BIOMARKERS FOR LIVER FIBROTIC INJURY

CROSS REFERENCE TO RELATED APPLICATIONS

[001] The present invention is related to and claims the benefit of, under U.S.C. §119(e), U.S. provisional patent application Serial No. 60/761,959, filed on January 24, 2006, which is expressly incorporated fully herein by reference.

FIELD OF THE INVENTION

[002] Herein described are methods of detecting liver fibrosis and/or cirrhosis based on the differential expression of various proteins, and kits for diagnosing fibrosis and/or cirrhosis.

BACKGROUND

[003] Liver fibrosis represents a continuous disease spectrum characterized by an increase in total liver collagen and other matrix proteins that disrupt the architecture of the liver and impair liver function (1, 2). The progression of fibrosis in the liver is a response to necroinflammatory changes. The overall liver fibrosis process is one of dynamic inflammation and repair and has the potential to be resolved (3). Fibrosis is seen as scar formation in the patient's liver. When the liver becomes permanently injured and scarred, the condition is called cirrhosis.

[004] Liver fibrosis and/or cirrhosis are the major risk factors of hepatocellular carcinoma (HCC). A range of factors, such as hepatitis B virus (HBV), hepatitis C virus (HCV), hepatotoxins, metabolic disorders and alcoholism, can induce both liver fibrosis and/or cirrhosis, which share similar phenotypes (3-7), with cirrhosis being the end-stage of fibrosis. However, it is not clear what types of genes are involved or how they act when liver injury and repair occur. Moreover, the cirrhosis caused by these risk factors often progresses insidiously. Patients with end-stage liver cirrhosis can die within one year unless they accept liver transplantation, which has a 75% five-year survival rate (3).

[005] Previous biochemical studies have reported that there are 39 well-known fibrosis or cirrhosis markers (3, 8, 9), but some markers are obtained through invasive sampling.

[006] Studies for additional markers have proceeded based on microarray analysis of transcriptomes as well as quantitative proteomics. Microarray technologies have been widely used for comprehensive gene expression analysis. In particular, large-scale microarray analysis of gene expression enables researchers to analyze simultaneous changes in thousands of genes and identify significant patterns (10, 11). To date, the most widely used technologies in differential proteomics research were two-dimensional gel electrophoresis (2DE) and liquid chromatography-based isotope-coded affinity tagging (ICAT) technologies (12). Recently, a variation of the ICAT technology, iTRAQ (isobaric tags for relative and absolute quantitation), has been introduced. Both ICAT and iTRAQ tagging permit online identification of multiple markers and relative quantification of these proteins.

SUMMARY OF THE INVENTION

[007] The inventors have used differential analysis such as microarray mRNA expression profiling and quantitative protein profiling to find additional unique and identifiable signatures potentially valuable for the diagnosis and treatment of liver fibrosis and/or cirrhosis.

[008] The invention provides a method of detecting liver fibrosis and/or cirrhosis comprising obtaining a biological sample from a patient and assaying the sample for differential expression of at least one gene encoding a protein chosen from SEQ ID NO: 1 – SEQ ID NO: 63 and human orthologs thereof, wherein the differential expression of at least one gene suggests the presence of liver fibrosis and/or cirrhosis.

[009] Also provided is a method of detecting liver fibrosis and/or cirrhosis by assaying the biological sample for differential expression of the human orthologs described above, wherein the human orthologs comprise at least one gene encoding a protein chosen from SEQ ID NO: 64 – SEQ ID NO: 120.

[010] Also provided is a kit for diagnosing liver fibrosis and/or cirrhosis, which comprises one or more nucleic acid probes that hybridize to nucleic acid molecules of at least one gene encoding a protein chosen from SEQ ID NO: 1 – SEQ ID NO: 63 and human orthologs thereof and packaging indicating use for detection of liver fibrosis and/or cirrhosis.

[011] In addition, a kit is provided for diagnosing liver fibrosis comprising one or more nucleic acid probes as described above, wherein the human orthologs comprise at least one gene encoding a protein chosen from SEQ ID NO: 64 – SEQ ID NO: 120.

[012] The invention further provides a kit for diagnosing liver fibrosis and/or cirrhosis that comprises one or more antibodies that specifically bind to at least one protein encoded by a sequence chosen from SEQ ID NO: 1 – SEQ ID NO: 63 and human orthologs thereof, and packaging indicating use for detection of liver fibrosis and/or cirrhosis.

[013] Also provided is a kit for diagnosis of liver fibrosis and/or cirrhosis comprising one or more antibodies as described above, wherein the human orthologs comprise at least one gene encoding a protein chosen from SEQ ID NO: 64 – SEQ ID NO: 120.

[014] The invention also provides a method of identifying a compound that decreases the differential expression of at least one gene encoding a protein chosen from SEQ ID NO: 7 – SEQ ID NO: 23 and SEQ ID NO: 42 – SEQ ID NO: 63 and human orthologs thereof, comprising providing a cell expressing at least one of the aforementioned genes, contacting the cell with a test compound, and determining whether the differential expression of the at least one gene is decreased in the presence of the test compound, wherein the decreased differential expression is an indication of halting or reversing liver fibrosis and/or cirrhosis.

[015] The invention further provides a method of identifying a compound that increases the differential expression of at least one gene encoding a protein chosen from SEQ ID NO: 1 – SEQ ID NO: 6 and SEQ ID NO: 24 – SEQ ID NO: 41 and human orthologs thereof, comprising providing a cell expressing at least one of the aforementioned genes, contacting the cell with a test compound, and determining whether the differential expression of the at least one gene is increased in the presence of the test compound, wherein the increased differential expression is an indication of halting or reversing liver fibrosis and/or cirrhosis.

[016] Additional objects and advantages of the invention will be set forth in part in the description which follows, and in part will be obvious from the description, or may be learned by practice of the invention. The objects and advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims.

[017] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and do not limit the invention as claimed.

[018] The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate several embodiments of the invention and together with the description, serve to explain the principles of the invention.

SUMMARY OF THE TABLES

[019] Table 1 is the result of a biochemical analysis of the serum of rats treated with dimethylnitrosamine (DMN) as compared with that of the control rats over a six-week time period.

[020] Table 2 is a list of genes showing significant changes in gene expression and their human orthologs. The genes were selected from the combined results of microarray analysis and quantitative proteomic analysis and those whose relationship to fibrosis are not yet disclosed in public literature.

[021] Table 3 is the result of interacting network analysis for the proteins showing significant changes in expression in oligonucleotide microarray (Table 3A) and iTRAQ proteomic (Table 3B) studies. Proteins in the network identified in the present study which have not previously been reported to associate with liver injury are shown in bold.

BRIEF DESCRIPTION OF THE DRAWINGS

[022] Figure 1A is a schematic illustration of DMN treatment for inducing fibrosis in rats. Each rat was either injected with DMN three times per week for three consecutive weeks (shown as inverted triangles) or with normal saline as control. Rats were weighed and sacrificed each week (shown as triangles, starting on day 11, which is referred to as first week, until week six). Blood samples were collected for biochemical assays (summary in Table 1) and livers were excised and weighed. Livers

were then either fixed in formaldehyde for histopathology or used to isolate RNA and protein for microarray and iTRAQ proteomic studies, respectively.

[023] Figure 1B is the quantitative real-time PCR (Q-RT-PCR) result of *Tgfβ1* (transforming growth factor-β1) in DMN-treated rats, showing a higher level of *Tgfβ1* expression, the strongest known inducer of fibrogenesis.

[024] Figure 2A shows the histopathological examination results of liver tissues from DMN-treated rats. The representative phenotype of the rat liver tissue was characterized by scoring the four histopathological features as follows: the necrosis scores are from N0 to N3 (the first row), the inflammation scores are from I0 to I3 (the second row), the fibrosis scores are from F0 to F3 (the third row), and the fatty change scores are presence or absence (+ and -) in the last row.

[025] Figure 2B is a summary of the histopathological scores for the rat model. The results are ranked by time course. The necroinflammatory change is divided into three grades: A0= none, A(1-3)= mild and A(4-6)= moderate necroinflammatory. The fibrosis change is divided into two grades: F(0-1)= normal to fibrous expansion of portal tracts, F(2-3)= bridge fibrosis to frequent bridging fibrosis with nodule formation. The fatty change is shown as presence or absence (+/-). The number of rats is counted and used to calculate the percentage of rats in each histopathological level at each time point.

[026] Figure 3A shows the dendrogram of two hundred and fifty-six (256) gene expression patterns from the DMN-induced fibrosis model in rats.

[027] Figure 3B shows the comparison of *Timp1* (tissue inhibitor of metalloproteinase 1) expression between the Q-RT-PCR result and the microarray data, and a good concordance between the two results is shown. For the Q-RT-PCR assay of *Timp 1*, expression levels (marked by the square) are relative to the mean of all gene expression, and are measured by the log scale on the right side of the plot. The expression levels of two *Timp1* transcripts, rc_AI169327_at and rc_AI169327_g_at (marked by circle and triangle), are relative to the mean of all gene expression levels, and are measured by the scale indicated on the left side of the plot.

[028] Figure 4A shows the vimentin expression ratios in DMN-treated and control rats at 2, 4, and 6 weeks as determined by iTRAQ proteomics. Vimentin expression increased 2.4- to 3.4-fold in the liver when rats were treated with DMN.

[029] Figure 4B shows the result of Western blot analysis of DMN-induced vimentin expression. The observed result matched with the trend measured by mass spectrometry in iTRAQ proteomic experiments.

[030] Figure 5 shows the result of Western blot analysis comparing the expression of carbonic anhydrase I in human sera from normal volunteers and from patients with hepatitis B-related cirrhosis.

DESCRIPTION OF THE EMBODIMENTS

Definitions

[031] As used herein, the term "biological sample" refers to any biological material collected from cells, tissues, or organs of the subject. The source of the biological sample may vary depending on the particular symptoms present in the subject to be diagnosed. The biological sample may be analyzed immediately after it is taken, or stored. If stored, the sample may be equilibrated with an appropriate storage buffer, and kept at 4°C, at -20°C, at -70°C, or even in cryogenic liquids, such as liquid nitrogen or liquid helium. In one embodiment, the biological sample may consist of blood, serum, or plasma. In another embodiment, the biological sample may consist of a biopsy or tissue sample. In additional embodiments of the invention, the biological sample may consist of aminotic fluid, milk, saliva, cerebrospinal fluid, lymph, sweat, mucus, synovial fluid, lacrimal fluid, or other clinical specimens or samples.

[032] As used herein, the term "patient" refers to a mammalian animal, including but not limited to human, primates, domestic mammals, laboratory mammals, etc.

[033] As used herein, the term "differential expression" refers to gene expression on the RNA/mRNA level, protein level, or both RNA/mRNA and protein levels as compared to normal gene

expression, e.g., an increased or decreased gene expression on the RNA/mRNA level, protein level, or both RNA/mRNA and protein levels.

[034] As used herein, the terms "gene" and "gene encoding a protein" refer to any nucleic acid sequence or portion thereof with a functional role in encoding or transcribing a protein. The gene may encompass all the nucleic acids responsible for encoding a functional protein of certain SEQ ID NO or only a portion of the nucleic acids responsible for encoding or expressing a protein of certain SEQ ID NO. The nucleic acid sequence may contain normal sequences as well as genetic abnormalities within exons, introns, initiation or termination regions, promoter sequences, other regulatory sequences or unique adjacent regions to the gene.

[035] As used herein, the term "human ortholog" refers to a corresponding human gene wherein the non-human gene and the human gene are derived from a single ancestral gene in the last common ancestor of human and non-human, and the human genes have the same function.

[036] As used herein, the term "probes" refers to hybridization probes that are oligonucleotides that bind in a base-specific manner to a complementary strand of nucleic acid. Such probes also include peptide nucleic acids, as described in Nielsen *et al.*, 1991, (13), and other nucleic acid analogs and nucleic acid mimetics. *See, e.g.,* U.S. Pat. No. 6,156,501, which is incorporated herein by reference.

[037] As used herein, the term "nucleic acid" refers to any DNA or RNA/mRNA, for example, chromosomal, mitochondrial, viral and/or bacterial nucleic acid present in tissue sample as well as synthetic nucleic acids. The term "nucleic acid" encompasses either or both strands of a double stranded nucleic acid molecule and includes any fragment or portion of an intact nucleic acid molecule.

[038] As used herein, the term "packaging" refers generally to packaging material comprising external labeling and internal material in the container, including but not limited to instructions for using the kit.

[039] As used herein, the term "antibody," refers to an immunoglobulin or a fragment or a derivative thereof, and encompasses any polypeptide comprising an antigen-binding site, whether

produced *in vitro* or *in vivo*. The term includes, but is not limited to, polyclonal, monoclonal, monospecific, polyspecific, humanized, single-chain, chimeric, synthetic, recombinant, hybrid, mutated, and grafted antibodies. The term "antibody" also includes antibody fragments such as Fab, $F(ab')_2$, Fv, scFv, Fd, dAb, and other antibody fragments that retain antigen-binding function, i.e., the ability to bind a specific antigen. Typically, such fragments would comprise an antigen-binding domain, *i.e.,* a part of an antibody molecule that comprises amino acids responsible for the specific binding between the antibody and the antigen. An antigen-binding domain typically comprises an antibody light chain variable region (V_L) and an antibody heavy chain variable region (V_H), however, it does not necessarily have to comprise both. For example, a so-called Fd antibody fragment consists only of a V_H domain, but still retains some antigen-binding function of the intact antibody.

[040] As used herein, the term "specifically binds," or the like, means that two molecules form a complex that is relatively stable under physiologic conditions (*e.g.*, a stable antigen/antibody complex). The term is also applicable where, for example, an antigen-binding domain is specific for a particular epitope, which is found on a number of molecules. Thus, an antibody may specifically bind multiple proteins when it binds to an epitope present in each. Specific binding is characterized by a selective interaction, often including high affinity binding with a low to moderate capacity. Nonspecific binding is usually a less selective interaction, and may have a low affinity with a moderate to high capacity. Typically, binding is considered specific when the affinity is at least $10^5\ M^{-1}$, $10^6\ M^{-1}$, $10^7\ M^{-1}$ or $10^8\ M^{-1}$. If necessary, non-specific binding can be reduced without substantially affecting specific binding by varying the binding conditions. Such conditions are known in the art, and a skilled artisan using routine techniques can select appropriate conditions. The conditions are usually defined in terms of concentrations of antibodies, ionic strength of the solution, temperature, time allowed for binding, concentration of non-related molecules (*e.g.*, blocking agents such as serum albumin or milk casein), and so forth. *See, e.g.*, Morgan et al., "The Matrix Effects on Kinetic Rate Constants of Antibody-Antigen

Interactions Reflect Solvent Viscosity," *J. Immunol. Meth.* 217:51:60 (1998); and Zhuang et al., "Measurement of Association Rate Constant of Antibody-Antigen Interactions in Solution Based on Enzyme-Linked Immunosorbent Assay," *J. Biosci. Bioeng.* 92(4):330-336 (2001).

