya-Chatterjee, and Foon, *Cancer Treat. Res.* 1998;94:51-68). If an anti-idiotypic antibody is introduced into a mammal or human subject, it should stimulate the production of anti-anti-idiotypic antibodies, which should be specific to the beta-ketoacyl synthase protein. Vaccines directed to a

disease characterized by beta-ketoacyl synthase expression may also be generated in this fashion.

In instances where the target antigen is intracellular and whole antibodies are used, internalizing antibodies may be preferred. Lipofectin or liposomes can be used to deliver the antibody or a fragment of the Fab region that binds to the target antigen into cells. Where fragments of the antibody are used, the smallest inhibitory fragment that binds to the target antigen is preferred. For example, peptides having an amino acid sequence corresponding to the Fv region of the antibody can be used. Alternatively, single chain neutralizing antibodies that bind to intracellular target antigens can also be administered. Such single chain antibodies can be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population (see e.g., Marasco et al., (1993, *Proc. Natl. Acad. Sci. USA* 90:7889-7893).

The identified compounds that inhibit target gene expression, synthesis and/or activity can be administered to a patient at therapeutically effective doses to prevent, treat or ameliorate beta-ketoacyl synthase disorders. A therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of the disorders.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds that exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects can be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage can vary within this range depending upon the dosage

form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (i.e., the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma can be measured, for example, by high performance liquid chromatography.

Another example of determination of effective dose for an individual is the ability to directly assay levels of "free" and "bound" compound in the serum of the test subject. Such assays may utilize antibody mimics and/or "biosensors" that have been created through molecular imprinting techniques. The compound which is able to modulate beta-ketoacyl synthase activity is used as a template, or "imprinting molecule", to spatially organize polymerizable monomers prior to their polymerization with catalytic reagents. The subsequent removal of the imprinted molecule leaves a polymer matrix which contains a repeated "negative image" of the compound and is able to selectively rebind the molecule under biological assay conditions. A detailed review of this technique can be seen in Ansell et al., (1996) *Current Opinion in Biotechnology* 7:89-94 and in Shea, (1994) *Trends in Polymer Science* 2:166-173. Such "imprinted" affinity matrixes are amenable to ligand-binding assays, whereby the immobilized monoclonal antibody component is replaced by an appropriately imprinted matrix. An example of the use of such matrixes in this way can be seen in Vlatakis et al., (1993) *Nature* 361:645-647. Through the use of isotope-labeling, the "free" concentration of compound which modulates the expression or activity of beta-ketoacyl synthase can be readily monitored and used in calculations of IC_{50}.

Such "imprinted" affinity matrixes can also be designed to include fluorescent groups whose photon-emitting properties measurably change upon local and selective binding of target compound. These changes can be readily assayed in real time using appropriate fiberoptic devices, in turn allowing the dose in a test subject to be quickly

optimized based on its individual IC_{50}. A rudimentary example of such a "biosensor" is discussed in Kriz et al., (1995) *Analytical Chemistry* 67:2142-2144.

Another aspect of the invention pertains to methods of modulating beta-ketoacyl synthase expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a beta-ketoacyl synthase or agent that modulates one or more of the activities of beta-ketoacyl synthase protein activity associated with the cell. An agent that modulates beta-ketoacyl synthase protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a beta-ketoacyl synthase protein (e.g., a beta-ketoacyl synthase substrate or receptor), a beta-ketoacyl synthase antibody, a beta-ketoacyl synthase agonist or antagonist, a peptidomimetic of a beta-ketoacyl synthase agonist or antagonist, or other small molecule.

In one embodiment, the agent stimulates one or beta-ketoacyl synthase activities. Examples of such stimulatory agents include active beta-ketoacyl synthase protein and a nucleic acid molecule encoding beta-ketoacyl synthase. In another embodiment, the agent inhibits one or more beta-ketoacyl synthase activities. Examples of such inhibitory agents include antisense beta-ketoacyl synthase nucleic acid molecules, anti-beta-ketoacyl synthase antibodies, and beta-ketoacyl synthaseinhibitors. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a beta-ketoacyl synthase protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) beta-ketoacyl synthase expression or activity. In another embodiment, the method involves administering a beta-ketoacyl synthase protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted beta-ketoacyl synthase expression or activity.

Stimulation of beta-ketoacyl synthase activity is desirable in situations in which beta-ketoacyl synthase is abnormally downregulated and/or in which increased beta-ketoacyl synthase activity is likely to have a beneficial effect. For example, stimulation of beta-ketoacyl synthase activity is desirable in situations in which a beta-ketoacyl synthase is downregulated and/or in which increased beta-ketoacyl synthase activity is likely to have a beneficial effect. Likewise, inhibition of beta-ketoacyl synthase activity is desirable in situations in which beta-ketoacyl synthase is abnormally upregulated and/or in which decreased beta-ketoacyl synthase activity is likely to have a beneficial effect.

The beta-ketoacyl synthase molecules can act as novel diagnostic targets and therapeutic agents for controlling one or more of cellular proliferative and/or differentiative disorders, cardiovascular disorders, as described above, as well as disorders associated with lipid metabolism, hematopoietic disorders, liver disorders, viral diseases, pain or metabolic disorders.

Examples of hematopoietic disorders include, but are not limited to, autoimmune diseases (including, for example, diabetes mellitus, arthritis (including rheumatoid arthritis, juvenile rheumatoid arthritis, osteoarthritis, psoriatic arthritis), multiple sclerosis, encephalomyelitis, myasthenia gravis, systemic lupus erythematosis, autoimmune thyroiditis, dermatitis (including atopic dermatitis and eczematous dermatitis), psoriasis, Sjögren's Syndrome, Crohn's disease, aphthous ulcer, iritis, conjunctivitis, keratoconjunctivitis, ulcerative colitis, asthma, allergic asthma, cutaneous lupus erythematosus, scleroderma, vaginitis, proctitis, drug eruptions,leprosy reversal reactions, erythema nodosum leprosum, autoimmune uveitis, allergic encephalomyelitis, acute necrotizing hemorrhagic encephalopathy, idiopathic bilateral progressive sensorineural hearing loss, aplastic anemia, pure red cell anemia, idiopathic thrombocytopenia, polychondritis, Wegener's granulomatosis, chronic active hepatitis, Stevens-Johnson syndrome, idiopathic sprue, lichen planus, Graves' disease, sarcoidosis, primary biliary cirrhosis, uveitis posterior, and interstitial lung fibrosis), graft-versus-host disease, cases of transplantation, and allergy such as, atopic allergy.

Disorders which may be treated or diagnosed by methods described herein include, but are not limited to, disorders associated with an accumulation in the liver of fibrous tissue, such as that resulting from an imbalance between production and degradation of the extracellular matrix accompanied by the collapse and condensation of preexisting fibers. The methods described herein can be used to diagnose or treat hepatocellular necrosis or injury induced by a wide variety of agents including processes which disturb homeostasis, such as an inflammatory process, tissue damage resulting from toxic injury or altered hepatic blood flow, and infections (e.g., bacterial, viral and parasitic). For example, the methods can be used for the early detection of hepatic injury, such as portal hypertension or hepatic fibrosis. In addition, the methods can be employed to detect liver fibrosis attributed to inborn errors of metabolism, for example, fibrosis resulting from a storage disorder such as Gaucher's disease (lipid abnormalities) or a glycogen storage disease, A1-antitrypsin deficiency; a disorder mediating the accumulation (e.g., storage) of an exogenous substance, for example, hemochromatosis (iron-overload syndrome) and copper storage diseases (Wilson's disease), disorders resulting in the accumulation of a toxic metabolite (e.g., tyrosinemia, fructosemia and galactosemia) and peroxisomal disorders (e.g., Zellweger syndrome). Additionally, the methods described herein may be useful for the early detection and treatment of liver injury associated with the administration of various chemicals or drugs, such as for example, methotrexate, isonizaid, oxyphenisatin, methyldopa, chlorpromazine, tolbutamide or alcohol, or which represents a hepatic manifestation of a vascular disorder such as obstruction of either the intrahepatic or extrahepatic bile flow or an alteration in hepatic circulation resulting, for example, from chronic heart failure, veno-occlusive disease, portal vein thrombosis or Budd-Chiari syndrome.

Additionally, beta-ketoacyl synthase molecules may play an important role in the etiology of certain viral diseases, including but not limited to, Hepatitis B, Hepatitis C and Herpes Simplex Virus (HSV). Modulators of beta-ketoacyl synthase activity could be used to control viral diseases. The modulators can be used in the treatment and/or diagnosis of viral infected tissue or virus-associated tissue fibrosis, especially liver and liver fibrosis. Also, beta-ketoacyl synthase modulators can be

used in the treatment and/or diagnosis of virus-associated carcinoma, especially hepatocellular cancer.

Additionally, beta-ketoacyl synthase may play an important role in the regulation of metabolism or pain disorders. Diseases of metabolic imbalance include, but are not limited to, obesity, anorexia nervosa, cachexia, lipid disorders, and diabetes. Examples of pain disorders include, but are not limited to, pain response elicited during various forms of tissue injury, e.g., inflammation, infection, and ischemia, usually referred to as hyperalgesia (described in, for example, Fields (1987) *Pain*, New York:McGraw-Hill); pain associated with muscoloskeletal disorders, e.g., joint pain; tooth pain; headaches; pain associated with surgery.

Pharmacogenomics

The beta-ketoacyl synthase molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on beta-ketoacyl synthase activity (e.g., beta-ketoacyl synthase gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) beta-ketoacyl synthase-associated disorders (e.g., cellular growth related disorders) associated with aberrant or unwanted beta-ketoacyl synthase activity. In conjunction with such treatment, pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a beta-ketoacyl synthase molecule or beta-ketoacyl synthase modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a beta-ketoacyl synthase molecule or beta-ketoacyl synthase modulator.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum et al. (1996) *Clin. Exp. Pharmacol. Physiol.*

23(10-11) :983-985 and Linder et al. (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (e.g., a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high-resolution map can be generated from a combination of some ten million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drug's target is known (e.g., a beta-ketoacyl synthase protein of the present

invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (e.g., a beta-ketoacyl synthase molecule or beta-ketoacyl synthase modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment of an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a beta-ketoacyl synthase molecule or beta-ketoacyl synthase modulator, such as a modulator identified by one of the exemplary screening assays described herein.