[041] As used herein, the term "protein" refers to a polymeric form of any length of amino acids, which can include naturally-occurring or synthetic amino acids and coded and non-coded amino acids, peptides, depsipeptides, polypeptides with cyclic, bicyclic, depsicyclic, or depsibicyclic peptide backbones, single chain protein as well as multimers, as well as any fragment or portion of the intact protein molecule.

[042] As used herein, the term "compound" refers to a substance comprising one or more chemical elements in any proportion and of any structure, including but not limited to cyclic, bicyclic, branched or straight chain. The compound may be organic or inorganic. The compound also refers to a composition comprising one or more chemical elements, one or more herbal/plant elements or herbal/plant extractions, or both chemical elements and herbal/plant elements or extractions.

Embodiments

[043] Chronic liver disease is a common and potentially lethal problem in Asia. The development of hepatocellular carcinoma (HCC) is generally preceded by hepatic cirrhosis, which occurs at the end stage of fibrosis. The same proteins are expressed in both cirrhosis and fibrosis. Changes in gene expression of these proteins during liver fibrosis are examined to identify markers of liver fibrosis to assist in the diagnosis of fibrosis and/or cirrhosis. The study begins with the establishment of a liver fibrosis model. Dimethylnitrosamine (DMN), a non-genotoxic hepatotoxin, is used to induce rat necroinflammatory and hepatic fibrosis, as described in Jezequel A.M. *et al.* (14), which is a known model for studying human liver damage (15). During a six-week time course, histopathological, biochemical and quantitative RT-PCR analyses confirmed the incidence of hepatic fibrosis in the rat model system.

[044] The microarray and the iTRAQ quantitative proteomics technology were used. The iTRAQ technology permits the identification of multiple proteins at the same time as well as the relative quantification of these proteins. Applied Biosystems iTRAQ Reagents are a multiplexed set of four isobaric reagents. The four reagents are amine-specific and yield labeled peptides. The labeled peptides are identical in mass and hence also identical in single MS mode. They produce strong, diagnostic, low-mass MS/MS signature ions allowing for quantification of up to four different samples simultaneously. Quantification is performed via the differences in abundance of four product ions, i.e., product ions weighing 114, 115, 116, and 117 daltons that are each cleaved from one of the four possible tags. Since all peptides are tagged, proteome coverage is expanded and analysis of multiple peptides per protein improves the confidence in those identified (16). The multi-sample capability of the iTRAQ technology provides a way to compare the protein expression profile of different liver states simultaneously.

[045] The model employed in the present study is very similar to the human liver study model, and thus findings from this study can have human clinical applications. This study employed DMN, a potent non-genotoxic hepatotoxin, to simulate liver fibrosis (17, 18). DMN has been demonstrated to induce liver damage rapidly and also has been empirically proven to be useful for the study of human fibrosis formation (14-15, 19) as mentioned above. Also, the serum markers that showed significant differences in expression in the rat liver fibrosis model established in this study as compared to the controls, see Table 1, are the same serum markers for human liver fibrosis, e.g., the thirteen serum markers of albumin, glutamic oxaloacetic transaminase (GOT), glutamic pyruvic transaminase (GPT), total bilirubin, alkaline phosphatase (AKP), acid phosphatases (ACP), α-fetoprotein (AFP), blood urea nitrogen (BUN), cholesterol (CHOL), lactate dehydrogenase (LDH), globulin, prothrombin time (PT) and blood platelets (PLT). Thus, the markers identified from this study can be used to diagnose liver fibrosis and/or cirrhosis in humans.

Detection Methods

[046] The present study identified nearly 97 genes as biomarkers for liver fibrotic injury. Results of a PubMed literature search indicate that approximately 30 % of the differentially expressed genes identified in this study by microarray and proteomics approach had been proven to be related to fibrosis or cirrhosis. The remaining 70% are not yet reported in any literature. Combining all the information obtained by both microarray and quantitative proteomics technology, sixty-three (63) genes were selected, all of which showed significant changes in expression between DMN-treated and DMN-untreated groups, but have not been previously reported in any literature in relation to liver fibrosis or cirrhosis (Table 2). The 63 genes in DMN-treated rats have a difference in gene expression on the RNA/mRNA level and/or the protein level when compared to normal gene expression. Some genes are up-regulated while others are down-regulated. Those 63 rat genes were converted into their human ortholog genes according to the ortholog assertions from the EnsEMBL and HomoloGene database (Table 2 – SEQ ID NO: 1 – SEQ ID NO: 120).

[047] Thus the invention provides a method for detecting liver fibrotic injury based on the liver injury-related differential gene expression study described herein. In the detection method, differential expression of a gene encoding a protein includes differential expression of the full length gene encoding a full length protein or a portion of the gene encoding a portion of the protein.

[048] Accordingly, the invention provides a method of detecting liver fibrosis and/or cirrhosis comprising obtaining a biological sample from a patient, and assaying the sample for differential expression of at least one gene encoding a protein chosen from SEQ ID NO: 1 to SEQ ID NO: 63 and human orthologs thereof, wherein the differential expression of at least one gene suggests the presence of liver fibrosis and/or cirrhosis. In another embodiment, the invention provides a method of detecting liver fibrosis and/or cirrhosis by assaying the sample for differential expression of at least one gene encoding a protein chosen from the human orthologs comprising SEQ ID NO: 64 to SEQ ID NO: 120.

[049] Human orthologs of proteins identified in this study can be derived from the search using gene/protein on the Ensembl database at http://www.ensembl.org/Rattus_norvegicus/ together with the HomoloGene NCBI database at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=homologene.

[050] In the practice of this invention, differential gene expression may be assayed by transcription analysis or quantitative proteomic analysis such as microarray, Q-RT-PCR, ICAT and iTRAQ, or any other appropriate methods known to one skilled in the art.

[051] In another embodiment, the invention relates to assaying the up-regulation or increased expression of one or more genes encoding a protein chosen from SEQ ID NO: 7 – SEQ ID NO: 23 and SEQ ID NO: 42 – SEQ ID NO: 63 and human orthologs thereof, wherein the increased expression suggests the presence of liver fibrosis and/or cirrhosis. In yet another embodiment, the invention relates to assaying for the increased expression of at least one gene encoding a protein chosen from the human orthologs comprising SEQ ID NO: 68 - SEQ ID NO: 82 and SEQ ID NO: 100 – SEQ ID NO: 120.

[052] The invention also provides, in another embodiment, assaying for the down-regulation or decreased expression of one or more genes encoding a protein chosen from SEQ ID NO: 1 – SEQ ID NO: 6 and SEQ ID NO: 24 – SEQ ID NO: 41 and human orthologs thereof, wherein the decreased expression suggests the presence of liver fibrosis and/or cirrhosis. In yet another embodiment, the invention relates to assaying for the decreased expression of at least one gene encoding a protein chosen from the human orthologs comprising SEQ ID NO: 64 – SEQ ID NO: 67 and SEQ ID NO: 83 – SEQ ID NO: 99.

[053] The invention also provides a clustering of the genes by functional groups. Thus, the invention provides a method of detecting liver fibrosis and/or cirrhosis comprising assaying a biological sample for differential expression of at least one gene encoding a protein that functions in cancer, cell cycle, and cell morphology chosen from sequences comprising SEQ ID NO: 9, SEQ ID NO: 69, SEQ ID NO: 12, SEQ ID NO: 72, SEQ ID NO: 4, SEQ ID NO: 65, SEQ ID NO: 17, SEQ ID NO: 76, SEQ ID NO: 7, SEQ ID NO: 68, SEQ ID NO: 19, SEQ ID NO: 78, SEQ ID NO: 5, SEQ ID NO: 66, SEQ ID NO: 20, SEQ ID NO: 79, SEQ ID NO: 22, SEQ ID NO: 81, SEQ ID NO: 21, and SEQ ID NO: 80; a protein

that functions in lipid metabolism, small molecule biochemistry, organismal injury and abnormalities chosen from sequences comprising SEQ ID NO: 10, SEQ ID NO: 70, SEQ ID NO: 3, SEQ ID NO: 64, SEQ ID NO: 14, SEQ ID NO: 73, SEQ ID NO: 15 and SEQ ID NO: 74; a protein that functions in hematological disease, endocrine system development and function, nervous system development and function comprising SEQ ID NO: 13; a protein that functions in cancer, cell morphology, dermatological diseases and conditions chosen from sequences comprising SEQ ID NO: 8, SEQ ID NO: 44, SEQ ID NO: 102, SEQ ID NO: 45, SEQ ID NO: 103, SEQ ID NO: 48, SEQ ID NO: 106, SEQ ID NO: 31, SEQ ID NO: 90, SEQ ID NO: 35, SEQ ID NO: 93, SEQ ID NO: 50, SEQ ID NO: 108, SEQ ID NO: 51, SEQ ID NO: 109, SEQ ID NO: 40, and SEQ ID NO: 98; a protein that functions in lipid metabolism, small molecule biochemistry, and molecular transport chosen from sequences comprising SEQ ID NO: 25, SEQ ID NO: 84, SEQ ID NO: 42, SEQ ID NO: 100, SEQ ID NO: 27, SEQ ID NO: 86, SEQ ID NO:47, SEQ ID NO: 105, SEQ ID NO: 36, SEQ ID NO: 94, SEQ ID NO: 16, SEQ ID NO: 75, SEQ ID NO: 53, SEQ ID NO: 111, SEQ ID NO: 37, SEQ ID NO: 95, SEQ ID NO: 38, SEQ ID NO: 96, and SEQ ID NO: 107; and a protein that functions in cancer, cellular movement, cellular growth and proliferation chosen from sequences comprising SEQ ID NO: 59, SEQ ID NO: 117, SEQ ID NO:46, SEQ ID NO: 104, SEQ ID NO: 30, SEQ ID NO: 89, SEQ ID NO: 52, SEQ ID NO: 110, SEQ ID NO: 58, SEQ ID NO: 116, SEQ ID NO: 55, SEQ ID NO: 113, SEQ ID NO: 39, SEQ ID NO: 97, SEQ ID NO: 56, SEQ ID NO: 114, SEQ ID NO: 57, and SEQ ID NO: 115.

Diagnostic Kits

[054] The invention also provides diagnostic kits based on the genes and/or proteins described above. In one embodiment, there is a kit for diagnosing liver fibrosis and/or cirrhosis that comprises one or more nucleic acid probes that hybridize to nucleic acid molecules of at least one gene encoding a protein marker chosen from SEQ ID NO: 1 – SEQ ID NO: 63 and human orthologs thereof and packaging indicating the use for detecting liver fibrosis. The kit can detect the differential expression, e.g., the up-regulation or down-regulation, of these genes, wherein the differential expression would

suggest liver fibrosis and/or cirrhosis. In another embodiment, the invention also provides a kit for diagnosing fibrosis and/or cirrhosis comprising nucleic acid probes that hybridize to nucleic acid molecules of at least one gene encoding a protein marker chosen from SEQ ID NO: 1 – SEQ ID NO: 63 and human orthologs thereof. The differential expression of the genes would suggest fibrosis and/or cirrhosis. In another embodiment, the diagnostic kit comprises one or more nucleic acid probes that hybridize to nucleic acid molecules of at least one gene encoding the human orthologs comprising proteins chosen from SEQ ID NO: 64 - SEQ ID NO: 120 for diagnosing liver fibrosis and/or cirrhosis.

[055] The kit may be prepared by techniques known to one skilled in the art. By way of example, the probes may be labeled, using labeling techniques that are known to one skilled in the art, to facilitate detection, including but not limited to radioisotope labels or fluorescent labels. The probes can hybridize to nucleic acid molecules that are either or both strands of a double stranded nucleic acid molecule and include any fragment or portion of an intact nucleic acid molecule.

[056] Nucleic acid hybridization is affected by such conditions as salt concentration, temperature, or organic solvents, in addition to the base composition, length of the complementary strands, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art. Stringency conditions depend on the length and base composition of the nucleic acid, which can be determined by techniques well known in the art. Generally, stringency can be altered or controlled by, for example, manipulating temperature and salt concentration during hybridization and washing. For example, a combination of high temperature and low salt concentration increases stringency. Such conditions are known to those skilled in the art and can be found in, for example, Strauss, W.M. "Hybridization With Radioactive Probes," *in* Current Protocols in Molecular Biology 6.3.1-6.3.6, (John Wiley & Sons, N.Y. 2000). Both aqueous and nonaqueous conditions as described in the art can be used.

[057] An example of stringent hybridization conditions is hybridization in 0.1×SSC (15 mM sodium chloride/1.5 mM sodium citrate) at 50°C or higher. Another example of stringent hybridization

conditions is hybridization overnight at 42°C in 50% formamide, 1X SSC (150 mM NaCl, 15 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5X Denhardt's solution, 10% (w/v) dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing in 0.1X SSC at about 65°C. Highly stringent conditions can include, for example, aqueous hybridization (*e.g.*, free of formamide) in 6X SSC (where 20X SSC contains 3.0 M NaCl and 0.3 M sodium citrate), 1% (w/v) sodium dodecyl sulfate (SDS) at 65°C for about 8 hours (or more), followed by one or more washes in 0.2 X SSC, 0.1% SDS at 65°C.

[058] Moderately stringent hybridization conditions permit a nucleic acid to bind a complementary nucleic acid that has at least about 60%, at least about 75%, at least about 85%, or greater than about 90% identity to the complementary nucleic acid. Stringency of hybridization is generally reduced by decreasing hybridization and washing temperatures, adding formamide to the hybridization buffer, or increasing salt concentration of the washing buffer, either individually or in combination. Moderately stringent conditions can include, for example, aqueous hybridization (*e.g.*, free of formamide) in 6X SSC, 1% (w/v) SDS at 65°C for about 8 hours (or more), followed by one or more washes in 2X SSC, 0.1% SDS at room temperature. Another exemplary hybridization under moderate stringency comprises hybridization in 6X SSC, 5X Denhardt's reagent, 0.5% (w/v) SDS, and optionally 100 μg/ml sonicated salmon or herring sperm DNA, at about 42°C, followed by washing in 2X SSC, 0.1% (w/v) SDS at 65°C. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the instant invention.

[059] In yet another embodiment, the invention provides a kit for diagnosing liver fibrosis and/or cirrhosis comprising one or more antibodies that specifically bind to at least one protein encoded by a sequence chosen from SEQ ID NO: 1 – SEQ ID NO: 63 and human orthologs thereof, and further comprises packaging that indicates use for detection of liver fibrosis and/or cirrhosis. In an alternative

embodiment, the kit comprises one or more antibodies that specifically bind to at least one human ortholog protein encoded by a sequence chosen from SEQ ID NO: 64 – SEQ ID NO: 120, and further comprises packaging that indicates use for detection of liver fibrosis and/or cirrhosis. Differential expression of these genes suggests the presence of fibrosis and/or cirrhosis.

[060] The antibodies specific to the proteins may be obtained by monoclonal or polyclonal techniques known to one skilled in the art. The antibody can be labeled with radioisotopes such as ^{35}S, ^{14}C, ^{125}I, ^{3}H, and ^{131}I, using the techniques described in *Current Protocols in Immunology* (20), as an example. The radioactivity can be measured using scintillation counting. The antibody may also be fluorescently labeled by, for example, rare earth chelates (europium chelates), Texas Red, rhodamine, fluorescein, dansyl, Lissamine, umbelliferone, phycocrytherin, phycocyanin, or commercially available fluorophores such SPECTRUM ORANGE® and SPECTRUM GREEN® and/or derivatives of any one or more of the above. The fluorescent labels can be conjugated to the antibody using the techniques disclosed in *Current Protocols in Immunology*, supra., for example. Fluorescence can be quantified using a fluorimeter.

[061] Differential protein expression may be assayed using antibodies with commonly used methods known in the art, such as Western blotting or enzyme-linked immunosorbent assay (ELISA). Western blotting begins with an electrophoresis step, where proteins from a biological sample of interest are separated on the basis of size and electromagnetic charge by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), according to standard methods known in the art. *See, e.g.,* SAMBROOK ET AL., 3 MOLECULAR CLONING: A LABORATORY MANUAL A8.40-A8.45 (2001) (describing various reagents and methods for electrophoresis of proteins by SDS-PAGE). The contents of the gel are then transferred to nitrocellulose, nylon, PVDF, or other membrane or filter suitable for fixation and Western blotting by standard methods also known in the art. The transfer may be by immersion, semi-dry blotting, or by other comparable methods known in the art. Next, the filters or membranes are fixed to prevent loss of the target proteins during the several hybridization, washing, and staining steps

comprising Western blotting. Fixation may be accomplished by heat, cross-linking with ultraviolet light, or by other comparable methods known in the art. *See, e.g.*, SAMBROOK ET AL., 3 MOLECULAR CLONING: A LABORATORY MANUAL A8.52-A8.55 (describing various reagents and methods for immunoblotting and detection of antigen/antibody complexes).

[062] Non-specific anitbody binding sites on the fixed filter or membrane are blocked with buffered solutions (*e.g.*, phosphate-buffered saline ("PBS") or the like) containing a blocking agent such as, for example, 0.5% (w/v) low-fat dry milk or 5% (w/v) bovine serum albumin (BSA). After blocking, the filter or membrane then undergoes the primary antibody incubation. After primary antibody incubation, the filter or membrane is washed, and the presence of antibody-antigen complexes detected using a secondary antibody labeled with chromogenic, fluorogenic, or chemiluminescent means. Antibody-antigen complexes are then detected colorimetrically (*e.g.*, with horseradish peroxidase and TMB), or by autoradiography (*e.g.*, alkaline phosphatase). If detected colorimetrically, or by chemiluminescence, the amount of color of fluorescence may be measured using a luminometer, a spectrophotometer, or other similar instruments. If detected autoradiographically, the amount of bound antibody may be measured from the exposed x-ray film using a densitometer, or similar instrument. *See, e.g.*, SAMBROOK ET AL., 3 MOLECULAR CLONING: A LABORATORY MANUAL A8.52-A8.55.

[063] Secondary antibodies used in Western blotting, whether polyclonal or monoclonal, may be labeled with a ligand (such as biotin) or a detectable marker (such as a fluorescent group or an enzyme) using conventional techniques. Suitable labels include fluorophores, chromophores, electron-dense reagents (*e.g.*, silver or gold), enzymes, and ligands having specific binding partners. Enzymes such as horseradish peroxidase or alkaline phosphatase are typically detected by their activity. For example, horseradish peroxidase can be detected by its ability to convert tetramethylbenzidine (TMB) to a blue pigment, quantifiable with a spectrophotometer. Other suitable ligands and/or detectable markers include bioting and avidin or streptavidin, IgG and protein A, and the numerous additional receptor-ligand

couples known in the art. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the instant invention.

[064] An ELISA begins with an antigen adsorption step, where the target antigen or antigens are adsorbed to the wells of a microtiter plate. *See, e.g.*, KIERKEGAARD & PERRY LABORATORIES, INC., TECHNICAL GUIDE FOR ELISA 9-13 (2003), *available at* http://www.kpl.com/docs/techdocs/chapters%201%20-%204.pdf (last accessed January 12, 2007). The most commonly used adsorption buffers for antibodies are 50 mM Carbonate, pH=9.6; 10 mM Tris-HCl, pH=8.5; and 10 mM PBS, pH=7.2. These buffers work well for many proteins. If the target antigens are not readily adsorbed to the surface of the microtiter plate, plates with surfaces modified or derivatized to permit covalent linkage of proteins to their surface by a variety of chemical means are widely available from commercial suppliers. Time and temperature are the most important factors affecting the amount of protein adsorbed.

[065] Once the wells of a microtiter plate are coated with the desired antigen or antigens, they are washed with a blocking buffer to block non-specific antibody binding and to minimize false positive results. *See, e.g., id.* at 13-14 (discussing methods and reagents for blocking microtiter plates). Commonly used blocking agents are either protein solutions, such as BSA (typically used at concentrations between 1% and 5% (w/v) in PBS, pH=7.0), non-fat dry milk, casein (the main protein component of non-fat dry milk), or caseinate (a more soluble version of casein, produced by partial digestion with sodium hydroxide), normal serum (typically used at concentrations between 1% and 5% (v/v)), and gelatin (normally used at concentrations between 1% and 5% (w/v)), or non-ionic detergents, such as Tween-20™ and Triton X-100™.

[066] Washing reagents are selected for their ability to disrupt low-affinity interactions between various reaction components that can affect the ability to detect specific antigen-antibody interactions. *See, e.g., id.* at 14-15 (discussing methods and reagents for washing microtiter plates). Wash solutions commonly contain a physiological buffer to prevent denaturation of antigens and their cognate

antibodies, and to preserve enzyme activity. Buffers such as PBS, Tris-saline, or imidizole-buffered saline at neutral pH are widely used. Specific buffers are typically selected based on the method of detection to be employed in a particular assay. Wash buffers should also include non-ionic detergents such as Tween-20™, Triton X-100™, or the like, at concentrations of between 0.01% to 0.05% (v/v), in order to disrupt low-affinity, non-specific interactions between reaction components.

[067] After the blocking step, the wells of the microtiter plate are washed. The adsorbed antigen then undergoes the primary antibody incubation, after which it is washed again. Antibody/antigen complexes are then detected using a secondary antibody labeled with chromogenic (*e.g.*, horseradish peroxidase and TMB), fluorescent or chemiluminescent (*e.g.*, alkaline phosphatase) means. *See, e.g., id.* at 15-21 (discussing antibody preparation and use, as well as commonly used detection molecules). The amount of color or fluorescence may be measured using a luminometer, a spectrophotometer, or other similar instruments. There are many common variations on the standard ELISA protocol, including competitive ELISA, sandwich ELISA, and numerous others. One of ordinary skill in the art will select the appropriate protocol to use, depending on the antigen to be detected, the source of antigen and/or primary antibody used in the assay, and any other relevant experimental parameters. These and many other permutations will be readily apparent to those of ordinary skill in the art, are considered as equivalents within the scope of the invention.

Screening Methods

[068] Also described is a method that identifies a compound that would decrease the differential expression of at least one gene encoding a protein chosen from SEQ ID NO: 7 – SEQ ID NO: 23 and SEQ ID NO: 42 – SEQ ID NO: 63 and human orthologs thereof. The method provides a cell expressing at least one of these genes, contacting the cell with a test compound to determine whether the differential expression is decreased in the presence of the test compound, wherein the differential expression is an indication of halting or reversing liver fibrosis and/or cirrhosis. The same method can also identify a compound that would increase the differential expression of at least one gene encoding a protein chosen

from SEQ ID NO: 1 – SEQ I D NO: 6 and SEQ ID NO: 24 – SEQ ID NO: 41 and human orthologs thereof.

[069] Other embodiments of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with the true scope and spirit of the invention being indicated by the following claims.

[070] Unless defined otherwise, the meanings of all technical and scientific terms used herein are those commonly understood by one of ordinary skill in the art to which this invention belongs. One of ordinary skill in the art will also appreciate that any methods and materials similar or equivalent to those described herein can also be used to practice or test the invention. Further, all publications mentioned herein are incorporated by reference.

[071] With respect to ranges of values, the invention encompasses the upper and lower limits and each intervening value between the upper and lower limits of the range to at least a tenth of the upper and lower limit's unit, unless the context clearly indicates otherwise. Further, the invention encompasses any other stated intervening values.

[072] Further, all numbers expressing quantities of ingredients, reaction conditions, percent purity, polypeptide and polynucleotide lengths, and so forth, used in the specification and claims, are modified by the term "about," unless otherwise indicated. Accordingly, the numerical parameters set forth in the specification and claims are approximations that may vary depending upon the desired properties of the present invention. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should at least be construed in light of the number of reported significant digits, applying ordinary rounding techniques. Nonetheless, the numerical values set forth in the specific examples are reported as precisely as possible. Any numerical value, however, inherently contains certain errors from the standard deviation of its experimental measurement.

[073] It must be noted that, as used herein and in the appended claims, the singular forms "a," "or," and "the" include plural referents and plural referents include singular forms unless the context clearly dictates otherwise. Thus, for example, reference to "a subject polypeptide" includes a plurality of such polypeptides, reference to "the agent" includes reference to one or more agents and equivalents thereof known to those skilled in the art, reference to "nucleic acid molecules" includes reference to one or more nucleic acid molecules, and reference to "antibodies" includes reference to one or more antibodies and so forth.

[074] The following examples further illustrate the invention. They are merely illustrative of the invention and disclose various beneficial properties of certain embodiments of the invention. The following examples should not be construed as limiting the invention.

EXAMPLES

[075] The practice of the present invention will employ, unless otherwise indicated, conventional techniques for the study of differential gene expression, differential proteomic expression, and histopathology, which are within the level of ordinary skill in the art. Such techniques are explained fully in the applicable literature.

[076] The following examples illustrate the study of the expression of different proteins in liver fibrosis through the use of animal models.

Example 1: *Establishment of the Rat Hepatic Fibrosis Model*

[077] A rat hepatic fibrosis model was set up through the use of DMN. DMN is a potent hepatoxin that specifically targets the liver and can cause liver fibrosis. The histopathological changes resulting from DMN treatment involve the rapid deposition of collagen, the major protein of fibrosis and important to the process of cirrhosis. DMN-induced liver injury in rats displays many features, such as portal hypertension ascites as well as a number of other histopathological and biochemical abnormalities. Previous study has suggested that DMN-induced liver injury in rats could reflect changes that occur in human hepatic fibrosis and/or cirrhosis and that it is an appropriate animal model for studying human

hepatic fibrosis and/or cirrhosis (15). The DMN-induced liver fibrosis model was performed as described in Jezequel A.M. *et al.* (14). Male Sprague–Dawley rats (Slc:SD; Japan SLC, Shizuoka, Japan), weighing from 300 to 350 grams, were used in the experiments. To induce hepatic fibrosis over a six-week time course experiment, the rats were given DMN (Sigma, Saint Louis, MO), dissolved in normal saline, three consecutive days a week for the first three consecutive weeks at the dosage of 6.7 mg/kg per body weight by intraperitoneal injection. Injection time points are shown as inverted triangles in Figure 1. Injections were at a much lower dosage than those used in other experiments where the dosage was 100 mg/kg/day. The higher dosage can cause toxicity in rat liver (21, 22).)

[078] Two to seven rats at each time point were treated with either DMN or with an equal volume of normal saline without DMN as the control (26 DMN-treated rats and 24 control rats). Rats were weighed and sacrificed on days 11, 18, 25, 32, 39, and 46 (sacrifice time points are shown as triangles in Figure 1A) respectively. These triangular time points were designated as weeks 1 through 6 (Figure 1A).

[079] To confirm the establishment of a rat liver fibrosis model, Q-RT-PCR, serum analysis, and histopathological analysis were employed.

Histopathological Examination of DMN-Induced Liver Damage

[080] Liver tissues were immediately removed after sacrifice and subject to histopathological examination. The fixed liver samples were then processed for paraffin embedding. Five micrometer (5 μm) sections were prepared for hematoxylin and eosin staining (to score necroinflammatory and fatty changes) and for Sirius red/fast green collagen staining (to score for fibrosis) (23). The fatty changes were classified as presence (+) or absence (-).

[081] Phenotypic changes resulting from DMN-induced liver damage are shown in Figure 2A. In Figure 2A, the representative phenotypes of the DMN-induced rat liver fibrosis were characterized by scoring the four histopathological features (necrosis, inflammation, fibrosis, and fatty change) as follows: the necrosis scores were from N0 to N3 (the first panel), and the inflammation scores were from I0 to I3.

The scoring system for examining the intensity of liver fibrosis was modified from the scoring system of the Hepatitis Activity Index (HAI) (29, 30). Fibrosis was divided into four scores: normal (F0), fibrous expansion of portal tracts (F1), bridging fibrosis (F2) and frequent bridging fibrosis with focal nodule formation (F3). The fatty change was scored as present (+) or absent (-), respectively (the last panel). The images of the fatty change are shown at 200X magnification, whereas the others are shown at 100X magnification (Fig. 2A).