The present invention further provides methods for identifying new agents, or combinations, that are based on identifying agents that modulate the activity of one or more of the gene products encoded by one or more of the beta-ketoacyl synthase genes of the present invention, wherein these products may be associated with resistance of the cells to a therapeutic agent. Specifically, the activity of the proteins encoded by the beta-ketoacyl synthase genes of the present invention can be used as a basis for identifying agents for overcoming agent resistance. By blocking the activity of one or more of the resistance proteins, target cells, e.g., cancer cells, will become sensitive to treatment with an agent that the unmodified target cells were resistant to.

Monitoring the influence of agents (e.g., drugs) on the expression or activity of a beta-ketoacyl synthase protein can be applied in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase beta-ketoacyl synthase gene expression, protein levels, or upregulate beta-ketoacyl synthase activity, can be monitored in clinical trials of subjects exhibiting decreased beta-ketoacyl synthase gene expression, protein levels, or downregulated beta-ketoacyl synthase activity. Alternatively, the effectiveness of an agent

determined by a screening assay to decrease beta-ketoacyl synthase gene expression, protein levels, or downregulate beta-ketoacyl synthase activity, can be monitored in clinical trials of subjects exhibiting increased beta-ketoacyl synthase gene expression, protein levels, or upregulated beta-ketoacyl synthase activity. In such clinical trials, the expression or activity of a beta-ketoacyl synthase gene, and preferably, other genes that have been implicated in, for example, a beta-ketoacyl synthase-associated disorder can be used as a "read out" or markers of the phenotype of a particular cell.

Other Embodiments

In another aspect, the invention features, a method of analyzing a plurality of capture probes. The method can be used, e.g., to analyze gene expression. The method includes: providing a two dimensional array having a plurality of addresses, each address of the plurality being positionally distinguishable from each other address of the plurality, and each address of the plurality having a unique capture probe, e.g., a nucleic acid or peptide sequence; contacting the array with a beta-ketoacyl synthase, preferably purified, nucleic acid, preferably purified, polypeptide, preferably purified, or antibody, and thereby evaluating the plurality of capture probes. Binding, e.g., in the case of a nucleic acid, hybridization with a capture probe at an address of the plurality, is detected, e.g., by signal generated from a label attached to the beta-ketoacyl synthase nucleic acid, polypeptide, or antibody.

The capture probes can be a set of nucleic acids from a selected sample, e.g., a sample of nucleic acids derived from a control or non-stimulated tissue or cell.

The method can include contacting the beta-ketoacyl synthase nucleic acid, polypeptide, or antibody with a first array having a plurality of capture probes and a second array having a different plurality of capture probes. The results of each hybridization can be compared, e.g., to analyze differences in expression between a first and second sample. The first plurality of capture probes can be from a control sample, e.g., a wild type, normal, or non-diseased, non-stimulated, sample, e.g., a biological fluid, tissue, or cell sample. The second plurality of capture probes can be from an experimental sample, e.g., a mutant type, at risk, disease-state or disorder-state, or stimulated, sample, e.g., a biological fluid, tissue, or cell sample.

The plurality of capture probes can be a plurality of nucleic acid probes each of which specifically hybridizes, with an allele of beta-ketoacyl synthase. Such methods can be used to diagnose a subject, e.g., to evaluate risk for a disease or disorder, to evaluate suitability of a selected treatment for a subject, to evaluate whether a subject has a disease or disorder. beta-ketoacyl synthase is associated with fatty acid synthase activity, thus it is useful for disorders associated with abnormal lipid metabolism.

The method can be used to detect SNPs, as described above.

In another aspect, the invention features, a method of analyzing a plurality of probes. The method is useful, e.g., for analyzing gene expression. The method includes: providing a two dimensional array having a plurality of addresses, each address of the plurality being positionally distinguishable from each other address of the plurality having a unique capture probe, e.g., wherein the capture probes are from a cell or subject which express or mis express beta-ketoacyl synthase or from a cell or subject in which a beta-ketoacyl synthase mediated response has been elicited, e.g., by contact of the cell with beta-ketoacyl synthase nucleic acid or protein, or administration to the cell or subject beta-ketoacyl synthase nucleic acid or protein; contacting the array with one or more inquiry probe, wherein an inquiry probe can be a nucleic acid, polypeptide, or antibody (which is preferably other than beta-ketoacyl synthase nucleic acid, polypeptide, or antibody); providing a two dimensional array having a plurality of addresses, each address of the plurality being positionally distinguishable from each other address of the plurality, and each address of the plurality having a unique capture probe, e.g., wherein the capture probes are from a cell or subject which does not express beta-ketoacyl synthase (or does not express as highly as in the case of the beta-ketoacyl synthase positive plurality of capture probes) or from a cell or subject which in which a beta-ketoacyl synthase mediated response has not been elicited (or has been elicited to a lesser extent than in the first sample); contacting the array with one or more inquiry probes (which is preferably other than a beta-ketoacyl synthase nucleic acid, polypeptide, or antibody), and thereby evaluating the plurality of capture probes. Binding, e.g., in the case of a nucleic acid,

hybridization with a capture probe at an address of the plurality, is detected, e.g., by signal generated from a label attached to the nucleic acid, polypeptide, or antibody.

In another aspect, the invention features, a method of analyzing beta-ketoacyl synthase, e.g., analyzing structure, function, or relatedness to other nucleic acid or amino acid sequences. The method includes: providing a beta-ketoacyl synthase nucleic acid or amino acid sequence; comparing the beta-ketoacyl synthase sequence with one or more preferably a plurality of sequences from a collection of sequences, e.g., a nucleic acid or protein sequence database; to thereby analyze beta-ketoacyl synthase.

Preferred databases include GenBank™. The method can include evaluating the sequence identity between a beta-ketoacyl synthase sequence and a database sequence. The method can be performed by accessing the database at a second site, e.g., over the internet.

In another aspect, the invention features, a set of oligonucleotides, useful, e.g., for identifying SNP's, or identifying specific alleles of beta-ketoacyl synthase. The set includes a plurality of oligonucleotides, each of which has a different nucleotide at an interrogation position, e.g., an SNP or the site of a mutation. In a preferred embodiment, the oligonucleotides of the plurality identical in sequence with one another (except for differences in length). The oligonucleotides can be provided with different labels, such that an oligonucleotides which hybridizes to one allele provides a signal that is distinguishable from an oligonucleotides which hybridizes to a second allele.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are incorporated herein by reference.

EXAMPLES

Example 1: Identification and Characterization of Human Beta-Ketoacyl synthase cDNAs

The human beta-ketoacyl synthase sequence (Figure 1A-B; SEQ ID NO:1), which is approximately 1489 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1380 nucleotides (nucleotides 58-1437 of SEQ ID NO:1; nucleotides 1-1380 of SEQ ID NO:3). The coding sequence encodes a 459 amino acid protein (SEQ ID NO:2).

PFAM analysis indicates that the 46619 polypeptide shares a high degree of sequence similarity with beta-ketoacyl synthase. For general information regarding PFAM identifiers, PS prefix and PF prefix domain identification numbers, refer to Sonnhammer *et al.* (1997) *Protein* 28:405-420 and http//www.psc.edu/general/software/packages/pfam/pfam.html. See also http://pfam.wustl.edu/.

As used herein, the term "beta-ketoacyl synthase domain" includes an amino acid sequence of about 415 amino acid residues in length and having a bit score for the alignment of the sequence to the beta-ketoacyl synthase domain (HMM) of at least 8. Preferably, a beta-ketoacyl synthase domain includes at least about 100-415 amino acids, more preferably about 130-300 amino acid residues, or about 160-200 amino acids and has a bit score for the alignment of the sequence to the beta-ketoacyl synthase domain (HMM) of at least 16 or greater. (http://pfam.wustl.edu/). An alignment of the beta-ketoacyl synthase domain (amino acids 44-459 of SEQ ID NO:2) of human beta-ketoacyl synthase polypeptide or protein with a consensus amino acid sequence derived from a hidden Markov model is depicted in Figure 6A-B.

In a preferred embodiment beta-ketoacyl synthase polypeptide or protein has a "beta-ketoacyl synthase domain" or a region which includes at least about 100-415, more preferably about 130-300 or 160-200 amino acid residues and has at least about 60%, 70%, 80%, 90%, 95%, 99%, or 100% sequence identity with an "beta-ketoacyl

synthase domain," e.g., the beta-ketoacyl synthase domain of human beta-ketoacyl synthase polypeptide or protein (e.g., amino acid residues 44-459 of SEQ ID NO:2).

To identify the presence of an "beta-ketoacyl synthase" domain in a 46619-like protein sequence, and make the determination that a polypeptide or protein of interest has a particular profile, the amino acid sequence of the protein can be searched against a database of HMMs (e.g., the Pfam database, release 2.1) using the default parameters (http://www.sanger.ac.uk/Software/Pfam/HMM_search). For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for MILPAT0063 and a score of 15 is the default threshold score for determining a hit. Alternatively, the threshold score for determining a hit can be lowered (e.g., to 8 bits). A description of the Pfam database can be found in Sonhammer *et al.* (1997) *Proteins 28(3)*:405-420 and a detailed description of HMMs can be found, for example, in Gribskov *et al.* (1990) *Meth. Enzymol. 183*:146-159; Gribskov *et al.* (1987) *Proc. Natl. Acad. Sci. USA 84*:4355-4358; Krogh *et al.* (1994) *J. Mol. Biol. 235*:1501-1531; and Stultz *et al.* (1993) *Protein Sci. 2*:305-314, the contents of which are incorporated herein by reference.

In one embodiment, a beta-ketoacyl synthase protein includes at least one transmembrane domain. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 15 amino acid residues in length that spans a phospholipid membrane. More preferably, a transmembrane domain includes about at least 16, 18, 20, 22, or 24 amino acid residues and spans a phospholipid membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an α-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, e.g., leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, http://pfam.wustl.edu/cgi-bin/getdesc?name=7tm-1, and Zagotta et al., (1996) *Annual Rev. Neuronsci.* 19: 235-63, the contents of which are incorporated herein by reference.

In a preferred embodiment, a beta-ketoacyl synthase polypeptide or protein has at least one transmembrane domain or a region which includes at least 16, 18, 20,

22, or 24 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% homology with a "transmembrane domain," e.g., at least one transmembrane domain of human beta-ketoacyl synthase (e.g., amino acid residues 44-63, 144-160, 230-248, or 386-404 of SEQ ID NO:2).