[082] Figure 2B shows a summary of histopathological scores for the rat model. The results were ranked by time course. The necroinflammatory scores were the sum of the necrosis and inflammation scores and range from A0 to A6. The necroinflammatory change was divided into three grades: A0= none, A(1-3)= mild and A(4-6)= moderate necroinflammation. The fibrosis change was divided into two grades: F(0-1)= normal to fibrous expansion of portal tracts, F(2-3)= bridge fibrosis to frequent bridging fibrosis with nodule formation. The fatty change was shown as presence (+) or absence (-). The number of rats was counted and used to calculate the percentage of each histopathological level at each time point.

[083] As shown in Figure 2B, seventy-five percent of the DMN-treated rats had none (F0) or low levels of fibrosis (F1) in the first two weeks. By the third to fourth week, nearly 90% of the DMN-treated rats had high levels of fibrosis, from bridging fibrosis (F2) to frequent bridging fibrosis with focal nodule formation (F3). In the last two weeks, F2 and F3 were still present in 78% of DMN-treated rats. The fatty changes were only present in a few treated rats (3.7%). In contrast, there were no abnormal pathological patterns present in the control group at all. In addition, no clear abnormality was found in the kidney or spleen of the DMN-treated and normal rats (data not shown). Together, the detailed necroinflammatory and fibrosis scoring systems of the process of the DMN-induced liver damage suggest that dramatic necrosis and inflammation took place during early liver damage progression (week 1-4), followed by fibrosis formation at 3-6 weeks. Collagen fiber deposition in rat liver could be

observed, along with bile duct proliferation, centrilobular necrosis, bridge fibrosis and fibrosis surrounding the central veins, after three weeks of DMN treatment.

Quantitative Real-Time Polymerase Chain Reaction (Q-RT-PCR)

[084] To gain additional information about the established animal model, Q-RT-PCR was used to evaluate the gene expression profile of transforming growth factor-β1 (*Tgfβ1*), which is the strongest known inducer of fibrogenesis in the effector cells of hepatic fibrosis that can stimulate the adipocyte transformation (24-27). The same total RNA samples were used for both microarray and Q-RT-PCR analyses. RNA preparation and analysis were performed according to Affymetrix's (Santa Clara, CA, U.S.A.) instructions. The TaqMan® assays were conducted in triplicate for each sample, and a mean value was used for calculation of expression levels. To standardize the quantification of the target genes, 18S ribosomal RNA (18S rRNA) from each sample was quantified on the same plate with the target genes. The Q-RT-PCR result showed that a higher level of mRNA expression of *Tgfβ1* was observed in DMN-treated rat livers than in the controls (Figure 1B). The changes were in agreement with observations described in Jezequel A.M. *et al.* (14). In sum, these examinations support the DMN-induced rat hepatic fibrosis model.

Analysis of Serum Biochemical Data from the DMN-Induced Hepatic Fibrosis Rat Model Experiment

[085] The serum of each rat, 50 rats in total, was subjected to various biochemical examinations related to liver damages. Blood samples, collected from the animals at necropsy, were used to measure serum concentrations or activity of albumin, glutamic oxaloacetic transaminase (GOT), glutamic pyruvic transaminase (GPT), total bilirubin, alkaline phosphatase (AKP), acid phosphatases (ACP), α-fetoprotein (AFP), blood urea nitrogen (BUN), cholesterol (CHOL), lactate dehydrogenase (LDH), globulin, prothrombin time (PT) and blood platelets (PLT) using an Hitachi 747 and ACL 3000 clinical chemistry analyzer system (MYCO, Renton, Washington). When the DMN-treated samples and controls were compared, thirteen serum markers showed significant differences and confirmed the DMN-treated group

had suffered liver damage. The biochemical data for the DMN-treated group suggest that there were changes in many serum markers and that the protein expression levels or physical responses are similar to human liver damage phenotypes (1, 2). The thirteen serum markers identified in these experiments are the same serum markers for human liver fibrosis.

Example 2: *Establishment of Gene Expression Profile During DMN-Induced Liver Fibrosis*

Microarray Analysis

[086] The quality of the total RNA for microarray analysis was determined using Spectra Max Plus (Molecular Devices, Sunnyvale, CA, U.S.A.) and had an A_{260}/A_{280} ratio ranging from 1.9 to 2.1. Protocols and reagents for hybridization, washing and staining followed Affymetrix's instructions (http://www.affymetrix.com/support/technical/manuals.affx). Labeled cRNA was hybridized to the Affymetrix GeneChip Test 3 Array to verify the quality prior to hybridization to the Affymetrix Rat Genome U34A Array.

Data Analysis and Clustering Algorithm

[087] The microarray images were transformed into text files containing intensity information using GeneChip® Operating Software (GCOS) developed by Affymetrix. The microarray datasets were then analyzed using GeneSpring® 7.2 software (Silicon Genetics, Redwood City, CA, U.S.A.).

Gene Expression Profiling

[088] Over the six week time course of the experiment, the liver tissues of 12 control animals and 12 DMN-treated rats (2 rats for each time point) were selected to perform microarray experiments. Before any statistical analyses were applied to the microarray data, reproducibility was assessed. Genes were selected as present when they were assigned a present call according to the perfect match (PM)/mismatch (MM) algorithm of Affymetrix in all gene chips (28). Of the 8799 probe sets analyzed, overall expression patterns for 2385 transcripts on the chips were reported to be present ($P< 0.05$). To verify that intra-sample variability did not obscure differences between the controls and DMN-treated groups, as well as to determine the fold-change that should be considered as significant, expression

profiles among the 24 control datasets were compared. Scattered graphs of expression levels of the 2385 transcripts represented on the microarray were compared with each other. The relationship between the experimental chips in the microarray analysis was analyzed by linear regression. Overall, there was no statistical difference, with 3.2 % of transcripts deviating more than 2-fold. To investigate the time course variability, the reliable signals of these 2385 probe sets between the first and sixth week of controls, were calculated. Again, there was no statistical difference, with 4.6 % of transcripts deviating more than 2-fold. In contrast, a significant scatter was found between controls and DMN-treated groups, with 28.7 % of transcripts deviating more than 2-fold.

[089] As a first step to minimize the likelihood of false positives, all transcripts were filtered by forming two independent clusters from the microarray data and identifying those that were differentially expressed. For detailed analysis, the first cluster generated 2385 transcripts as previously described. Of these, 268 were differentially expressed transcripts either higher or lower by 1.5 fold or more when comparing the controls and DMN-treated groups. The second method, which used the "detection flag" selection (24), reported 23 transcripts as "present" in the DMN-treated groups but not in the controls. In contrast, there was only one transcript reported "absent" in all DMN-treated groups but "present" in the controls.

[090] Altogether, 292 transcripts representing 256 genes, including 137 up-regulated and 119 down-regulated genes, exhibited differentially expressed gene expression patterns when the DMN-treated groups and controls were compared. Hierarchical clustering generated a dendrogram for the gene expression patterns of these 292 transcripts across the 24 samples as shown in Figure 3A. The rows represent individual transcripts and columns represent time course samples. The color in each cell reflects the expression level of the corresponding tissue, relative to its mean expression level. The scale extends from fluorescence radios of 0.25 to 4 relative to the mean level for all samples.

Quantitative-Real-Time-Reverse-Transcriptase Polymerase Chain Reaction (Q-RT-PCR)

[091] To validate the microarray data, Q-RT-PCR analysis was performed for tissue inhibitor of metalloproteinase 1 (*Timp1*), tissue inhibitor of metalloproteinase 2 (*Timp2*), matrix metalloproteinase 3 (*Mmp3*) and gamma-glutamyl transpeptidase (*Ggtp*). These genes were chosen for validation because these genes occurred both in this GeneChip study and in previous studies.

[092] As determined by Q-RT-PCR, *Timp1*(Figure 3B), *Timp2*, *Mmp3*, *Ggtp* (data not shown) and *Tgfβ1* (shown previously in Figure 1B) were elevated in DMN-treated samples. The results of Q-RT-PCR analysis of these five genes were consistent with previous reports examining these individual markers (8, 9).

[093] Moreover, a good concordance based on the fold changes between microarray data and Q-RT-PCR results was observed. As shown in Figure 3B, *Timp1* expression was elevated over 20-fold in DMN-treated rats in both microarray and Q-RT-PCR experiments. TaqMan® assays were conducted in triplicate for each sample, and a mean value was used for calculation of expression levels (marked by the square). To standardize the quantification of the *Timp1* transcript, 18S rRNA from each sample was quantified on the same plate as the target gene, as indicated by the log scale on the right side of plot. The expression levels of the two *Timp1* transcripts, rc_AI169327_at and rc_AI169327_g_at (marked by a circle and a triangle, respectively), were set forth relative to the mean of all gene expression levels as indicated by the scale on the left side of plot. The expression pattern of *Timp1* was highly correlated between the Q-RT-PCR results and the GeneChip analysis (the Pearson's correlation coefficients were 0.79 and 0.92, respectively) (Figure 3B), suggesting that the gene expression results were reliable when subject to more detailed analysis.

Fibrosis Candidate Genes

[094] Necroinflammatory and fibrosis have been suggested to play important roles in the progression of liver cirrhosis in the rat model (8, 23, 24, 29, 30). To clarify the factors responsible for the histopathological phenotype, all rat samples were classified by histopathological evaluation with

histopathological scores for necroinflammation (A0-A6) and fibrosis (F0-F3) as described earlier in Figure 2B. The student *t*-test statistic analysis was used for fibrosis-related genes analysis as it was based on the two-subgroup (F(0-1) and F(2-3)) variation in fibrosis score. A *P*-value of less than 0.05 was considered to be statistically significant.

[095] Using the student's *t*-test, the inventors analyzed those 256 genes with expression either higher or lower by 1.5 fold or more in DMN-induced rats than in control rats. A total of 62 differentially expressed genes (32 up regulated and 30 down regulated) between the F(0-1) and F(2-3) level of fibrosis were identified, estimated using only two subgroup variations for the fibrosis score, at the 5 % significance level. In agreement with previous studies, three genes, including *Timp1*, CD63 and annexin A1 (*Anxa1*), exhibited similar gene expression patterns during liver fibrosis (31-34).

[096] *Timp1* is a well-known fibrosis marker and has been proven to play a significant role in the progression of liver fibrosis. In the later stages of liver injury, hepatic stellate cells (HSCs) express a combination of matrix metalloproteinases (MMPs) that have the ability to degrade normal liver matrix, while inhibiting degradation of the fibrillar collagens that accumulate in liver fibrosis. An increase in expression of *Timp1* leads to a more global inhibition of degradation of fibrillar liver collagens by interstitial collagenases (MMP-1/MMP-13) (35). As shown in Fig. 3B, expression of *Timp1* was elevated more than twenty fold in DMN treated rats in microarray.

[097] CD63, a transmembrane protein, one of the fibrosis gene signatures observed, is also up-regulated after DMN treatment. Following chronic injury, HSCs activate or differentiate into myofibroblast-like cells, acquiring contractile, proinflammatory, and fibrogenic properties. Activated HSCs migrate and accumulate at the sites of tissue repair, secreting large amounts of extracellular matrix proteins during the progression of fibrosis. Activated HSCs have been identified as major collagen-producing cells (collagen is an extracellular matrix protein) and an initiator of liver fibrosis when the liver is injured (7). It has been demonstrated that inhibition of CD63 might be a novel diagnostic marker for the injured liver.

[098] Annexin A1 (*Anxa1*) is highly expressed after liver injury in an alcoholic liver disease (ALD) study (34). Alcohol initiated liver injury occurs via inflammation. ALD progression involves continuing liver injury, fibrosis, and impaired liver regeneration. It has been suggested that *Anxa1* might play a role in the progression of fibrosis.

[099] The fact that the present study agrees with previous studies for the three genes as described above indicated that oligonucleotide microarray analysis is a powerful approach for monitoring molecular events during liver injury and repair where the pathogenesis is unknown. Also, the signature genes identified in this study could discriminate successfully between the low-score and the high-score histopathology groups. Together, the genes would seem to be responsible for fibrosis formation and are possible markers for the detection of fibrosis.

Example 3: *Protein Expression Profiling During DMN-Induced Liver Fibrosis*

[0100] Experiments using iTRAQ tagging were performed to identify differentially expressed proteins in liver fibrosis.

Isobaric Tagging Method : Sample Preparation and Reagent Labeling

[0101] The liver tissues of 6 control and 6 DMN-treated rats were used to perform duplicate iTRAQ labeling experiments. Liver tissue was frozen in liquid nitrogen and then pulverized to powder using a mortar and pestle pre-cooled with liquid nitrogen. Then, the liquid nitrogen was allowed to evaporate, and the powdered tissue was separated into tubes and stored at –80°C. Twenty volumes (w/v) of lysis buffer (containing 2% (w/v) SDS in 20mM phosphate buffer at pH 7.6) was added to the powdered tissue and incubated at room temperature for 1 hour. Samples were centrifuged at 12000 rpm for 10 minutes and the supernatant was taken for acetone precipitation by adding six volumes of cold acetone to the sample tubes. Tubes were incubated at –20°C until precipitate formed. After decanting the acetone, pellets were resuspended with 0.1% SDS and 6M urea in dissolution buffer (provided by iTRAQ Reagents Kit, Applied Biosystems, Foster City, CA, U.S.A.).

[0102] The total protein contents were determined using the Coomassie Plus Protein Assay Reagent (PIERCE, Rockford, IL, U.S.A.). Two hundred (200) μg of proteins were taken and diluted by an equal volume (w/v) of dissolution buffer containing 0.1 % (w/v) SDS. Proteins were reduced and cysteines blocked as described in the iTRAQ protocol (Applied Biosystems, Foster City, CA, U.S.A.). Four volumes of dissolution buffer were added to each tube and incubated with 20 μg Trypsin at 37°C for 16 hours. Tryptic peptides extracted from liver tissue of controls and DMN-treated rats were labeled with iTRAQ 114 and 115, respectively, at each time point of the experiment.