In another embodiment, a beta-ketoacyl synthase polypeptide or protein includes at least one "non-transmembrane domain." As used herein, "non-transmembrane domains" are domains that reside outside of the membrane. When referring to plasma membranes, non-transmembrane domains include extracellular domains (i.e., outside of the cell) and intracellular domains (i.e., within the cell). When referring to membrane-bound proteins found in intracellular organelles (e.g., mitochondria, endoplasmic reticulum, peroxisomes and microsomes), non-transmembrane domains include those domains of the protein that reside in the cytosol (i.e., the cytoplasm), the lumen of the organelle, or the matrix or the intermembrane space (the latter two relate specifically to mitochondria organelles). The C-terminal amino acid residue of a non-transmembrane domain is adjacent to an N-terminal amino acid residue of a transmembrane domain in a naturally occurring beta-ketoacyl synthase polypeptide or protein.

In a preferred embodiment, a beta-ketoacyl synthase polypeptide or protein has a "non-transmembrane domain" or a region which includes at least about 1-130, preferably about 25-100, more preferably about 50-90, and even more preferably about 60-80 amino acid residues, and has at least about 60%, 70% 80% 90% 95%, 99% or 100% sequence identity with a "non-transmembrane domain", e.g., a non-transmembrane domain of human beta-ketoacyl synthase polypeptide or protein (e.g., residues 1-43; 64-143; 161- 229; 249-385; 405-459 of SEQ ID NO:2). Preferably, a non-transmembrane domain is capable of catalytic activity (e.g., ketoacyl synthase).

A non-transmembrane domain located at the N-terminus of a beta-ketoacyl synthase polypeptide or protein is referred to herein as an "N-terminal non-transmembrane domain." As used herein, an "N-terminal non-transmembrane domain" includes an amino acid sequence having about 1-43, preferably about 10-40, more preferably about 20-35, or even more preferably about 25-30 amino acid residues in length and is located outside the boundaries of a membrane. For example,

an N-terminal non-transmembrane domain is located at about amino acid residues 1-43 of SEQ ID NO:2.

Similarly, a non-transmembrane domain located at the C-terminus of a beta-ketoacyl synthase polypeptide or protein is referred to herein as a "C-terminal non-transmembrane domain." As used herein, an "C-terminal non-transmembrane domain" includes an amino acid sequence having about 1-54, preferably about 15-50, preferably about 20-45, more preferably about 25-40 amino acid residues in length and is located outside the boundaries of a membrane. For example, an C-terminal non-transmembrane domain is located at about amino acid residues 405-459 of SEQ ID NO:2.

A beta-ketoacyl synthase polypeptide or protein molecule can further include a signal sequence. As used herein, a "signal sequence" refers to a peptide of about 20-80 amino acid residues in length which occurs at the N-terminus of secretory and integral membrane proteins and which contains a majority of hydrophobic amino acid residues. For example, a signal sequence contains at least about 12-25 amino acid residues, preferably about 30-50 amino acid residues, more preferably about 58 amino acid residues, and has at least about 40-70%, preferably about 50-65%, and more preferably about 55-60% hydrophobic amino acid residues (e.g., alanine, valine, leucine, isoleucine, phenylalanine, tyrosine, tryptophan, or proline). Such a "signal sequence", also referred to in the art as a "signal peptide", serves to direct a protein containing such a sequence to a lipid bilayer. For example, in one embodiment, a beta-ketoacyl synthase polypeptide or protein contains a signal sequence of about amino acids 1-58 of SEQ ID NO:2. The "signal sequence" is cleaved during processing of the mature protein. The mature beta-ketoacyl synthase polypeptide or protein corresponds to amino acids 59-459 of SEQ ID NO:2.

Example 2: Tissue Distribution of Beta-Ketoacyl Synthase mRNA

Expression levels of 46619 in various tissue and cell types were determined by quantitative RT-PCR (Reverse Transcriptase Polymerase Chain Reaction; Taqman® brand PCR kit, Applied Biosystems). The quantitative RT-PCR reactions were performed according to the kit manufacturer's instructions. In the prostate, colon, and

lung, 46619 expression levels were found to be elevated in tumor tissue as compared with normal tissue. See Figure 3. Expression of 46619 was also observed in ovary tumor, normal liver, proliferating human umbilical vein endothelial cells, normoxic colon cancer, hypoxic colon cancer (Figure 4), and breast tumor (Figure 5).

Example 3: Recombinant Expression of Beta-Ketoacyl Synthase in Bacterial Cells

In this example, beta-ketoacyl synthase is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, beta-ketoacyl synthase is fused to GST and this fusion polypeptide is expressed in *E. coli*, e.g., strain PEB199. Expression of the GST-beta-ketoacyl synthase fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

Example 4: Expression of Recombinant Beta-Ketoacyl Synthase Protein in COS Cells

To express the beta-ketoacyl synthase gene in COS cells, the pcDNA/Amp vector by Invitrogen Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire beta-ketoacyl synthase protein and an HA tag (Wilson et al. (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

To construct the plasmid, the beta-ketoacyl synthase DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the beta-ketoacyl synthase

coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the beta-ketoacyl synthase coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, MA). Preferably the two restriction sites chosen are different so that the beta-ketoacyl synthase gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5α, SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the beta-ketoacyl synthase-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook et al., *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the beta-ketoacyl synthase polypeptide is detected by radiolabelling (^{35}S-methionine or ^{35}S-cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labeled for 8 hours with ^{35}S-methionine (or ^{35}S-cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the beta-ketoacyl synthase coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate

restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the beta-ketoacyl synthase polypeptide is detected by radiolabelling and immunoprecipitation using a beta-ketoacyl synthase specific monoclonal antibody.

All publications and patent applications mentioned in the specification are indicative of the level of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

Applicant's or agent's file reference	35800/235514	International application No.	PCT/US01/

INDICATIONS RELATING TO DEPOSITED MICROORGANISM OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13*bis*)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 2, line 28

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depository institution
American Type Culture Collection

Address of depositary institution *(including postal code and country)*
10801 University Boulevard
Manassas, VA 20110-2209 USA

Date of deposit	Accession Number
	PTA-

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

Page 3, lines 1, 5 and 28; page 4, line 3; page 9, line 9; page 22, line 15; page 23, line 6; page 26, line 26; page 27, lines 7, 11, 13, 19 and 30; page 29 lines 7 and 10; page 30, lines 1, 15 and 20; page 31, lines 8 and 18; page 32, line 18; page 68, line 10; page 107, lines 8, 13, 17, 21, 24 and 28; page 108, lines 8 and 12; page 109, lines 2, 5, 11, 28, and 31; page 110, lines 3 and 7.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indicators are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

Accession Number of Deposit and Date of Deposit

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application Sonya D Barnes	☐ This sheet was received with the International Bureau on:
Authorized officer Sonya D. Barnes PCT/Internat'l Appl Processing Div (703) 305-3665	Authorized officer

Form PCT/RO/134 (July 1998)

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule selected from the group consisting of:

a) a nucleic acid molecule comprising a nucleotide sequence having at least 70% sequence identity to the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the cDNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein said nucleotide sequence encodes a polypeptide having biological activity;

b) a nucleic acid molecule comprising a fragment of at least 20 nucleotides of the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the cDNA insert of the plasmid deposited with ATCC as Accession Number _____,

c) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____;

d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____,

e) a nucleic acid molecule which encodes a naturally occurring allelic variant of a biologically active polypeptide comprising the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising the complement of SEQ ID NO:1 or SEQ ID NO:3 under stringent conditions; and

f) a nucleic acid molecule comprising the complement of a), b), c), d), or e).

2. The isolated nucleic acid molecule of claim 1, which is selected from the group consisting of:

a) a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, the nucleotide sequence of the cDNA insert of the plasmid deposited with ATCC as Accession Number _____, or a complement thereof; and

b) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____, or a complement thereof.

3. The nucleic acid molecule of claim 1 further comprising vector nucleic acid sequences.

4. The nucleic acid molecule of claim 1 further comprising nucleic acid sequences encoding a heterologous polypeptide.

5. A host cell which contains the nucleic acid molecule of claim 3.

6. The host cell of claim 5 which is a mammalian host cell.

7. A non-human mammalian host cell containing the nucleic acid molecule of claim 1.

8. An isolated polypeptide selected from the group consisting of:

a) a biologically active polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence having at least 70% identity to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or

the nucleotide sequence of the cDNA insert of the plasmid deposited with ATCC as Accession Number _____;

b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising the complement of SEQ ID NO:1 or SEQ ID NO:3 under stringent conditions; and,

c) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2; and

d) a polypeptide having at least 70% sequence identity to the amino acid sequence SEQ ID NO:2, wherein the polypeptide has biological activity.

9. The isolated polypeptide of claim 8 comprising the amino acid sequence of SEQ ID NO:2.

10. The polypeptide of claim 8 further comprising heterologous amino acid sequences.

11. An antibody which selectively binds to a polypeptide of claim 8.

12. A method for producing a polypeptide selected from the group consisting of:

a) a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____;

b) a polypeptide comprising a fragment of the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____, wherein the

fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____;

c) a biologically active naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC as Accession Number _____, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising the complement of SEQ ID NO:1 or SEQ ID NO:3,

d) a polypeptide having at least 70% sequence identity to the amino acid sequence of SEQ ID NO:2, wherein said polypeptide has biological activity;

comprising culturing the host cell of claim 5 under conditions in which the nucleic acid molecule is expressed.

13. A method for detecting the presence of a polypeptide of claim 8 in a sample, comprising:

a) contacting the sample with a compound which selectively binds to a polypeptide of claim 8; and

b) determining whether the compound binds to the polypeptide in the sample.

14. The method of claim 13, wherein the compound which binds to the polypeptide is an antibody.

15. A kit comprising a compound which selectively binds to a polypeptide of claim 8 and instructions for use.

16. A method for detecting the presence of a nucleic acid molecule of claim 1 in a sample, comprising the steps of:

a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and

b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample.

17. The method of claim 16, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

18. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of claim 1 and instructions for use.