Peptide Separation and Analysis

[0103] Peptides with different isobaric tags were pooled and acidified by mixing with 10 mM phosphoric acid (in 25% (v/v) acetonitrile and 75% H_2O) to a total volume of 4.0 mL for strong cation exchange (SCX) chromatography. The resulting sample was injected into a liquid chromatography system (Ettan, GE Healthcare Bio-Science, Umea, Sweden) using a 2.1 mm × 200 mm Polysulfoethyl A column packed with 5 micron 300 Å beads (PolyLC, Columbia, MD, U.S.A.) at a flow rate of 0.08 mL/min. A guard column of the same material was plumbed upstream from the analytical column. The buffers used were 10 mM KH_2PO_4, 25% (v/v) acetonitrile at pH 3.0 for buffer A and 10 mM KH_2PO_4, 1M KCl, 25% (v/v) acetonitrile at pH 3.0 for buffer B. The elution gradient was changed linearly from 0 to 40% buffer B within 16 mL and then up to 100% buffer B in another 4 mL. A total of 30 fractions were collected and dried by speed vacuuming centrifugation. All the fractions were desalted on pepClean™ C-18 spin columns (PIERCE, Rockford, IL, U.S.A.). The desalted peptide from each SCX fraction was dried by speed vacuuming centrifugation and then analyzed by nanoLC hybrid mass spectrometry. The nanoLC system was from LC Packings (Amsterdam, The Netherlands), coupled to an API QSTAR Pulsar Hybrid QqTOF mass spectrometer (Applied Biosystems/MDS Sciex, Foster City, CA, U.S.A.). Peptide separation was performed using a reversed-phase C_{18} column (75 μm in diameter x 15 cm in length, 3 μm particles) with a two-step linear gradient of 5-50% Buffer A over 45 minutes and 50-95% Buffer B over 10 minutes at a flow rate of 0.2 μl/minute (Buffer A: 2% (v/v) acetonitrile, 0.1%

(w/v) formic acid; Buffer B: 80% (v/v) acetonitrile, 0.1% (w/v) formic acid). The m/z scanning range for MS and tendem MS (MS/MS) were 400-1200 and 75-1500 Da, respectively. MS/MS mass spectra were analyzed in continuous flow mode with a 10 mm I.D. fused silica tip. MS/MS spectra were analyzed against rat.fasta databases from NCBI by ProQUANT 2.0 (Applied Biosystems/MDS Sciex, Foster City, CA, U.S.A.). The accuracy tolerance for peptide identification was 0.5 Da for MS and 0.5 Da for MS/MS, respectively. The cut off for the confidence settings was at 95 and for the score was at 20. Relative quantification of peptide was performed on the MS/MS scans, using ratio of the areas under the peaks. Relative quantification of protein was obtained by averaging the constituted peptides identified.

[0104] To validate the method and results obtained from iTRAQ proteomic studies, the expression changes of vimentin quantified by mass spectrometry in iTRAQ experiments were compared with those obtained through Western blot with anti-vimentin antibody. Proteins were extracted with 2% (w/v) SDS in 20 mM phosphate buffer at pH 7.6 from control and DMN-induced rat livers. After resolving by 12% SDS-PAGE, proteins were transferred to polyvinylidene difluoride membranes and subsequently immunostained with mouse anti-vimentin monoclonal antibody (Chemicon, Temecula, CA, U.S.A.) for 1 hour, then with horseradish peroxidase-labeled goat anti-rat IgG for another hour. The immunoreactive bands were detected using an enhanced chemiluminescence system (ECL, Perkin-Elmer Life Sciences, Wellesley, MA, U.S.A.). As demonstrated in Figure 4, the results observed in Western blots matched the trend as measured by mass spectrometry.

Results

[0105] The numbers of distinct peptides identified range from 624-1591 when the confidence threshold was set at 95%. A significant number of these peptides were identified more than once. More than 351 unique proteins were identified in each experiment. Thirty-nine proteins show significant and consistent differences (1.5 fold or more) in protein expression in DMN-treated liver tissue with F2 and F3 fibrosis score compared to control rats. Among these 39 proteins, several have previously been reported to associate with fibrosis formation, including, fibrinogen, and fibronectin.

[0106] Fibrinogen was found to be up-regulated in DMN-induced rats in the present study. This is consistent with the results of fibrinogen gene expression in a model CCl_4-induced rat liver damage. The CCl_4-induced rat liver damage also showed increased fibrinogen mRNA levels and fibrinogen/fibrin deposition during short-term liver injury and liver fibrogenesis. This may suggest that fibrinogen involves a "clotting-like process" in short term liver damage and liver fibrosis (36). Fibronectin consists of two polypeptide chains. It mediates adhesion of collagen, fibrin and heparin to cells and is thus involved in the organization of thrombi and in wound healing by inducing attachment of these structures to cells. Fibronectin and collagen types I and III have been used as a characteristic feature of a liver cirrhotic state in the CCl_4-induced liver cirrhosis rat model (37).

[0107] In addition, glycine N-methyltransferase (*Gnmt*), aldolase A(*Aldoa*), Myosin light polypeptide 6 (*Myl6*) and, cytoplasmic γ actin, (*Actγ*)) were identified in both RNA expression microarray and proteomic studies as exhibiting differential expression.

Example 4: *Network Analysis*

[0108] To further refine the genes with differential expression in the fibrotic liver, we carried out network analysis using the Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems, Mountain View, CA, U.S.A.). The genes, which were identified by microarray and iTRAQ proteomic analyses, were subjected to interacting network analyses. Gene accession numbers were imported into IPA software. These networks described functional relationships between gene products based on known interactions in the literature. The IPA tool then associated these networks with known biologic pathways. Of sixty-two genes identified from the microarray study, forty-two genes fall into 3 groups, i.e., gene networks associated with 1) cancer, cell morphology, and dermatological diseases and conditions; 2) lipid metabolism, small molecule biochemistry, molecular transport; and 3) cancer, cellular movement, cellular growth and profileration. *See* Table 3A. Twenty-two out of thirty-nine proteins identified in the proteomic study fall into another 3 networks associated with 1) cancer, cell cycle and cell morphology; 2) lipid metabolism, small molecule biochemistry, organismal injury and abnormalities;

and 3) hematological disease, endocrine system development and function, nervous system development and function. *See* Table 3B. Table 3A and 3B list the networks' protein functions and members. Genes identified in the present study are underlined. The underlined genes that have not been reported to associate with liver injury are shown in bold.

Example 5: *Clinical Validation of Biomarkers Identified in the DMN-Induced Liver Fibrosis Rat Model with Human Samples*

[0109] To confirm that the biomarkers identified in the DMN-induced liver fibrosis rat model could be used to diagnose liver fibrosis or cirrhosis in humans, we examined expression of the enzyme carbonic anhydrase I (*see* Table 2, SEQ ID NOS: 11 and 71) in human sera by Western blot analysis. Serum samples (8 μl from each patient) from normal volunteers and from patients suffering from hepatitis B-related cirrhosis were first treated to remove the twenty high abundance serum proteins using the ProteoPrep® 20 Plasma Immunodepletion Kit (Sigma-Aldrich, St. Louis, MO, U.S.A.). After depletion, the low abundance serum proteins were precipitated with 5 volumes of cold acetone at -20°C for 2 hours. The protein/acetone mixture was then centrifuged at 12,000 rpm for 10 minutes. The protein pellet was resuspended, separated by electrophoresis on a 12% SDS-PAGE gel, and transferred to 0.45 μm polyvinylidene difluoride (PVDF) membranes (Millipore, Billerica, MA, U.S.A.). After blocking in 5% (w/v) nonfat milk, the membranes were washed with TBST buffer (20 mM Tris-HCl, pH 7.6, 137 mM NaCl, 0.1% (v/v) Tween-20) and incubated with goat polyclonal antibody specific for carbonic anhydrase I (Abcam, Cambridge, UK) for 1 hour. The membranes were washed again with TBST buffer and immunostained with horseradish peroxidase-conjugated donkey anti-goat IgG (Jackson ImmunoResearch Laboratories, PA, U.S.A.). The immunoreactive bands were visualized using an enhanced chemiluminescence system (ECL, Perkin-Elmer Life Sciences, Wellesley, MA, U.S.A.). The Western blots clearly show that some serum samples from the cirrhosis patients have higher concentrations of carbonic anhydrase I than samples from normal individuals not suffering from

cirrhosis. This result confirms that the markers identified in the rat model for liver fibrosis and/or cirrhosis are also associated with liver fibrosis and/or cirrhosis in humans.

References

[0110] The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates, which may need to be independently confirmed.

1. Friedman, S.L. Molecular regulation of hepatic fibrosis, an integrated cellular response to tissue injury, *J. Biol. Chem.* 275:2247-2250 (2000).

2. Friedman, S.L. Seminars in medicine of the Beth Israel Hospital, Boston. The cellular basis of hepatic fibrosis. Mechanisms and treatment strategies, *N. Engl. J. Med.* 328:1828-1835 (1993).

3. Iredale, J.P. Cirrhosis: new research provides a basis for rational and targeted treatments, *BMJ* 327:143-147 (2003).

4. Okita, K., Sakaida, I., and Hino, K. Current strategies for chemoprevention of hepatocellular carcinoma, *Oncology* 62 Suppl. 1:24-28 (2002).

5. Iizuka, N., Oka, M., Yamada-Okabe, H., Mori, N., Tamesa, T., Okada, T., and Takemoto, N. Differential gene expression in distinct virologic types of hepatocellular carcinoma: association with liver cirrhosis, *Oncogene* 22:3007-3014 (2003).

6. Day, C.P. Non-alcoholic steatohepatitis (NASH): where are we now and where are we going?, *Gut* 50:585-588 (2002).

7. Bataller, R., and Brenner, D.A. Liver fibrosis, *J. Clin. Invest.* 115:209-218 (2005).

8. Friedman, S.L. Liver fibrosis - from bench to bedside, *J. Hepatol.* 38 Suppl. 1:S38-53 (2003).

9. Hayasaka, A., and Saisho, H. Serum markers as tools to monitor liver fibrosis, *Digestion* 59:381-384 (1998).

10. Hippo, Y., Taniguchi, H., Tsutsumi, S., Machida, N., Chong, J.-M., and Fukayama, M. Global gene expression analysis of gastric cancer by oligonucleotide microarrays, *Cancer Res.* 62:233-240 (2002).

11. Ji, J., Chen, X., Leung, S. Y., Chi, J.-T. A., Chu, K. M., and Yuen, S. T. Comprehensive analysis of the gene expression profiles in human gastric cancer cell lines, *Oncogene* 21:6549–6556 (2002).

12. Wright, M.E., Han, D.K., and Aebersold, R. Mass spectrometry-based expression profiling of clinical prostate cancer, *Mol. Cell. Proteomics* 4(4):545-54 (2005).

13. Nielsen, P.E., Egholm, M., Berg, R.H., and Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide, *Science.* 254(5037):1497-500 (1991).

14. Jezequel, A.M., Mancini, R., Rinaldesi, M.L., Macarri, G., Venturini, C., and Orlandi, F. A morphological study of the early stages of hepatic fibrosis induced by low doses of dimethylnitrosamine in the rat, *J. Hepatol.* 5:174-181 (1987).

15. George, J., Rao, K.R., Stern, R., and Chandrakasan, G. Dimethylnitrosamine-induced liver injury in rats: the early deposition of collagen, *Toxicol.* 156:129-138 (2001).

16. DeSouza, L., Diehl, G., Rodrigues, M.J., Guo, J., Romaschin, A.D., Colgan, T.J., and Siu, K.W. Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry, *J. Proteome Res.* 4(2):377-86 (2005).

17. Mirsalis, J.C., and Butterworth, B.E. Detection of unscheduled DNA synthesis in hepatocytes isolated from rats treated with genotoxic agents: an in vivo- in vitro assay for potential carcinogens and mutagens, *Carcinogenesis* 1:621-625 (1980).

18. Haggerty, H.G., and Holsapple, M.P. Role of metabolism in dimethylnitrosamine-induced immunosuppression: a review, *Toxicol.* 63:1-23 (1990).

19. Ala-Kokko, L., Pihlajaniemi, T., Myers, J.C., Kivirikko, K.I., and Savolainen, E.R. Gene expression of type I, III and IV collagens in hepatic fibrosis induced by dimethylnitrosamine in the rat, *Biochem. J.* 244:75-79 (1987).

20. Coligen, J.E.. *Current Protocols in Immunology*, Volumes 1 and 2, (Wiley-Interscience, New York, N.Y., 1991).

21. Waring, J.F., Jolly, R.A., Ciurlionis, R., Lum, P.Y., Praestgaard, J.T., Morfitt, D.C., and Buratto, B. Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles, *Toxicol. Appl. Pharmacol.* 175:28-42 (2001).

22. Waring, J.F., Ciurlionis, R., Jolly, R.A., Heindel, M., and Ulrich, R.G. Microarray analysis of hepatotoxins in vitro reveals a correlation between gene expression profiles and mechanisms of toxicity *Toxicol. Lett.* 120:359-368 (2001).

23. Lopez-De Leon, A., and Rojkind, M.. A simple micromethod for collagen and total protein determination in formalin-fixed paraffin-embedded sections, *J. Histochem. Cytochem.* 33:737-743 (1985).

24. Schuppan, D., Krebs, A, Bauer, M., and Hahn, E.G. Hepatitis C and liver fibrosis, *Cell Death Differ.* 10 Suppl. 1:S59-67 (2003).

25. Gressner, A.M., Weiskirchen, R., Breitkopf, K., and Dooley, S. Roles of TGF-β in hepatic fibrosis, *Front. Biosci.* 7:d793-807 (2002).

26. Bauer, M., and Schuppan, D. TGFβ1 in liver fibrosis: time to change paradigms?, *FEBS Lett.* 502:1-3 (2001).

27. Dooley, S., Delvoux, B., Streckert, M., Bonzel, L., Stopa, M., ten Dijke, P., and Gressner, A.M. Transforming growth factor beta signal transduction in hepatic stellate cells via Smad2/3 phosphorylation, a pathway that is abrogated during in vitro progression to myofibroblasts. TGFβ signal transduction during transdifferentiation of hepatic stellate cells, *FEBS Lett.* 502:4-10 (2001).

28. Marvanova, M., Menager, J., Bezard, E., Bontrop, R.E., Pradier, L., and Wong, G. Microarray analysis of nonhuman primates: validation of experimental models in neurological disorders, *FASEB J.* 17:929-931 (2003).

29. Knodell, R.G., Ishak, K.G., Black, W.C., Chen, T.S., Craig, R., Kaplowitz, N., and Kiernan, T.W. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis, *Hepatol.* 1:431-435 (1981).

30. Ishak, K., Baptista, A., Bianchi, L., Callea, F., De Groote, J., Gudat, F., and Denk, H., et al. Histological grading and staging of chronic hepatitis, *J. Hepatol.* 22:696-699 (1985).