19. A method for identifying a compound which binds to a polypeptide of claim 8 comprising the steps of:

a) contacting a polypeptide, or a cell expressing a polypeptide of claim 8 with a test compound; and

b) determining whether the polypeptide binds to the test compound.

20. The method of claim 19, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

a) detection of binding by direct detecting of test compound/polypeptide binding;

b) detection of binding using a competition binding assay;

c) detection of binding using an assay for 46619-mediated beta-ketoacyl synthase activity.

21. A method for modulating the activity of a polypeptide of claim 8 comprising contacting a polypeptide or a cell expressing a polypeptide of claim 8 with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

22. A method for identifying a compound which modulates the activity of a polypeptide of claim 8, comprising:

a) contacting a polypeptide of claim 8 with a test compound; and

b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound that modulates the activity of the polypeptide.

SUBSTITUTE SHEET (RULE 26)

Input file Fbh46619FL.seq; Output File 46619.trans
Sequence length 1489

```
                                                                  M   S   N   C   L       5
AAGTGTCCGGGGTAGCCCCGTTACAGGAATGTGTTTCTGATCATCTGAATCTTAATC ATG TCC AAC TGC CTG      15

 Q   N   F   L   K   I   T   S   T   R   L   L   C   S   R   L   C   Q   Q   L    25
CAA AAT TTC CTG AAA ATT ACA AGC ACT CGT CTT CTA TGT TCA AGA TTA TGC CAA CAG TTA   75

 R   S   K   R   K   F   F   G   T   V   P   I   S   R   L   H   R   R   V   V    45
AGA AGT AAA AGG AAG TTT TTC GGA ACT GTG CCA ATA TCC AGA TTG CAT AGG CGA GTT GTC  135

 I   T   G   I   G   L   V   T   P   L   G   V   G   T   H   L   V   W   D   R    65
ATT ACA GGC ATT GGC TTA GTG ACT CCT CTT GGT GTT GGA ACT CAC CTG GTT TGG GAT CGT  195

 L   I   G   G   E   S   G   I   V   S   L   V   G   E   E   Y   K   S   I   P    85
CTT ATC GGA GGA GAG AGT GGA ATT GTT TCA CTG GTT GGT GAA GAG TAT AAG AGT ATC CCT  255

 C   S   V   A   A   Y   V   P   R   G   S   D   E   G   Q   F   N   E   Q   N   105
TGC AGT GTT GCT GCT TAT GTG CCA AGA GGT AGT GAT GAA GGT CAG TTC AAT GAA CAA AAC  315

 F   V   S   K   S   D   I   K   S   M   S   S   P   T   I   M   A   I   G   A   125
TTT GTG TCC AAA TCA GAT ATC AAG TCC ATG TCT TCT CCC ACC ATC ATG GCC ATT GGG GCT  375

 A   E   L   A   M   K   D   S   G   W   H   P   Q   S   E   A   D   Q   V   A   145
GCA GAA TTA GCC ATG AAG GAT TCT GGC TGG CAT CCT CAG TCA GAA GCT GAT CAA GTG GCT  435

 T   G   V   A   I   G   M   G   M   I   P   L   E   V   V   S   E   T   A   L   165
ACT GGT GTT GCA ATT GGC ATG GGA ATG ATT CCT CTT GAA GTT GTT TCT GAA ACT GCT TTG  495

 N   F   Q   T   K   G   Y   N   K   V   S   P   F   F   V   P   K   I   L   V   185
AAT TTT CAG ACA AAA GGT TAC AAT AAA GTT AGC CCA TTT TTT GTC CCT AAG ATT CTG GTC  555

 N   M   A   A   G   Q   V   S   I   R   Y   K   L   K   G   P   N   H   A   V   205
AAT ATG GCA GCA GGC CAG GTC AGC ATT CGA TAT AAA CTC AAG GGC CCA AAT CAT GCA GTA  615
```

FIG. 1A.

```
S  T  A  C  T  T  G  A  H  A  V  G  D  S  F  R  F  I  A  H      225
TCC ACA GCC TGT ACC ACA GGA GCT CAT GCT GTG GGA GAC TCA TTT AGA TTT ATA GCC CAT      675

G  D  A  D  V  M  V  A  G  G  T  D  S  C  I  S  P  L  S  L      245
GGT GAT GCT GAT GTG ATG GTG GCT GGA GGT ACA GAT TCT TGT ATT AGC CCT TTA TCT CTT      735

A  G  F  S  R  A  R  A  L  S  T  N  S  D  P  K  L  A  C  R      265
GCT GGG TTT TCC AGA GCC CGG GCT CTG AGC ACA AAC TCA GAT CCC AAG TTG GCA TGT CGA      795

P  F  H  P  K  R  D  G  F  V  M  G  E  G  A  A  V  L  V  L      285
CCA TTT CAT CCA AAG AGA GAT GGT TTT GTA ATG GGA GAA GGT GCA GCT GTG CTG GTG CTG      855

E  E  Y  E  H  A  V  Q  R  R  A  R  I  Y  A  E  V  L  G  Y      305
GAA GAA TAT GAA CAT GCT GTT CAA AGA AGA GCC CGG ATC TAT GCA GAA GTT TTG GGC TAT      915

G  L  S  G  D  A  G  H  I  T  A  P  D  P  E  G  E  G  A  L      325
GGA CTC TCA GGT GAT GCT GGT CAC ATA ACT GCC CCT GAT CCT GAA GGA GAA GGT GCC TTA      975

R  C  M  A  A  A  L  K  D  A  G  V  Q  P  E  E  I  S  Y  I      345
AGG TGT ATG GCT GCT GCT TTA AAA GAT GCA GGT GTG CAG CCT GAG GAG ATA TCC TAT ATC     1035

N  A  H  A  T  S  T  P  L  G  D  A  A  E  N  K  A  I  K  H      365
AAT GCA CAT GCT ACT TCC ACA CCA TTG GGA GAT GCT GCT GAA AAC AAA GCT ATC AAA CAT     1095

L  F  K  D  H  A  Y  A  L  A  V  S  S  T  K  G  A  T  G  H      385
CTC TTC AAA GAC CAT GCA TAT GCC CTT GCA GTT TCC TCA ACT AAG GGA GCA ACA GGA CAT     1155

L  L  G  A  A  G  A  V  E  A  A  F  T  T  L  A  C  Y  Y  Q      405
CTG CTG GGA GCT GCA GGG GCA GTC GAG GCA GCT TTT ACC ACA TTA GCT TGT TAT TAT CAA     1215

K  L  P  P  T  L  N  L  D  C  S  E  P  E  F  D  L  N  Y  V      425
AAA CTA CCA CCT ACT TTA AAC CTG GAT TGT TCG GAA CCA GAA TTT GAT CTC AAC TAT GTT     1275

P  L  K  A  Q  E  W  K  T  E  K  R  F  I  G  L  T  N  S  F      445
CCA CTA AAG GCA CAG GAA TGG AAA ACT GAG AAA AGA TTT ATT GGC CTC ACC AAT TCC TTT     1335

G  F  G  G  T  N  A  T  L  C  I  A  G  L  *                  460
GGT TTT GGT GGT ACT AAT GCA ACA CTT TGT ATT GCT GGA CTG TAG                        1380

AACATATAATTTGTAATTAAATACTGATTTTTAAATGTTAAAAAAAAAAAAA
```

FIG. 1B.