31. Iredale, J.P.. Tissue inhibitors of metalloproteinases in liver fibrosis, *Int. J. Biochem. Cell Biol.* 29:43-54 (1997).

32. Arthur, M.J., Mann, D.A., and Iredale, J.P. Tissue inhibitors of metalloproteinases, hepatic stellate cells and liver fibrosis, *J. Gastroenterol. Hepatol.* 13 Suppl.:S33-38 (1998).

33. Mazzocca, A., Carloni, V., Sciammetta, S., Cordella, C., Pantaleo, P., Caldini, A., and Gentilini, P. Expression of transmembrane 4 superfamily (TM4SF) proteins and their role in hepatic stellate cell motility and wound healing migration, *J. Hepatol.* 37:322-330 (2002).

34. Seth, D., Leo, M.A., McGuinness, P.H., Lieber, C.S., Brennan, Y., Williams, R., and Wang, X.M., Gene expression profiling of alcoholic liver disease in the baboon (*Papio hamadryas*) and human liver, *Am. J. Pathol.* 163:2303-2317 (2003).

35. Benyon, R.C. and Arthur, M.J. Extracellular matrix degradation and the role of hepatic stellate cells, *Semin. Liver Dis.* 21(3):373-84 (2001).

36. Neubauer, K., Knittel, T., Armbrust, T., and Ramadori, G. Accumulation and cellular localization of fibrinogen/fibrin during short-term and long-term rat liver injury, *Gastroenterol.* 108(4):1124-35 (1995).

37. Hernandez-Munoz, R., Diaz-Munoz, M., Suarez-Cuenca, J.A., Trejo-Solis, C., Lopez, V., Sanchez-Sevilla, L., Yanez, L., De Sanchez, V.C. Adenosine reverses a preestablished CCl_4-induced micronodular cirrhosis through enhancing collagenolytic activity and stimulating hepatocyte cell proliferation in rats, *Hepatol.*34(4 Pt 1):677-87 (2001).

38. Morgan et al., "The Matrix Effects on Kinetic Rate Constants of Antibody-Antigen Interactions Reflect Solvent Viscosity," *J. Immunol. Meth.* 217:51-60 (1998).

39. Zhuang et al., "Measurement of Association Rate Constant of Antibody-Antigen Interactions in Solution Based on Enzyme-Linked Immunosorbent Assay," *J. Biosci. Bioeng.* 92(4):330-336 (2001).

40. Strauss, W.M. "Hybridization With Radioactive Probes," *in* Current Protocols in Molecular Biology 6.3.1-6.3.6, (John Wiley & Sons, N.Y. 2000).

41. SAMBROOK ET AL., 3 MOLECULAR CLONING: A LABORATORY MANUAL A8.40-A8.45, and A8.52-A8.55 (2001).

WHAT IS CLAIMED IS:

1. A method of detecting liver fibrosis and/or cirrhosis comprising:

assaying a biological sample from a patient for differential expression of at least one gene encoding a SEQ ID NO: 11 or SEQ ID NO: 71 protein,

wherein the differential expression of at least one gene suggests the presence of liver fibrosis and/or cirrhosis.

2. A method of detecting liver fibrosis and/or cirrhosis comprising:

determining an amount of at least one protein chosen from SEQ ID NO: 11 and SEQ ID NO: 71 in a first biological sample from a first subject without liver fibrosis and/or cirrhosis;

determining an amount of at least one protein chosen from SEQ ID NO: 11 and SEQ ID NO: 71 in a second biological sample from a second subject known or suspected of having liver fibrosis and/or cirrhosis; and

comparing the amount of said protein present in the first biological sample from the first subject with the amount of the protein present in the second biological sample from the second subject.

3. The method as claimed in claim 2, wherein the amount of at least one protein is determined by Western blot or ELISA.

4. The method as claimed in claim 3, wherein the at least one protein is the SEQ ID NO: 11 protein.

5. The method as claimed in claim 3, wherein the at least one protein is the SEQ ID NO: 71 protein.

Table 1. Clinical, chemical and fibrosis parameters in treated and untreated groups of rats

	control			DMN-treatment		
Numeric variable	**1-2 wk (n)[a]**	**3-4 wk (n)[a]**	**5-6 wk (n)[a]**	**1-2 wk (n)[a]**	**3-4 wk (n)[a]**	**5-6 wk (n)[a]**
Albumin (g/dl)	4.4±0.4 (7)	4.6±0.2 (8)	4.7±0.2 (8)	3.9±0.7 (7)	3.5±0.6 (11)	3.2±0.1 (7)
GPT (U/l)	61.1±26.7 (8)	65.9±19.7 (7)	50.3±4.9 (8)	459.5±78.5 (8)	566.6±313.5 (11)	763.6±405.2 (7)
GOT (U/l)	110.3±37.6 (8)	84.0±23.5 (7)	109.1±23.5 (8)	661.5±134.4 (8)	1006.1±749.6 (11)	1572.9±965.3 (7)
Bilirubin (mg/dl)	0.13±0.05 (8)	0.10±0.01 (8)	0.13±0.05 (8)	0.72±0.53 (8)	1.01±0.74 (11)	1.13±1.00 (7)
AKP (KA)	46.0±3.7 (4)	44.8±2.2 (4)	47.0±13.6 (4)	600.8±93.0 (4)	668.3±222.0 (3)	468±12.7 (2)
LDH (IU/l)	262.3±75.1 (4)	289.3±31.7 (3)	292.3±31.3 (4)	414.8±102.7 (4)	562.0±120.8 (3)	853.5±91.2 (2)
Globulin (g/dl)	6.9±0.3 (3)	6.9±0.5 (4)	7.3±0.2 (4)	6.7±0.1 (2)	5.0±0.8 (4)	3.6±0.3 (2)
AFP (ng/dl)	0.32±0.04 (4)	0.2±0.01 (2)	0.24±0.03 (4)	0.40±0.19 (4)	0.38±0.05 (4)	0.35±0.07 (2)
CHOL (mg/dl)	88±5 (4)	71±20 (4)	91±5 (4)	77±8 (4)	70±13 (6)	67±18 (5)
BUN (mg/dl)	31±2 (4)	25±6 (4)	26±9 (4)	33±4 (4)	36±2 (4)	31±5 (2)
ACP (mg/dl)	2.3±0.8 (4)	2.6±0.5 (4)	2.3±0.8 (4)	1.9±0.6 (4)	6.2±1.1 (4)	8.2±0.6 (2)
PT (sec)	14±1 (7)	13±1 (8)	13±1 (7)	18±4 (8)	20±4 (9)	22±5 (6)
PLT (10^3/-l)	741±245 (8)	981±124 (8)	893±109 (8)	407±72 (7)	300±165 (11)	229±302 (7)

GPT, glutamic pyruvic transaminase; GOT, glutamic oxaloacetic transaminase; Bilirubin, total bilirubin; AKP, alkaline phosphatase; LDH, lactate dehydrogenase; AFP, a-fetoprotein; CHOL, cholesterol; BUN, blood urea nitrogen; ACP, acid phosphatase; PT, prothrombin time; PLT, blood platelets.

a: Mean±SD of value from 1-2, 3-4 or 5-6 week in treated and untreated groups.

Table 2. The 63 selected genes and their human orthologs

ID	Method	Expression	Protein name-Rat	Official Symbol	Gene ID	ID	Human orthologs-name	Official Symbol	Human Homolog GeneID
SEQ ID NO: 1	Proteomics	down	cytochrome P450, family 2, subfamily c, polypeptide 23	Cyp2c23	83790				-
SEQ ID NO: 2	Proteomics	down	cytochrome P450, family 2, subfamily d, polypeptide 13	Cyp2d13	24303				-
SEQ ID NO: 3	Proteomics	down	fatty acid synthase	Fasn	50671	SEQ ID NO: 64	fatty acid synthase	FASN	2194
SEQ ID NO: 4	Proteomics	down	hydroxysteroid (17-beta) dehydrogenase 2	Hsd17b2	79243	SEQ ID NO: 65	hydroxysteroid (17-beta) dehydrogenase 2	HSD17B2	3294
SEQ ID NO: 5	Proteomics	down	Estrogen sulfotransferase, isoform 2	Ste2	-	SEQ ID NO: 66	sulfotransferase, estrogen-preferring	SULT1E1	6783
SEQ ID NO: 6	Proteomics	down	similar to cDNA sequence BC022133		362399	SEQ ID NO: 67	ZXD family zinc finger C	ZXDC	79364
SEQ ID NO: 7	Proteomics	up	S100 calcium binding protein A9 (calgranulin B)	S100a9	94195	SEQ ID NO: 68	S100 calcium-binding protein A9	S100A9	6280
SEQ ID NO: 8	Proteomics	up	actin, gamma, cytoplasmic	Actg(Actg1)	287876	-	-		-
SEQ ID NO: 9	Proteomics	up	aldehyde dehydrogenase family 1, member A1	Aldh1a1	24188	SEQ ID NO: 69	aldehyde dehydrogenase 1A1	ALDH1A1	216

ID	Method	Expression	Protein name-Rat	Official Symbol	Gene ID	ID	Human orthologs-name	Official Symbol	Human Homolog GeneID
SEQ ID NO: 10	Proteomics	up	Rho GDP dissociation inhibitor (GDI) alpha	Arhgdia	360678	SEQ ID NO: 70	Rho GDP dissociation inhibitor (GDI) alpha	ARHGDIA	396
SEQ ID NO: 11	Proteomics	up	carbonic anhydrase I (predicted)	Ca1_predicted	310218	SEQ ID NO: 71	carbonic anhydrase I	CA1	759
SEQ ID NO: 12	Proteomics	up	carbonic anhydrase 2	Ca2	54231	SEQ ID NO: 72	carbonic anhydrase II	CA2	760
SEQ ID NO: 13	Proteomics	up	cytochrome P450 2c13	Cyp2c13	171521	-	-		-
SEQ ID NO: 14	Proteomics	up	thyroid autoantigen	G22p1	25019	SEQ ID NO: 73	ATP-dependent DNA helicase II, 70 kDa subunit	XRCC6	2547
SEQ ID NO: 15	Proteomics	up	myosin, heavy polypeptide 9	Myh9	25745	SEQ ID NO: 74	myosin, heavy polypeptide 9, non-muscle	MYH9	4627
SEQ ID NO: 16	Proteomics	up	Myosin light polypeptide 6	Myl6	-	SEQ ID NO: 75	similar to Myosin light polypeptide 6 (Myosin light chain alkali 3) (Myosin light chain 3) (MLC-3) (LC17)		124685

ID	Method	Expression	Protein name-Rat	Official Symbol	Gene ID	ID	Human orthologs-name	Official Symbol	Human Homolog GeneID
SEQ ID NO: 17	Proteomics	up	prosaposin	Psap	25524	SEQ ID NO: 76	prosaposin	PSAP	5660
SEQ ID NO: 18	Proteomics	up	similar to Myosin regulatory light chain 2-A, smooth muscle isoform (Myosin RLC-A) (predicted)	RGD1309537_predicted	501203	SEQ ID NO: 77	myosin regulatory light chain MRCL3		10627
SEQ ID NO: 19	Proteomics	up	solute carrier family 4, member 1	Slc4a1	24779	SEQ ID NO: 78	solute carrier family 4, anion exchanger, member 1 (erythrocyte membrane protein band 3, Diego blood group)	SLC4A1	6521
SEQ ID NO: 20	Proteomics	up	transgelin 2 (predicted)	Tagln2	304983	SEQ ID NO: 79	transgelin 2	TAGLN2	8407
SEQ ID NO: 21	Proteomics	up	tropomyosin 3, gamma	Tpm3	117557	SEQ ID NO: 80	tropomyosin 3	TPM3	*7170*
SEQ ID NO: 22	Proteomics	up	talin	Tln	313494	SEQ ID NO: 81	talin 1	TLN1	7094

ID	Method	Expression	Protein name-Rat	Official Symbol	Gene ID	ID	Human orthologs-name	Official Symbol	Human Homolog GeneID
SEQ ID NO: 23	Proteomics	up	similar to hypothetical protein D4Ertd765e (predicted)	RGD1308525_ predicted	298381	SEQ ID NO: 82	enoyl Coenzyme A hydratase domain containing 2	ECHDC2	55268
SEQ ID NO: 24	Microarray	down	aminoadipate aminotransferase	Aadat	29416	SEQ ID NO: 83	alpha-aminoadipate aminotransferase	AADAT	51166
SEQ ID NO: 25	Microarray	down	acyl-CoA synthetase long-chain family member 1	Acsl1	25288	SEQ ID NO: 84	acyl-CoA synthetase long-chain family member 1	ACSL1	2180
SEQ ID NO: 26	Microarray	down	aldehyde dehydrogenase 1 family, member L1	Fthfd	64392	SEQ ID NO: 85	aldehyde dehydrogenase 1 family, member L1	ALDH1L1	10840
SEQ ID NO: 27	Microarray	down	arginosuccinate synthetase	Ass	25698	SEQ ID NO: 86	argininosuccinate synthetase	ASS	445
SEQ ID NO: 28	Microarray	down	dimethylglycine dehydrogenase	Dmgdh	245961	SEQ ID NO: 87	dimethylglycine dehydrogenase	DMGDH	29958
SEQ ID NO: 29	Microarray	down	dihydropyrimidinase	Dpys	65135	SEQ ID NO: 88	dihydropyrimidinase	DPYS	1807
SEQ ID NO: 30	Microarray	down	phosphodiesterase I/nucleotide pyrophosphatase 2	Enpp2	84050	SEQ ID NO: 89	ectonucleotide pyrophosphatase/ phosphodiesterase 2 (autotaxin)	ENPP2	5168

ID	Method	Expression	Protein name-Rat	Official Symbol	Gene ID	ID	Human orthologs-name	Official Symbol	Human Homolog GeneID
SEQ ID NO: 31	Microarray	down	fatty acid amide hydrolase	Faah	29347	SEQ ID NO: 90	fatty acid amide hydrolase	FAAH	2166
SEQ ID NO: 32	Microarray	down	guanidinoacetate methyltransferase	Gamt	25257	SEQ ID NO: 91	guanidinoacetate N-methyltransferase	GAMT	2593
SEQ ID NO: 33	Microarray	down	L-gulono-gamma-lactone oxidase	Gulo	60671	-	-		
SEQ ID NO: 34	Microarray	down	3-hydroxyisobutyrate dehydrogenase	Hibadh	63938	SEQ ID NO: 92	3-hydroxyisobutyrate dehydrogenase	HIBADH	11112
SEQ ID NO: 35	Microarray	down	ketohexokinase	Khk	25659	SEQ ID NO: 93	ketohexokinase	KHK	3795
SEQ ID NO: 36	Microarray	down	kynureninase (L-kynurenine hydrolase)	Kynu	116682	SEQ ID NO: 94	kynureninase (L-kynurenine hydrolase)	KYNU	8942
SEQ ID NO: 37	Microarray	down	Pyruvate carboxylase	Pc	25104	SEQ ID NO: 95	pyruvate carboxylase	PC	5091
SEQ ID NO: 38	Microarray	down	pyruvate dehydrogenase kinase 2 subunit p45 (PDK2)	Pdk2	81530	SEQ ID NO: 96	pyruvate dehydrogenase kinase, isoenzyme 2	PDK2	5164

ID	Method	Expression	Protein name-Rat	Official Symbol	Gene ID	ID	Human orthologs-name	Official Symbol	Human Homolog GeneID
SEQ ID NO: 39	Microarray	down	solute carrier family 10, member 1	Slc10a1	24777	SEQ ID NO: 97	solute carrier family 10 (sodium/bile acid cotransporter family), member 1	SLC10A1	6554
SEQ ID NO: 40	Microarray	down	solute carrier family 37 (glycerol-6-phosphate transporter), member 4	Slc37a4	29573	SEQ ID NO: 98	solute carrier family 37 (glycerol-6-phosphate transporter), member 4	SLC37A4	2542
SEQ ID NO: 41	Microarray	down	similar to Sulfotransferase K2 (rSULT1C2A)	-	501086	SEQ ID NO: 99	sulfotransferase family, cytosolic, 1C, member 1	SULT1C1	6819
SEQ ID NO: 42	Microarray	up	arachidonate 5-lipoxygenase activating protein	Alox5ap	29624	SEQ ID NO: 100	arachidonate 5-lipoxygenase-activating protein	ALOX5AP	241
SEQ ID NO: 43	Microarray	up	adenine phosphoribosyl transferase (predicted)	Aprt_predicted	292072	SEQ ID NO: 101	adenine phosphoribosyltransferase	APRT	353
SEQ ID NO: 44	Microarray	up	brain abundant, membrane attached signal protein 1	Basp1	64160	SEQ ID NO: 102	brain abundant, membrane attached signal protein 1	BASP1	10409
SEQ ID NO: 45	Microarray	up	benzodiazepin receptor	Bzrp	24230	SEQ ID NO: 103	peripheral benzodiazapine receptor	BZRP	706

ID	Method	Expression	Protein name-Rat	Official Symbol	Gene ID	ID	Human orthologs-name	Official Symbol	Human Homolog GeneID
SEQ ID NO: 46	Microarray	up	carbonyl reductase	Cbr1	29224	SEQ ID NO: 104	carbonyl reductase 1	CBR1	873
SEQ ID NO: 47	Microarray	up	CD47 antigen (Rh-related antigen, integrin-associated signal transducer)	Cd47	29364	SEQ ID NO: 105	CD47 antigen	CD47	961
SEQ ID NO: 48	Microarray	up	embigin	Emb	114511	SEQ ID NO: 106	embigin homolog	EMB	133418
SEQ ID NO: 49	Microarray	up	myeloid differentiation primary response gene 116	Myd116	171071	SEQ ID NO: 107	protein phosphatase 1, regulatory subunit 15A	PPP1R15A	23645
SEQ ID NO:50	Microarray	up	N-myc downstream regulated gene 2	Ndrg2	171114	SEQ ID NO: 108	N-myc downstream-regulated gene 2	NDRG2	57447
SEQ ID NO:51	Microarray	up	expressed in non-metastatic cells 1, protein (NM23A) (nucleoside diphosphate kinase)	Nme1	191575	SEQ ID NO: 109	non-metastatic cells 1, protein (NM23A) expressed in	NME1	4830
SEQ ID NO: 52	Microarray	up	neuregulin 1	Nrg1	112400	SEQ ID NO: 110	neuregulin 1	NRG1	3084
SEQ ID NO: 53	Microarray	up	oxidized low density lipoprotein (lectin-like) receptor 1	Oldlr1	140914	SEQ ID NO: 111	oxidised low density lipoprotein (lectin-like) receptor 1	OLR1	4973

ID	Method	Expression	Protein name-Rat	Official Symbol	Gene ID	ID	Human orthologs-name	Official Symbol	Human Homolog GeneID
SEQ ID NO: 54	Microarray	up	protein tyrosine kinase 9	Ptk9	315265	SEQ ID NO: 112	twinfilin-like protein	PTK9L	11344
SEQ ID NO: 55	Microarray	up	syndecan binding protein	Sdcbp	83841	SEQ ID NO: 113	syntenin	SDCBP	6386
SEQ ID NO: 56	Microarray	up	serine protease inhibitor, Kazal type 1	Spink1	24833	SEQ ID NO: 114	serine protease inhibitor, Kazal type 1	SPINK1	6690
SEQ ID NO: 57	Microarray	up	tyrosine 3/tryptophan 5 - monooxygenase activation protein, eta polypeptide	Ywhah	25576	SEQ ID NO: 115	tyrosine 3/tryptophan 5 - monooxygenase activation protein, eta polypeptide	YWHAH	7533
SEQ ID NO: 58	Microarray	up	S100 calcium binding protein A10	S100a10	81778	SEQ ID NO: 116	S100 calcium-binding protein A10	S100A10	6281
SEQ ID NO: 59	Microarray	up	B-cell translocation gene 2	Btg2	29619	SEQ ID NO: 117	B-cell translocation gene 2	BTG2	7832
SEQ ID NO: 60	Microarray	up	3'-phosphoadenosine 5'-phosphosulfate synthase 2 (predicted)	Papss2_ predicted	294103	SEQ ID NO: 118	3'-phosphoadenosine 5'-phosphosulfate synthase 1	PAPSS1	9061

ID	Method	Expression	Protein name-Rat	Official Symbol	Gene ID	ID	Human orthologs-name	Official Symbol	Human Homolog GeneID
SEQ ID NO: 61	Microarray	up	biogenesis of lysosome-related organelles complex-1, subunit 2	Bloc1s2	293938	SEQ ID NO: 119	biogenesis of lysosome-related organelles complex-1, subunit 2	BLOC1S2	282991
SEQ ID NO: 62	Microarray	up	Similar to RIKEN cDNA 6720467C03 (predicted)	RGD1310681_ predicted	297903		-		-
SEQ ID NO: 63	Microarray	up	transformed mouse 3T3 cell double minute 2 (predicted)	Mdm2_ predicted	314856	SEQ ID NO: 120	mouse double minute 2 homolog	MDM2	4193

Table 3A. Interacting network analysis of proteins selected from microarray study

	Genes*	Top functions
1	**ACTG1**, ADARB1, AHCY, AKT1, ALDOA, **BASP1, BZRP**, CD63, Crisp1 (MGI:102553), **EMB, FAAH**, GLI1, GNMT, HRAS, IFNG, **KHK**, KRT18, MYC, MYCN, **NDRG2, NME1**, OTC, PPARA, RHOB, RPL23, RPL27, RPL35, RPL38, RPL41, Rpm12, RPS16, RPS27, RT1-A, SARDH, **SLC37A4**	**Cancer, Cell Morphology, Dermatological Diseases and Conditions**
2	**ACSL1, ALOX5AP**, ANXA1, **ASS**, ATF3, **CD47**, CIDEC, CP, DF, FABP1, GPD1, HOXA11, **KYNU**, MMP16, MTP, **MYL6, OLR1, PC, PDK2**, PF4, PLIN, PLP1, PPARG, **PPP1R15A**, PROC, PTGES, RGS7, SLC27A1, SPHK1, TIMP1, TNF, UBQLN2, UCP2, UCP3, Wap	**Lipid Metabolism, Small Molecule Biochemistry, Molecular Transport**
3	ACO2, AKR1B10, **BTG2, CBR1**, DAD1, DUSP5, ELF3, **ENPP2**, FGF2, GBP1, GLG1, GPI, GRIN2C, HSPA1A, IGFALS, IL1B, JUN, KCNH2, MAOB, MT1L, MT2A, NR3C1, **NRG1**, P53AIP1, PHKG1, **S100A10, SDCBP**, SHBG, **SLC10A1**, SLC10A2, **SPINK1**, SPRY1, TP53, **YWHAH**, ZFP36L2	**Cancer, Cellular Movement, Cellular Growth and Proliferation**

* Genes identified in the present study are underlined. The underlined genes that have not been reported to associate with liver injury are in bold.

Table 3B. Interacting network analysis of proteins selected from iTRAQ proteomic study

	Genes*	Top functions
1	ADD1, ALB, **ALDH1A1**, **CA2**, CASP3, CNN2, CTNNB1, DIA1, ENPP1, GBA, GM2A, HGF, **HSD17B2**, IL1B, ITPA, JTV1, KRT17, LMNA, LMNB2, MAP4, MYC, OAT, PLS3, **PSAP**, **S100A9**, SERPINH1, **SLC4A1**, SRM, **SULT1E1**, **TAGLN2**, TGFB1, **TLN1**, **TPM3**, UGDH, VIM	**Cancer, Cell Cycle, Cell Morphology**
2	ACADM, ACTC, ADCY7, ALDH2, ALDOA, ALDOB, ARG2, **ARHGDIA**, CD276, CNR2, CPS1, CROT, Cyp4a10, Cyp4a14, EHHADH, EPHX1, FAAH, **FASN**, FN1, **G22P1**, GNMT, GOT2, HOXD4, HPX, HRSP12, Ifi1, IFNG, IL4, LEP, Mug1, **MYH9**, MYOD1, PPARA, SORD, TGM2	**Lipid Metabolism, Small Molecule Biochemistry, Organismal Injury and Abnormalities**
3	ALB, ANXA1, CLU, **CYP2C13**, CYP3A2, FGA, FGB, FGG, FLI1, GAL, GFAP, GH1, GHRH, GHRL, HBA1, HBA2, HBB, Hbb-ar, Hbb-b1, Hbb-b2, Hbb-bh1, Hbb-y, HBD, HBE1, HBG1, HBG2, HBQ1, HBZ, IL6, NTS, SOCS2, Stat5, TF, TG, TTR	**Hematological Disease, Endocrine System Development and Function, Nervous System Development and Function**

* Genes identified in the present study are underlined. The underlined genes that have not been reported to associate with liver injury are in bold.

SEQ ID 1 cytochrome P450, family 2, subfamily c, polypeptide 23

MELLGRTTLALVVSVTCLSLLSVWTKLRTRGRLPPGPHPPSHYWESTATEPQGHPASLSKLAKEYGPVYT
LYFGTSPTVVLHGYDVVKEALLQQGDEFLGRGPLPIIEDTHKGYGLIFSNGERWKVMRRFSLMTLRNFGM
GKRSLEERVQEEAWCLVEELQKTKAQPFDPTFILACAPCNVICSILFNDRFQYNDKTFLNLMDLLNKNFQ
QVNSVWCQMYNLWPTIIKYLPGKHIEFAKRIDDVKNFILEKVKEHQKSLDPANPRDYIDCFLSKIEEEKD
NLKSEPHLENLAVCGSNLFTAGTETTSTTLRFGLLLLMKYPEVQAKVHEELDRVIGRHQPPSMKDKMKLP
YTDAVLHEIQRYITLVGSSLPHAVVQDTKFRDYVIPKGTTVLPMLSSVMLDQKEFANPEKFDPGHFLDKN
GCFKKTDYFVPFSLGKRACVGESLARMELFLFFTTLLQKFSLKTLVEPKDLDIKPITTGIINLPPPYKLC
LVPR

SEQ ID 2 cytochrome P450, family 2, subfamily d, polypeptide 13

MELLAGTGLWPMAIFTVIFILLVDLMHRRQRWTSRYPPGPVPWPVLGNLLQVDLCNMPYSMYKLQNRYGD
VFSLQMGWKPVVVINGLKAVQELLVTCGEDTADRPEMPIFQHIGYGHKAKGVVLAPYGPEWREQRRFSVS
TLRNFGVGKKSLEQWVTDEASHLCDALTAEAGRPLDPYTLLNKAVCNVIASLIYARRFDYGDPDFIKVLK
ILKESMGEQTGLFPEVLNMFPVLLRIPGLADKVFPGQKTFLTMVDNLVTEHKKTWDPDQPPRDLTDAFLA
EIEKAKGNPESSFNDANLRLVVNDLFGAGMVTTSITLTWALLLMILHPDVQCRVQQEIDEVIGQVRHPEM
ADQAHMPFTNAVIHEVQRFADIVPMNLPHKTSRDIEVQGFLIPKGTTLIPNLSSVLKDETVWEKPLRFHP
EHFLDAQGNFVKHEAFMPFSAGRRACLGEPLARMELFLFFTCLLQRFSFSVPTGQPRPSDYGVFAFLLSP
SPYQLCAFKR