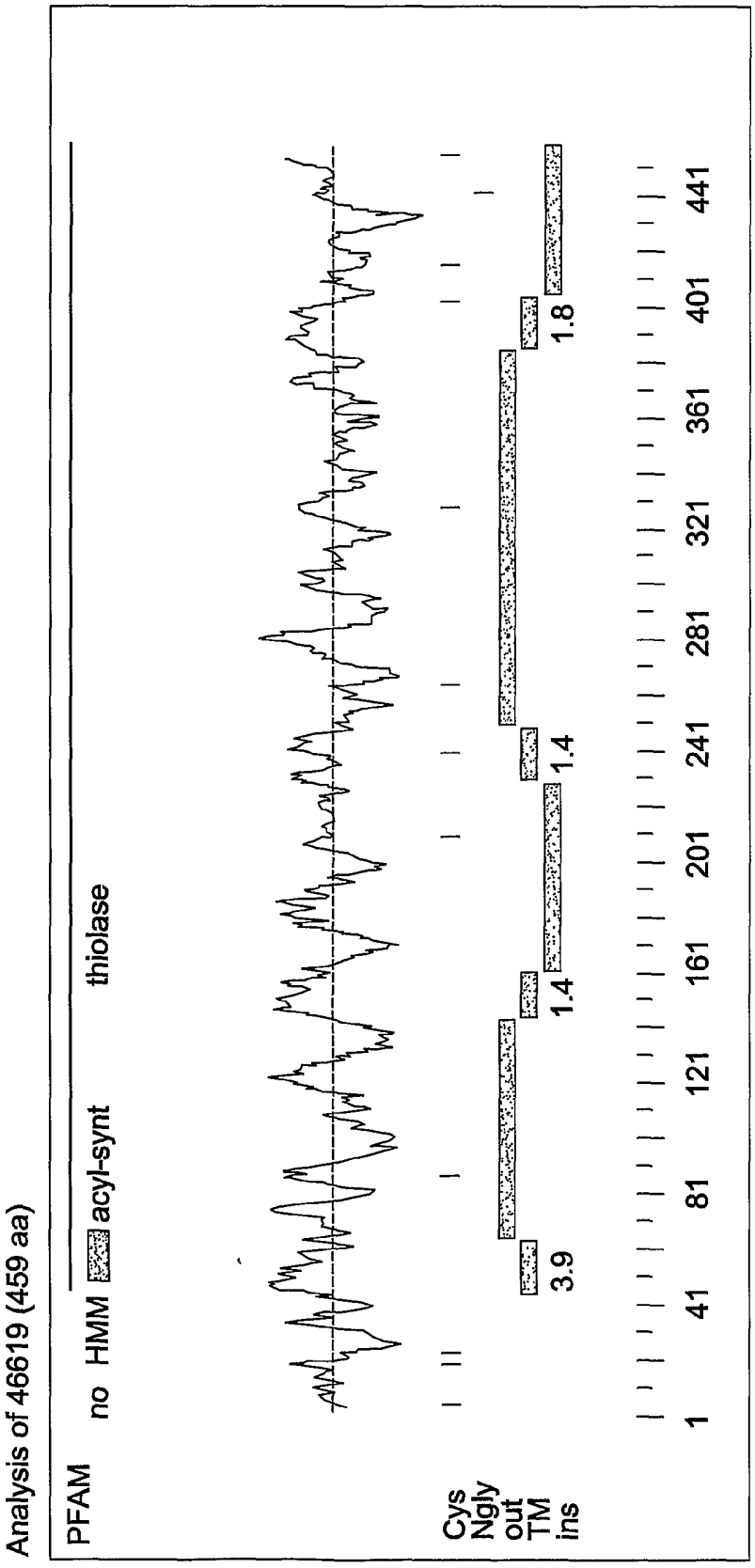
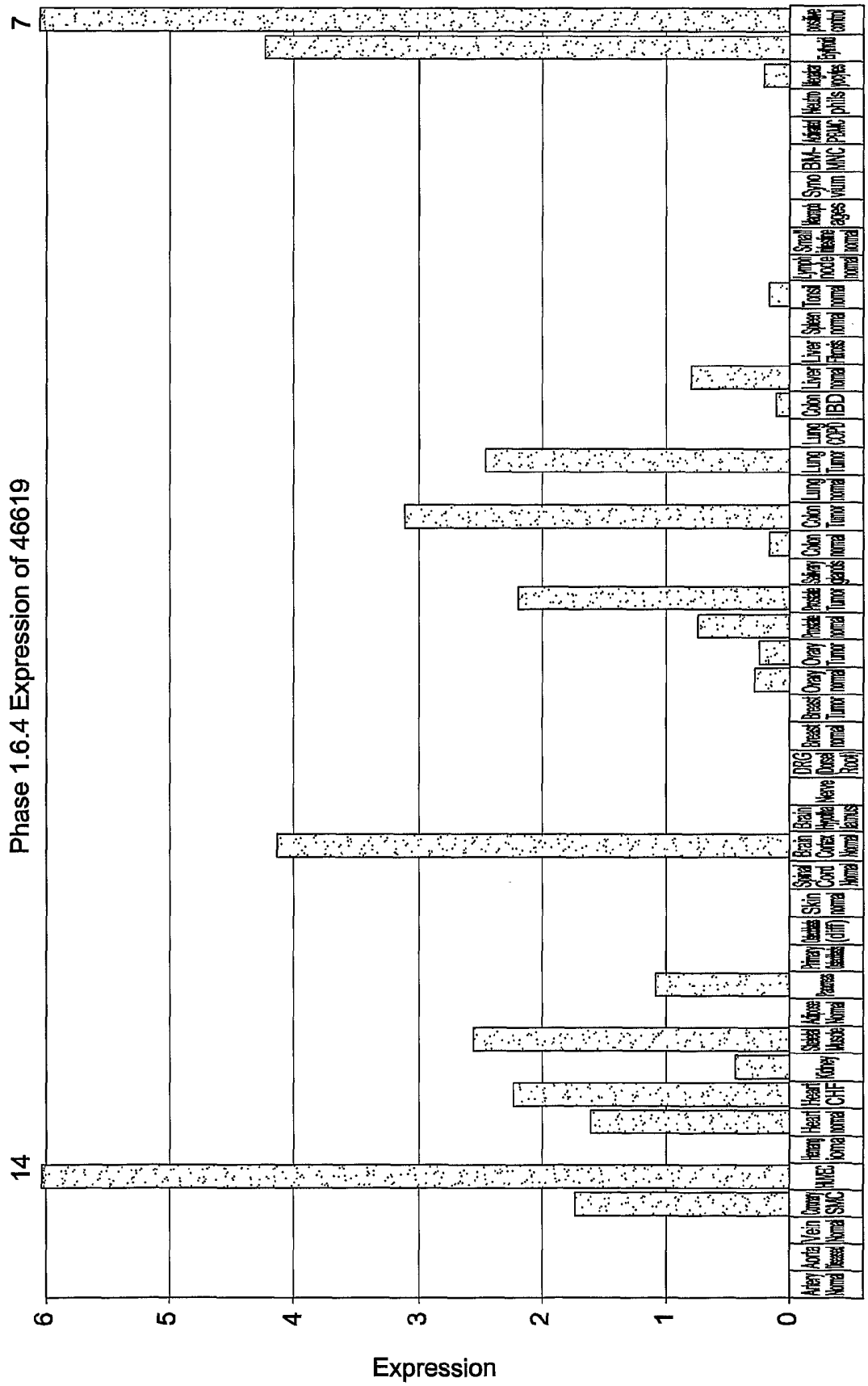

Analysis of 46619 (459 aa)
PFAM
no HMM
acyl-synt
thiolase
Cys
Ngly
out
TM
ins
3.9
1.4
1.4
1.8
1
41
81
121
161
201
241
281
321
361
401
441

FIG. 2.


Phase 1.6.4 Expression of 46619
7
14
6
5
4
3
2
1
0
Expression
Adrenal Normal
Aorta Normal
Vein Normal
Bladder Normal
Cervix Normal
Uterus Normal
Ovary Normal
Testes Normal
Prostate Normal
Salivary gland Normal
Colon Normal
Small Intestine Normal
Spleen Normal
Tonsil Normal
Lymph node Normal
Stomach Normal
Pancreas Normal
Liver Normal
Liver Fibrosis
Colon IBD
Lung COPD
Lung Tumor
Lung Normal
Colon Tumor
Colon Normal
Salivary gland Tumor
Prostate Tumor
Prostate Normal
Ovary Tumor
Ovary Normal
Breast Tumor
Breast Normal
DRG (Dorsal Root)
Nerve
Brain Hypothalamus
Brain Cortex Normal
Spinal Cord Normal
Skin Normal
Testes (diff)
Primary Osteoblasts
Uterus
Adipose Normal
Skeletal Muscle
Kidney
Heart CHF
Heart Normal
Heart
SMC
Coronary
HUVEC
Aorta
Vein
Bladder
Cervix
Uterus
Ovary
Testes
Prostate
Salivary gland
Colon
Small Intestine
Spleen
Tonsil
Lymph node
Stomach
Pancreas
Liver
Liver Fibrosis
Colon IBD
Lung COPD
Lung Tumor
Lung
Colon Tumor
Colon
Salivary gland
Prostate Tumor
Prostate
Ovary Tumor
Ovary
Breast Tumor
Breast
DRG (Dorsal Root)
Nerve
Brain Hypothalamus
Brain Cortex
Spinal Cord
Skin
Testes (diff)
Primary Osteoblasts
Uterus
Adipose
Skeletal Muscle
Kidney
Heart CHF

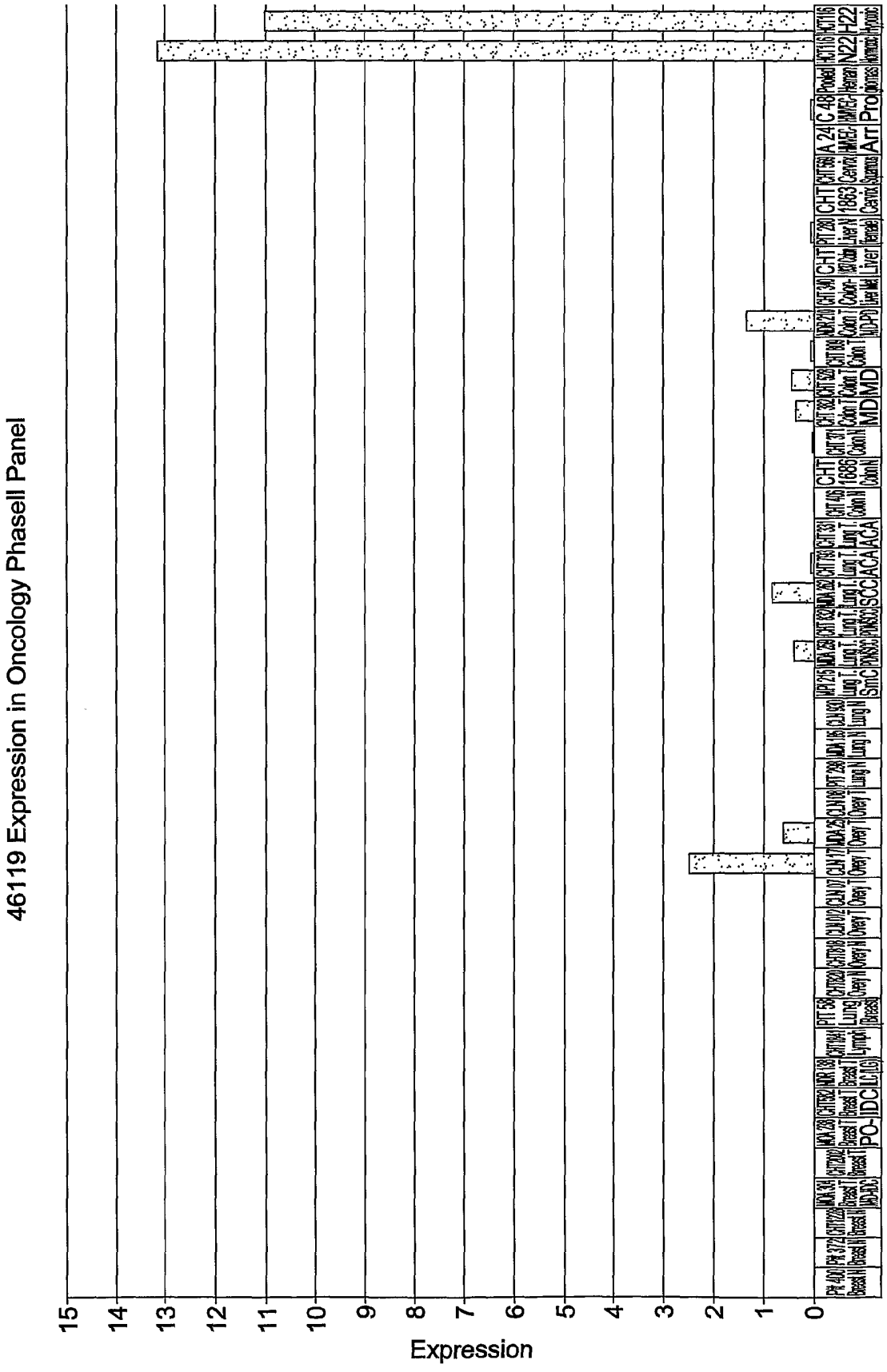

46119 Expression in Oncology Phasell Panel
Expression
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

FIG. 4.

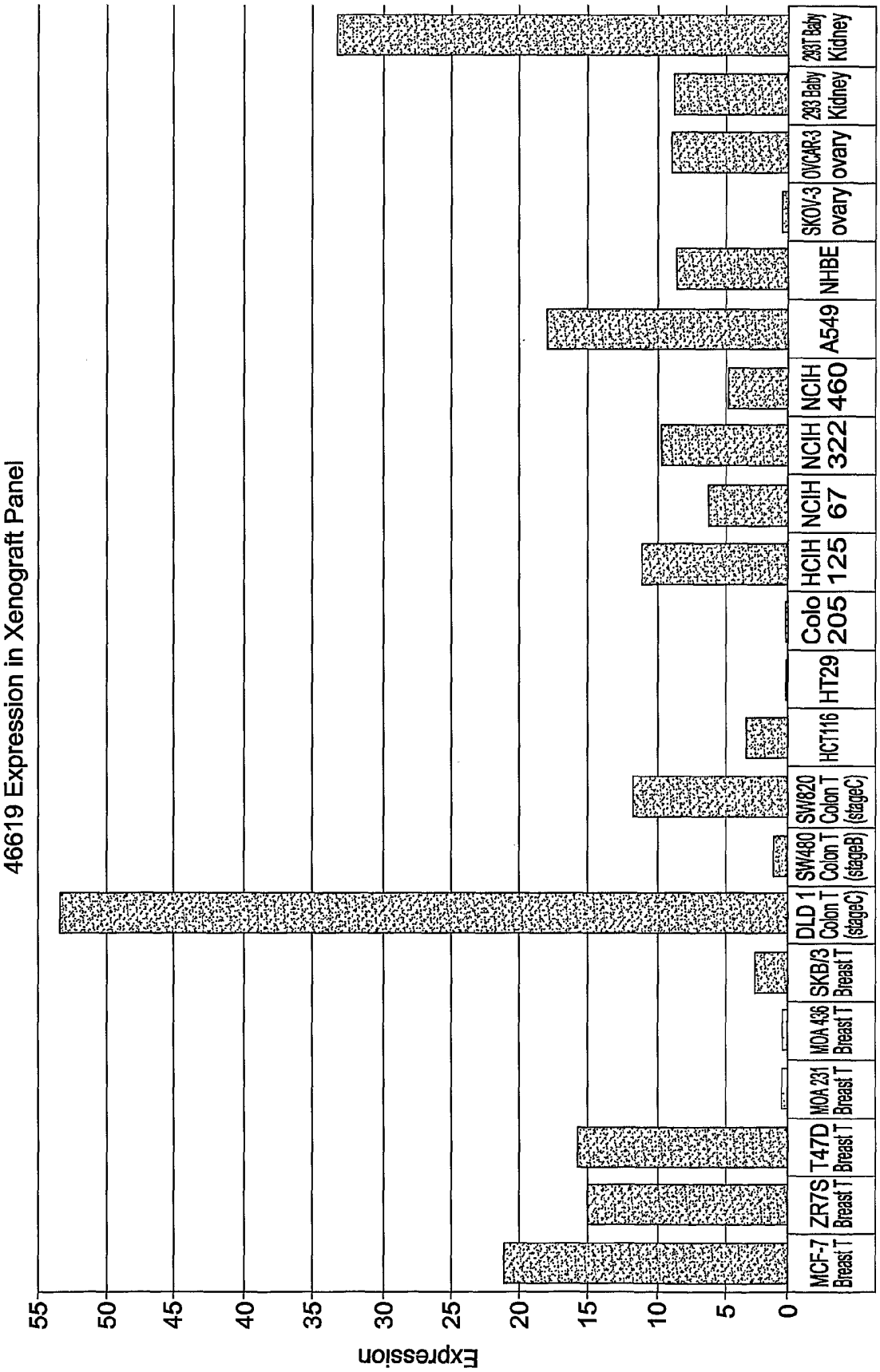

46619 Expression in Xenograft Panel
Expression
0
5
10
15
20
25
30
35
40
45
50
55
MCF-7 Breast T
ZR7S Breast T
T47D Breast T
MDA 231 Breast T
MDA 436 Breast T
SKBI/3 Breast T
DLD 1 Colon T (stageC)
SW480 Colon T (stageB)
SW620 Colon T (stageC)
HCT116
HT29
Colo 205
HCIH 125
NCIH 67
NCIH 322
NCIH 460
A549
NHBE
SKOV-3 ovary
OVCAR3 ovary
293 Baby Kidney
293T Baby Kidney

FIG. 5.

SUBSTITUTE SHEET (RULE 26)

Protein Family / Domain Matches, HMMer version 2

Searching for complete domains in PFAM
hmmpfam - search a single seq against HMM database

Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL).

HMM file: /prod/ddm/seqanal/PFAM/pfam5.3/pfam
Sequence file: /prod/ddm/wspace/orfanal/oa-script.28220. seq

Query: 46619

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E- value	N
ketoacyl-synt	Beta-ketoacyl synthase	226.2	4.8e-64	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
ketoacyl-synt	1/1	44	459	.]	1	452	[]	226.2	4.8e-64

FIG. 6A.

```
ketoacyl-synt: domain 1 of 1, from 44 to 459: score 226.2, E = 4.8e-64
                *->IAIvGmscRfPgavnspeefWdlLreGRdai.......reiPedRG
                   ++I+G++ + P +v+  + +Wd L +G ++i +  +++   iP +
 46619   44       VVITGIGLVTPLGVG-THLVWDRLIGGESGIvslvgeeyKSIPCSV-  88

                WDldrlrdpdadnpgtsyskyGgfldddadrFDpqFqFgisPrEaaamDPQ
                   ++++    ++     s+ G f +  + +  ++   +s+    a  ++
 46619   89    ---AAYV----PR----GSDEGQFNEQNFVSKSDI-KSMSSPTIMAIGAA 126

                qRlfLetaw....ealEdAGytpkslrgskhaekrdrvGVFvGvmsqeYg
                ++++   +w++++ea  +A  ++ + +  +          V v+++   ++
 46619   127   ELAMKDSGWhpqsEADQVATGVAIGMGMIP-------LEV-VSETALNFQ 168

                pdgareeedvegylvtggnasvlagRIaytfgleGPsisVDTACSSSLvA
                + g     +v + +v++   + +ag ++   +l+GP+ +V+TAC+++++A
 46619   169   TKGYN---KVSPFFVPKILVNMAAGQVSIRYKLKGPNHAVSTACTTGAHA 215

                iHlAcqsLrrGecdlAlAGGVsvmltPetyveFdrqgaLSpkyDdGEkrC
                ++     +++G++d+ AGG ++ ++P + ++F+r++aLS+ + d+  +C
 46619   216   VGDSFRFIAHGDADVMVAGGTDSCISPLSLAGFSRARALST-NSDPKLAC 264

                kaFdagADGFvlgEGvGaVVLkrLsdAradGdhIYaVIkGsAvNqDGasn
                ++F   +DGFv+gEG++++VL++ ++A+ ++++IYa++ G++  +D+ +
 46619   265   RPFHPKRDGFVMGEGAAVLVLEEYEHAVQRRARIYAEVLGYGLSGDAGHI 314

                GlsaPsgeaQarvikkALeeaGvdpedIsYVEAHGTGTqlGDpiElaAll
                +   P+ge++ r++  AL  aGv pe+IsY++AH T+T+lGD+ E +A++
 46619   315   TAPDPEGEGALRCMAAALKDAGVQPEEISYINAHATSTPLGDAAENKAIK 364

                evfgedrddkqplalGSVKSNIGHtlaAAGvagviKVllaLkhgvlpPsL
                  f +++   + la +S+K  +GH l+AGG  ++    La  +++lpP+L
 46619   365   HLFKDHA---YALAVSSTKGATGHLLGAAGAVEAAFTTLACYYQKLPPTL 411

                hieelnphnFPiDlesgplyvlrEteeWerpvidsaGkrvelPRrAgv.S
                + +   p+   +Dl     yv   ++eW+       ++     R +g+++
 46619   412   NLDCSEPE---FDLN----YVPLKAQEWK-------TEK----RFIGLtN 443

                sFGasGTNAHlilleeakP<-*
                sFG++GTNA+l l+++
 46619   444   SFGFGGTNATLCIAGL--    459
```

FIG. 6B.

SEQUENCE LISTING

<110> Meyers, Rachel
Williamson, Mark

<120> 46619, A Novel Human Beta-Ketoacyl Synthase and Uses Thereof

<130> 35800/235514

<160> 4

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1489
<212> DNA
<213> homo sapiens

<220>
<221> CDS
<222> (58)...(1437)

<400> 1
aagtgtccgg ggtagccccg ttacaggaat gtgtttctga tcatctgaat cttaatc atg 60
Met
1

tcc aac tgc ctg caa aat ttc ctg aaa att aca agc act cgt ctt cta 108
Ser Asn Cys Leu Gln Asn Phe Leu Lys Ile Thr Ser Thr Arg Leu Leu
5 10 15

tgt tca aga tta tgc caa cag tta aga agt aaa agg aag ttt ttc gga 156
Cys Ser Arg Leu Cys Gln Gln Leu Arg Ser Lys Arg Lys Phe Phe Gly
20 25 30

act gtg cca ata tcc aga ttg cat agg cga gtt gtc att aca ggc att 204
Thr Val Pro Ile Ser Arg Leu His Arg Arg Val Val Ile Thr Gly Ile
35 40 45

ggc tta gtg act cct ctt ggt gtt gga act cac ctg gtt tgg gat cgt 252
Gly Leu Val Thr Pro Leu Gly Val Gly Thr His Leu Val Trp Asp Arg
50 55 60 65

ctt atc gga gga gag agt gga att gtt tca ctg gtt ggt gaa gag tat 300
Leu Ile Gly Gly Glu Ser Gly Ile Val Ser Leu Val Gly Glu Glu Tyr
70 75 80

aag agt atc cct tgc agt gtt gct gct tat gtg cca aga ggt agt gat 348
Lys Ser Ile Pro Cys Ser Val Ala Ala Tyr Val Pro Arg Gly Ser Asp
85 90 95

gaa ggt cag ttc aat gaa caa aac ttt gtg tcc aaa tca gat atc aag 396
Glu Gly Gln Phe Asn Glu Gln Asn Phe Val Ser Lys Ser Asp Ile Lys
100 105 110

tcc atg tct tct ccc acc atc atg gcc att ggg gct gca gaa tta gcc 444
Ser Met Ser Ser Pro Thr Ile Met Ala Ile Gly Ala Ala Glu Leu Ala
115 120 125

atg aag gat tct ggc tgg cat cct cag tca gaa gct gat caa gtg gct 492
Met Lys Asp Ser Gly Trp His Pro Gln Ser Glu Ala Asp Gln Val Ala
130 135 140 145

act ggt gtt gca att ggc atg gga atg att cct ctt gaa gtt gtt tct 540
Thr Gly Val Ala Ile Gly Met Gly Met Ile Pro Leu Glu Val Val Ser
150 155 160

```
gaa act gct ttg aat ttt cag aca aaa ggt tac aat aaa gtt agc cca   588
Glu Thr Ala Leu Asn Phe Gln Thr Lys Gly Tyr Asn Lys Val Ser Pro
            165                 170                 175

ttt ttt gtc cct aag att ctg gtc aat atg gca gca ggc cag gtc agc   636
Phe Phe Val Pro Lys Ile Leu Val Asn Met Ala Ala Gly Gln Val Ser
        180                 185                 190

att cga tat aaa ctc aag ggc cca aat cat gca gta tcc aca gcc tgt   684
Ile Arg Tyr Lys Leu Lys Gly Pro Asn His Ala Val Ser Thr Ala Cys
    195                 200                 205

acc aca gga gct cat gct gtg gga gac tca ttt aga ttt ata gcc cat   732
Thr Thr Gly Ala His Ala Val Gly Asp Ser Phe Arg Phe Ile Ala His
210                 215                 220                 225

ggt gat gct gat gtg atg gtg gct gga ggt aca gat tct tgt att agc   780
Gly Asp Ala Asp Val Met Val Ala Gly Gly Thr Asp Ser Cys Ile Ser
                230                 235                 240

cct tta tct ctt gct ggg ttt tcc aga gcc cgg gct ctg agc aca aac   828
Pro Leu Ser Leu Ala Gly Phe Ser Arg Ala Arg Ala Leu Ser Thr Asn
            245                 250                 255

tca gat ccc aag ttg gca tgt cga cca ttt cat cca aag aga gat ggt   876
Ser Asp Pro Lys Leu Ala Cys Arg Pro Phe His Pro Lys Arg Asp Gly
        260                 265                 270

ttt gta atg gga gaa ggt gca gct gtg ctg gtg ctg gaa gaa tat gaa   924
Phe Val Met Gly Glu Gly Ala Ala Val Leu Val Leu Glu Glu Tyr Glu
    275                 280                 285

cat gct gtt caa aga aga gcc cgg atc tat gca gaa gtt ttg ggc tat   972
His Ala Val Gln Arg Arg Ala Arg Ile Tyr Ala Glu Val Leu Gly Tyr
290                 295                 300                 305

gga ctc tca ggt gat gct ggt cac ata act gcc cct gat cct gaa gga   1020
Gly Leu Ser Gly Asp Ala Gly His Ile Thr Ala Pro Asp Pro Glu Gly
                310                 315                 320

gaa ggt gcc tta agg tgt atg gct gct gct tta aaa gat gca ggt gtg   1068
Glu Gly Ala Leu Arg Cys Met Ala Ala Ala Leu Lys Asp Ala Gly Val
            325                 330                 335

cag cct gag gag ata tcc tat atc aat gca cat gct act tcc aca cca   1116
Gln Pro Glu Glu Ile Ser Tyr Ile Asn Ala His Ala Thr Ser Thr Pro
        340                 345                 350

ttg gga gat gct gct gaa aac aaa gct atc aaa cat ctc ttc aaa gac   1164
Leu Gly Asp Ala Ala Glu Asn Lys Ala Ile Lys His Leu Phe Lys Asp
    355                 360                 365

cat gca tat gcc ctt gca gtt tcc tca act aag gga gca aca gga cat   1212
His Ala Tyr Ala Leu Ala Val Ser Ser Thr Lys Gly Ala Thr Gly His
370                 375                 380                 385

ctg ctg gga gct gca ggg gca gtc gag gca gct ttt acc aca tta gct   1260
Leu Leu Gly Ala Ala Gly Ala Val Glu Ala Ala Phe Thr Thr Leu Ala
                390                 395                 400

tgt tat tat caa aaa cta cca cct act tta aac ctg gat tgt tcg gaa   1308
Cys Tyr Tyr Gln Lys Leu Pro Pro Thr Leu Asn Leu Asp Cys Ser Glu
            405                 410                 415

cca gaa ttt gat ctc aac tat gtt cca cta aag gca cag gaa tgg aaa   1356
Pro Glu Phe Asp Leu Asn Tyr Val Pro Leu Lys Ala Gln Glu Trp Lys
        420                 425                 430
```

```
act gag aaa aga ttt att ggc ctc acc aat tcc ttt ggt ttt ggt ggt   1404
Thr Glu Lys Arg Phe Ile Gly Leu Thr Asn Ser Phe Gly Phe Gly Gly
    435                 440                 445

act aat gca aca ctt tgt att gct gga ctg tag aacatataat ttgtaattaa 1457
Thr Asn Ala Thr Leu Cys Ile Ala Gly Leu  *
450                 455

atactgattt ttaaatgtta aaaaaaaaaa aa                               1489

<210> 2
<211> 459
<212> PRT
<213> homo sapiens

<400> 2
Met Ser Asn Cys Leu Gln Asn Phe Leu Lys Ile Thr Ser Thr Arg Leu
 1               5                  10                  15
Leu Cys Ser Arg Leu Cys Gln Gln Leu Arg Ser Lys Arg Lys Phe Phe
            20                  25                  30
Gly Thr Val Pro Ile Ser Arg Leu His Arg Arg Val Val Ile Thr Gly
        35                  40                  45
Ile Gly Leu Val Thr Pro Leu Gly Val Gly Thr His Leu Val Trp Asp
    50                  55                  60
Arg Leu Ile Gly Gly Glu Ser Gly Ile Val Ser Leu Val Gly Glu Glu
65                  70                  75                  80
Tyr Lys Ser Ile Pro Cys Ser Val Ala Ala Tyr Val Pro Arg Gly Ser
                85                  90                  95
Asp Glu Gly Gln Phe Asn Glu Gln Asn Phe Val Ser Lys Ser Asp Ile
            100                 105                 110
Lys Ser Met Ser Ser Pro Thr Ile Met Ala Ile Gly Ala Ala Glu Leu
        115                 120                 125
Ala Met Lys Asp Ser Gly Trp His Pro Gln Ser Glu Ala Asp Gln Val
    130                 135                 140
Ala Thr Gly Val Ala Ile Gly Met Gly Met Ile Pro Leu Glu Val Val
145                 150                 155                 160
Ser Glu Thr Ala Leu Asn Phe Gln Thr Lys Gly Tyr Asn Lys Val Ser
                165                 170                 175
Pro Phe Phe Val Pro Lys Ile Leu Val Asn Met Ala Ala Gly Gln Val
            180                 185                 190
Ser Ile Arg Tyr Lys Leu Lys Gly Pro Asn His Ala Val Ser Thr Ala
        195                 200                 205
Cys Thr Thr Gly Ala His Ala Val Gly Asp Ser Phe Arg Phe Ile Ala
    210                 215                 220
His Gly Asp Ala Asp Val Met Val Ala Gly Gly Thr Asp Ser Cys Ile
225                 230                 235                 240
Ser Pro Leu Ser Leu Ala Gly Phe Ser Arg Ala Arg Ala Leu Ser Thr
                245                 250                 255
Asn Ser Asp Pro Lys Leu Ala Cys Arg Pro Phe His Pro Lys Arg Asp
            260                 265                 270
Gly Phe Val Met Gly Glu Gly Ala Ala Val Leu Val Leu Glu Glu Tyr
        275                 280                 285
Glu His Ala Val Gln Arg Arg Ala Arg Ile Tyr Ala Glu Val Leu Gly
    290                 295                 300
Tyr Gly Leu Ser Gly Asp Ala Gly His Ile Thr Ala Pro Asp Pro Glu
305                 310                 315                 320
Gly Glu Gly Ala Leu Arg Cys Met Ala Ala Ala Leu Lys Asp Ala Gly
                325                 330                 335
Val Gln Pro Glu Glu Ile Ser Tyr Ile Asn Ala His Ala Thr Ser Thr
            340                 345                 350
Pro Leu Gly Asp Ala Ala Glu Asn Lys Ala Ile Lys His Leu Phe Lys
        355                 360                 365
Asp His Ala Tyr Ala Leu Ala Val Ser Ser Thr Lys Gly Ala Thr Gly
    370                 375                 380
His Leu Leu Gly Ala Ala Gly Ala Val Glu Ala Ala Phe Thr Thr Leu
385                 390                 395                 400
Ala Cys Tyr Tyr Gln Lys Leu Pro Pro Thr Leu Asn Leu Asp Cys Ser
```

```
                405                 410                 415
Glu Pro Glu Phe Asp Leu Asn Tyr Val Pro Leu Lys Ala Gln Glu Trp
            420                 425                 430
Lys Thr Glu Lys Arg Phe Ile Gly Leu Thr Asn Ser Phe Gly Phe Gly
        435                 440                 445
Gly Thr Asn Ala Thr Leu Cys Ile Ala Gly Leu
    450                 455


<210> 3
<211> 1380
<212> DNA
<213> homo sapiens

<220>
<221> CDS
<222> (1)...(1380)

<400> 3
atg tcc aac tgc ctg caa aat ttc ctg aaa att aca agc act cgt ctt   48
Met Ser Asn Cys Leu Gln Asn Phe Leu Lys Ile Thr Ser Thr Arg Leu
 1               5                  10                  15

cta tgt tca aga tta tgc caa cag tta aga agt aaa agg aag ttt ttc   96
Leu Cys Ser Arg Leu Cys Gln Gln Leu Arg Ser Lys Arg Lys Phe Phe
            20                  25                  30

gga act gtg cca ata tcc aga ttg cat agg cga gtt gtc att aca ggc   144
Gly Thr Val Pro Ile Ser Arg Leu His Arg Arg Val Val Ile Thr Gly
        35                  40                  45

att ggc tta gtg act cct ctt ggt gtt gga act cac ctg gtt tgg gat   192
Ile Gly Leu Val Thr Pro Leu Gly Val Gly Thr His Leu Val Trp Asp
    50                  55                  60

cgt ctt atc gga gga gag agt gga att gtt tca ctg gtt ggt gaa gag   240
Arg Leu Ile Gly Gly Glu Ser Gly Ile Val Ser Leu Val Gly Glu Glu
65                  70                  75                  80

tat aag agt atc cct tgc agt gtt gct gct tat gtg cca aga ggt agt   288
Tyr Lys Ser Ile Pro Cys Ser Val Ala Ala Tyr Val Pro Arg Gly Ser
                85                  90                  95

gat gaa ggt cag ttc aat gaa caa aac ttt gtg tcc aaa tca gat atc   336
Asp Glu Gly Gln Phe Asn Glu Gln Asn Phe Val Ser Lys Ser Asp Ile
            100                 105                 110

aag tcc atg tct tct ccc acc atc atg gcc att ggg gct gca gaa tta   384
Lys Ser Met Ser Ser Pro Thr Ile Met Ala Ile Gly Ala Ala Glu Leu
        115                 120                 125

gcc atg aag gat tct ggc tgg cat cct cag tca gaa gct gat caa gtg   432
Ala Met Lys Asp Ser Gly Trp His Pro Gln Ser Glu Ala Asp Gln Val
    130                 135                 140

gct act ggt gtt gca att ggc atg gga atg att cct ctt gaa gtt gtt   480
Ala Thr Gly Val Ala Ile Gly Met Gly Met Ile Pro Leu Glu Val Val
145                 150                 155                 160

tct gaa act gct ttg aat ttt cag aca aaa ggt tac aat aaa gtt agc   528
Ser Glu Thr Ala Leu Asn Phe Gln Thr Lys Gly Tyr Asn Lys Val Ser
                165                 170                 175

cca ttt ttt gtc cct aag att ctg gtc aat atg gca gca ggc cag gtc   576
Pro Phe Phe Val Pro Lys Ile Leu Val Asn Met Ala Ala Gly Gln Val
            180                 185                 190

agc att cga tat aaa ctc aag ggc cca aat cat gca gta tcc aca gcc   624
```

```
Ser Ile Arg Tyr Lys Leu Lys Gly Pro Asn His Ala Val Ser Thr Ala
        195                 200                 205

tgt acc aca gga gct cat gct gtg gga gac tca ttt aga ttt ata gcc   672
Cys Thr Thr Gly Ala His Ala Val Gly Asp Ser Phe Arg Phe Ile Ala
    210                 215                 220

cat ggt gat gct gat gtg atg gtg gct gga ggt aca gat tct tgt att   720
His Gly Asp Ala Asp Val Met Val Ala Gly Gly Thr Asp Ser Cys Ile
225                 230                 235                 240

agc cct tta tct ctt gct ggg ttt tcc aga gcc cgg gct ctg agc aca   768
Ser Pro Leu Ser Leu Ala Gly Phe Ser Arg Ala Arg Ala Leu Ser Thr
                245                 250                 255

aac tca gat ccc aag ttg gca tgt cga cca ttt cat cca aag aga gat   816
Asn Ser Asp Pro Lys Leu Ala Cys Arg Pro Phe His Pro Lys Arg Asp
            260                 265                 270

ggt ttt gta atg gga gaa ggt gca gct gtg ctg gtg ctg gaa gaa tat   864
Gly Phe Val Met Gly Glu Gly Ala Ala Val Leu Val Leu Glu Glu Tyr
        275                 280                 285

gaa cat gct gtt caa aga aga gcc cgg atc tat gca gaa gtt ttg ggc   912
Glu His Ala Val Gln Arg Arg Ala Arg Ile Tyr Ala Glu Val Leu Gly
    290                 295                 300

tat gga ctc tca ggt gat gct ggt cac ata act gcc cct gat cct gaa   960
Tyr Gly Leu Ser Gly Asp Ala Gly His Ile Thr Ala Pro Asp Pro Glu
305                 310                 315                 320

gga gaa ggt gcc tta agg tgt atg gct gct gct tta aaa gat gca ggt   1008
Gly Glu Gly Ala Leu Arg Cys Met Ala Ala Ala Leu Lys Asp Ala Gly
                325                 330                 335

gtg cag cct gag gag ata tcc tat atc aat gca cat gct act tcc aca   1056
Val Gln Pro Glu Glu Ile Ser Tyr Ile Asn Ala His Ala Thr Ser Thr
            340                 345                 350

cca ttg gga gat gct gct gaa aac aaa gct atc aaa cat ctc ttc aaa   1104
Pro Leu Gly Asp Ala Ala Glu Asn Lys Ala Ile Lys His Leu Phe Lys
        355                 360                 365

gac cat gca tat gcc ctt gca gtt tcc tca act aag gga gca aca gga   1152
Asp His Ala Tyr Ala Leu Ala Val Ser Ser Thr Lys Gly Ala Thr Gly
    370                 375                 380

cat ctg ctg gga gct gca ggg gca gtc gag gca gct ttt acc aca tta   1200
His Leu Leu Gly Ala Ala Gly Ala Val Glu Ala Ala Phe Thr Thr Leu
385                 390                 395                 400

gct tgt tat tat caa aaa cta cca cct act tta aac ctg gat tgt tcg   1248
Ala Cys Tyr Tyr Gln Lys Leu Pro Pro Thr Leu Asn Leu Asp Cys Ser
                405                 410                 415

gaa cca gaa ttt gat ctc aac tat gtt cca cta aag gca cag gaa tgg   1296
Glu Pro Glu Phe Asp Leu Asn Tyr Val Pro Leu Lys Ala Gln Glu Trp
            420                 425                 430

aaa act gag aaa aga ttt att ggc ctc acc aat tcc ttt ggt ttt ggt   1344
Lys Thr Glu Lys Arg Phe Ile Gly Leu Thr Asn Ser Phe Gly Phe Gly
        435                 440                 445

ggt act aat gca aca ctt tgt att gct gga ctg tag                  1380
Gly Thr Asn Ala Thr Leu Cys Ile Ala Gly Leu  *
    450                 455
```

<210> 4
<211> 458
<212> PRT
<213> Artificial Sequence

<220>
<223> Pfam beta-ketoacyl synthase domain.

<400> 4

```
Ile Ala Ile Val Gly Met Ser Cys Arg Phe Pro Gly Ala Val Asn Ser
 1               5                  10                  15
Pro Glu Glu Phe Trp Asp Leu Leu Arg Glu Gly Arg Asp Ala Ile Arg
            20                  25                  30
Glu Ile Pro Glu Asp Arg Gly Trp Asp Leu Asp Arg Leu Arg Asp Pro
        35                  40                  45
Asp Ala Asp Asn Pro Gly Thr Ser Tyr Ser Lys Tyr Gly Gly Phe Leu
    50                  55                  60
Asp Asp Ala Asp Arg Phe Asp Pro Gln Phe Gln Phe Gly Ile Ser Pro
65                  70                  75                  80
Arg Glu Ala Ala Ala Met Asp Pro Gln Gln Arg Leu Phe Leu Glu Thr
                85                  90                  95
Ala Trp Glu Ala Leu Glu Asp Ala Gly Tyr Thr Pro Lys Ser Leu Arg
            100                 105                 110
Gly Ser Lys His Ala Glu Lys Arg Asp Arg Val Gly Val Phe Val Gly
        115                 120                 125
Val Met Ser Gln Glu Tyr Gly Pro Asp Gly Ala Arg Glu Glu Glu Asp
    130                 135                 140
Val Glu Gly Tyr Leu Val Thr Gly Gly Asn Ala Ser Val Leu Ala Gly
145                 150                 155                 160
Arg Ile Ala Tyr Thr Phe Gly Leu Glu Gly Pro Ser Ile Ser Val Asp
                165                 170                 175
Thr Ala Cys Ser Ser Ser Leu Val Ala Ile His Leu Ala Cys Gln Ser
            180                 185                 190
Leu Arg Arg Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Ser Val
        195                 200                 205
Met Leu Thr Pro Glu Thr Tyr Val Glu Phe Asp Arg Gln Gly Ala Leu
    210                 215                 220
Ser Pro Lys Tyr Asp Asp Gly Glu Lys Arg Cys Lys Ala Phe Asp Ala
225                 230                 235                 240
Gly Ala Asp Gly Phe Val Leu Gly Glu Gly Val Gly Ala Val Val Leu
                245                 250                 255
Lys Arg Leu Ser Asp Ala Arg Ala Asp Gly Asp His Ile Tyr Ala Val
            260                 265                 270
Ile Lys Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Ser
        275                 280                 285
Ala Pro Ser Gly Glu Ala Gln Ala Arg Val Ile Lys Lys Ala Leu Glu
    290                 295                 300
Glu Ala Gly Val Asp Pro Glu Asp Ile Ser Tyr Val Glu Ala His Gly
305                 310                 315                 320
Thr Gly Thr Gln Leu Gly Asp Pro Ile Glu Leu Ala Ala Leu Leu Glu
                325                 330                 335
Val Phe Gly Glu Asp Arg Asp Asp Lys Gln Pro Leu Ala Leu Gly Ser
            340                 345                 350
Val Lys Ser Asn Ile Gly His Thr Leu Ala Ala Ala Gly Val Ala Gly
        355                 360                 365
Val Ile Lys Val Leu Leu Ala Leu Lys His Gly Val Leu Pro Pro Ser
    370                 375                 380
Leu His Ile Glu Glu Leu Asn Pro His Asn Phe Pro Ile Asp Leu Glu
385                 390                 395                 400
Ser Gly Pro Leu Tyr Val Leu Arg Glu Thr Glu Glu Trp Glu Arg Pro
                405                 410                 415
Val Ile Asp Ser Ala Gly Lys Arg Val Glu Leu Pro Arg Arg Ala Gly
            420                 425                 430
Val Ser Ser Phe Gly Ala Ser Gly Thr Asn Ala His Leu Ile Ile Glu
        435                 440                 445
Glu Ala Lys Pro Arg Thr Ala Arg Thr Ala
    450                 455
```