SEQ ID 3 fatty acid synthase

MEEVVIAGMSGKLPESENLQEFWANLIGGVDMVTDDDRRWKAGLYGLPKRSGKLKDLSKFDASFFGVHPK
QAHTMDPQLRLLLEVSYEAIVDGGINPASLRGTNTGVWVGVSGSEASEALSRDPETLLGYSMVGCQRAMM
ANRLSFFFDFKGPSIALDTACSSSLLALQNAYQAIRSGECPAAIVGGINLLLKPNTSVQFMKLGMLSPDG
TCRSFDDSGNGYCRAEAVVAVLLTKKSLARRVYATILNAGTNTDGCKEQGVTFPSGEAQEQLIRSLYQPG
GVAPESLEYIEAHGTGTKVGDPQELNGITRSLCAFRQSPLLIGSTKSNMGHPEPASGLAALTKVLLSLEN
GVWAPNLHFHNPNPEIPALLDGRLQVVDRPLPVRGGIVGINSFGFGGANVHVILQPNTQQAPAPAPHAAL
PHLLHASGRTMEAVQGLLEQGRQHSQDLAFVSMLNDIAATPTAAMPFRGYTVLGVEGHVQEVQQVPASQR
PLWFICSGMGTQWRGMGLSLMRLDSFRESILRSDEALKPLGVKVSDLLLSTDEHTFDDIVHSFVSLTAIQ
IALIDLLTSMGLKPDGIIGHSLGEVACGYADGCLSQREAVLAAYWRGQCIKDANLPAGSMAAVGLSWEEC
KQRCPPGVVPACHNSEDTVTISGPQAAVNEFVEQLKQEGVFAKEVRTGGLAFHSYFMEGIAPTLLQALKK
VIREPRPRSARWLSTSIPEAQWQSSLARTSSAEYNVNNLVSPVLFQEALWHVPEHAVVLEIAPHALLQAV
LKRGVKPSCTIIPLMKRDHKDNLEFFLTNLGKVHLTGIDINPNALFPPVEFPVPRGTPLISPHIKWDHSQ
TWDIPVAEDFPNGSSSSSATVYNIDASSESSDHYLVDHCIDGRVLFPGTGYLYLVWKTLARSLSLSLEET
PVVFENVTFHQATILPRTGTVPLEVRLLEASHAFEVSDSGNLIVSGKVYQWEDPDSKLFDHPEVPIPAES
ESVSRLTQGEVYKELRLRGYDYGPHFQGVYEATLEGEQGKLLWKDNWVTFMDTMLQISLGFSKQSLQLP
TRVTAIYIDPATHLQKVYMLEGDTQVADVTTSRCLGVTVSGGVYISRLQTTATSRRQQEQLVPTLEKFVF
TPHVEPECLSESAILQKELQLCKGLAKALQTKATQQGLKMTVPGLEDLPQHGLPRLLAAACQLQLNGNLQ
LELGEVLARERLLLPEDPLISGLLNSQALKACIDTALENLSTLKMKVVEVLAGEGHLYSHISALLNTQPM
LQLEYTATDRHPQALKDVQTKLQQHDVAQGQWDPSGPAPTNLGALDLVVCNCALATLGDPALALDNMVAA
LKDGGFLLMHTVLKGHALGETLACLPSEVQPGPSPLSQEEWESLFSRKALHLVGLKKSFYGTALPLCRRL
SPQDKPIPLPVEDTSFQWVDSLKSILATSSSQPVWLTAMNCPTSGVVGLVNCLRKEPGGHRIRCILLSNL
SSTSHVPKLDPGSSELQKVLESDLVMNVYRDGAWGAERHFQLEQDKPEEQTAHAFVNVLTRGDLASIRWV
SSPLKHMQPPSSSGAQLCTVYYASLNFRDIMLATGKLSPDAIPGKWASRDCMLGMEFSGRDKCGRRVMGL
VPAEGLATSVLLSPDFLWDVPSSWTLEEAASVPVVYTTAYYSLVVRGRIQHGETVLIHSGSGGVGQAAIS
IALSLGCRVFTTVGSAEKRAYLQARFPQLDDTSFANSRDTSFEQHVLLHTGGKGVDLVLNSLAEEKLQAS
VRCLAQHGRFLEIGKFDLSNNHPLGMAIFLKNVTFHGILLDALFEGANDSWREVAELLKAGIRDGVVKPL
KCTVFPKAQVEDAFRYMAQGKHIGKVLVQVREEDPEAMLPGAQPTLISAISKTFCPEHKSYIITGGLGGF
GLELARWLVLRGAQRLVLTSRSGIRTGYQAKHVREWRRQGIHVLVSTSNVSSLEGARALIAEATKLGPVG
GVFNLAMVLRDAMLENQTPELFQDVNKPKYNGTLNLDRATREACPELDYFVAFSSVSCGRGNAGQSNYGF
ANSTMERICEQRRHDGLPGLAVQWGAIGDVGIILEAMGTNDTVVGGTLPQRISSCMEVLDLFLNQPHAVL
SSFVLVEKKAVAHGDGEAQRDLVKAVAHILGIRDLAGINLDSSLADLGLDSLMGVEVRQILEREHDLVLP
IREVRQLTLRKLQEMSSKAGSDTELAAPKSKNDTSLKQAQLNLSILLVNPEGPTLTRLNSVQSSERPLFL

VHPIEGSITVFHSLAAKLSVPTYGLQCTQAAPLDSIPNLAAYYIDCIKQVQPEQPHRVAGYSFGACVAFE
MCSQLQAQQGPAPAHNNLFLFDGSHTYVLAYTQSYRAKLTPGCEAEAEAEAICFFIKQFVDAEHSKVLEA
LLPLKSLEDRVAAAVDLITRSHQSLDRRDLSFAAVSFYYKLRAADQYKPKAKYHGNVILLRAKTGGTYGE
DLGADYNLSQVCDGKVSVHIEGDHRTLLEGRGLESIINIIHSSLAEPRVSVREG

SEQ ID 4 hydroxysteroid (17-beta) dehydrogenase 2

MNPFSSESAWLCLTATAVLGGMLLCKAWSSGQLRSQVVCLAGLWGGACLLSLSLLCSLFLLSVSCFFLLY
VSSSDQDLLPVDQKAVLVTGADSGFGHALAKHLDKLGFTVFAGVLDKEGPGAEELRKNCSERLSVLQMDV
TKPEQIKDVHSEVAEKIQDKGLWAVVNNAGVLHFPIDGELIPMTVYRKCMAVNFFGAVEVTKVFLPLLRK
SKGRLVNVSSMGAMIPFQMVAAYASTKAAISMFSAVIRQELAKWGVKVVTIHPGGFQTNIVGSQDSWDKM
EKELDHFSKEIQENYGQEYVHTQKLALPVMREMSNPDITPVLRDIQHAICAKNPSSFYCSGRMTYLWIC
FAAYSPISLLDYILKNYFTPKLMPRALRTAS

SEQ ID 5 Estrogen sulfotransferase, isoform 2

METSMPEYYEVFGDFHGFLMDKLFTKYWEDVETFSARPDDLLVVTYPKSGSTWIGEIVDMIYKEGDVEKC
KEDAIFNRIPYLECRNEDLINGIKQLKEKESPRIVKTHLPAKLLPASFWEKNCKIIYLCRNAKDVVVSYY
YFFLIMKSYPNPKSFSEFVEKFMEGQVPYGSWYDHVKSWWEKSKNSRVLFMFYEDMKEDIRREVVKLIEF
LERDPSAELVDRIIQHTSFQEMKNNPCTNYSMLPETMIDLKVSPFMRKGIVGDWKNHFPEALRERFEEHY
QQQMKDCPVKFRAEL

SEQ ID 6 similar to cDNA sequence BC022133

MSSLQELCSGLPLRPLPENRGRWAGVPHAPVRTPDLSPEEEQLALRNALRYFPPDVQKLLALEFAQELRQ
FGHIYMYRFCPSIEMRAYPIEQYPCRTRAAAAIMHMIMNNLDPAVAQFPQELVTYGGNGQVFSNWAQVWV
QGASSHLQSPPCALSIQFWLTMSYLSKMTEEQTLVMYSGHPLGLPPSSSHAPRLVITNGMVIPNYSSRTE
YEKLFALGVTMYGQMTAGSYCYIGPQGIVHGTVLTVLNAGRRYLGIEDLAGKVFVTSGLGGMSGAQAKAA
VIVGCIGVIAEVDKAALVKRHQQGWLMEVTDSLDHCIARLREARKKKEVLSLGYHGNVVDLWERLVHELD
TTGELLVDLGSDQTSCHNPFNGGYYPVQLSFSEAQSLMSSNPAAFKNLVQESLRRHISAINRLAQEKFFF
WDYGNAFLLEAQRAGADVEKKGANKTEFRYPSYVQHIMGDIFSQGFGPFRWVCTSGNPQDLTVTDHLATS
VLEKAIADGGDTSGWVVLRSDSCFCGALGGLCQILQESKDHRLLHKGTSSTFRSHSLKAVLMKASVKLQY
MDNIRWIREAAKHQLVVGSQARILYSDQKGRVAIAVAINQAIASGKIKAPVVLSRDHHDVSGTDSPFRET
SNIYDGSAFCADMAVQNFVGDACRGATWVALHNGGGVGWGEVINGGFGLVLDGTPEAEQKARMMLSWDVS
NGVARRCWSGNPKAYEIICQTMHDNSGLVVTLPHEVADEQVLRQDLLQGKPGLRALQRGHREDEKQSREV
QSLSASSGPGDRRLGLKAKESQLGGPAPAPPPGSLVTTVPTALARHAHAHGPRARAARDATLGCTLRLAL
TETPGGRTLGVARMPGRAGPLFSGGCGSDASAGPKMDLPAVLAAPATRGDQHGGGPSRLRRGTGPSLGAG
PGRRRLLLLRSPEDGGPGPRQEEAPGPSPPPPEDGGDSFVVLLEVPRAADTHEQEETEPDSGASPTEQVP
AAAPGAALAGTVTIHNQDLLVRFDRGVFTLAAAPAPATQGLHPATTPGLEPSSAAASRRGPVAASAGSPA
YRCPEPQCALSFAKKHQLKVHLLTHGSLQGRRPFKCPLDGCGWAFTTSYKLKRHLQSHDKLRPFSCPVGG
CGKKFTTVYNLKAHMKGHEQESLFKCEVCAERFPTHAKLNSHQRSHFEPERPYKCDFPGCEKTFITVSAL
FSHNRAHFREQELFSCSFPGCNKQYDKACRLKIHLRSHTGERPFICDSDSCGWTFTSMSKLLRHKRKHDD
DRRFTCPVEGCGKSFTRAEHLKGHSITHLGTKPFECPVEGCCARFSARSSLYIHSKKHLQDVGTPKSRCP
VSSCNRLFTSKHSMKAHMVRQHSRRQDLVPQLEAPSSLTPSSELSSPGQSELTNIDLAALFSDTPANSSS
STAGSDEALNSGILTIDVTSVSSSLGGNLPTNNNSLGPMDPLVLVAHSDMPPSLDSPLVLGTPATVLQPG
SFSADDSQAMSTGAVGCLVALPMRNLSQDSPALTPSNNLTALGTTPTSSDTTQETASVPDLLVPIKVEQD
LSPVPDVVQQQKESHGPSVLSSSAERLGAQKDSELSAGTGSLYLV

SEQ ID 7 S100 calcium binding protein A9 (calgranulin B)

MAAKTGSQLERSISTIINVFHQYSRKYGHPDTLNKAEFKEMVNKDLPNFLKREKRNENLLRDIMEDLDTN
QDNQLSFEECMMLMGKLIFACHEKLHENNPRGHDHRHGKGCGK

SEQ ID 8 actin, gamma, cytoplasmic

MEEEIAALVIDNGSGMCKAGFAGDDAPRAVFPSIVGRPRHQGVMVGMGQKDSYVGDEAQSKRGILTLKYP
IEHGIVTNWDDMEKIWHHTFYNELRVAPEEHPVLLTEAPLNPKANREKMTQIMFETFNTPAMYVAIQAVL
SLYASGRTTGIVMDSGDGVTHTVPIYEGYALPHAILRLDLAGRDLTDYLMKILTERGYSFTTTAEREIVR

DIKEKLCYVALDFEQEMATAASSSSLEKSYELPDGQVITIGNERFRCPEALFQPSFLGMESCGIHETTFN
SIMKCDVDIRKDLYANTVLSGGTTMYPGIADRMQKEITALAPSTMKIKIIAPPERKYSVWIGGSILASLS
TFQQMWISKQEYDESGPSIVHRKCF

SEQ ID 9 aldehyde dehydrogenase family 1, member A1

MSSPAQPAVPAPLANLKIQHTKIFINNEWHDSVSGKKFPVLNPATEEVICHVEEGDKADVDKAVKAARQA
FQIGSPWRTMDASERGRLLNKLADLMERDRLLLATIEAINGGKVFANAYLSDLGGSIKALKYCAGWADKI
HGQTIPSDGDIFTFTRREPIGVCGQIIPWNFPLLMFIWKIGPALSCGNTVVVKPAEQTPLTALHMASLIK
EAGFPPGVVNIVPGYGPTAGAAISSHMDVDKVAFTGSTQVGKLIKEAAGKSNLKRVTLELGGKSPCIVFA
DADLDIAVEFAHHGVFYHQGQCCVAASRIFVEESVYDEFVRKSVERAKKYVLGNPLTQGINQGPQIDKEQ
HDKILDLIESGKKEGAKLECGGGRWGNKGFFVQPTVFSNVTDEMRIAKEEIFGPVQQIMKFKSIDDVIKR
ANNTTYGLAAGVFTKDLDRAITVSSALQAGVVWVNCYMILSAQCPFGGFKMSGNGRELGEHGLYEYTELK
TVAMKISQKNS

SEQ ID 10 Rho GDP dissociation inhibitor (GDI) alpha

MAEQEPTAEQLAQIAAENEEDEHSVNYKPPAQKSIQEIQELDKDDESLRKYKEALLGRVAVSADPNVPNV
IVTRLTLVCSTAPGPLELDLTGDLESFKKQSFVLKEGVEYRIKISFRVNREIVSGMKYIQHTYRKGVKID
KTDYMVGSYGPRAEEYEFLTPMEEAPKGMLARGSYNIKSRFTDDDKTDHLSWEWNLTIKKEWKD

SEQ ID 11 carbonic anhydrase I (predicted)

MASADWGYDSKNGPDQWSKLYPIANGNNQSPIDIKTSEAKHDSSLKPVSVSYNPATAKEIVNVGHSFHVV
FDDSSNQSVLKGGPLADSYRLTQFHFHWGNSNDHGSEHTVDGAKYSGELHLVHWNSAKYSSAAEAISKAD
GLAIIGVLMKVGPANPNLQKVLDALSSVKTKGKRAPFTNFDPSSLLPSSLDYWTYPGSLTHPPLHESVTW
VICKESISLSPEQLAQLRGLLSSAEGEPAVPVLSNHRPPQPLKGRTVRASF

SEQ ID 12 carbonic anhydrase 2

MSHHWGYSKSNGPENWHKEFPIANGDRQSPVDIDTGTAQHDPSLQPLLICYDKVASKSIVNNGHSFNVEF
DDSQDFAVLKEGPLSGSYRLIQFHFHWGSSDGQGSEHTVNKKKYAAELHLVHWNTKYGDFGKAVQHPDGL
AVLGIFLKIGPASQGLQKITEALHSIKTKGKRAAFANFDPCSLLPGNLDYWTYPGSLTTPPLLECVTWIV
LKEPITVSSEQMSHFRKLNFNSEGEAEELMVDNWRPAQPLKNRKIKASFK

SEQ ID 13 cytochrome P450 2c13

MDPVVVLLLSLFFLLFLSLWRLSSGRGKLPPGPTPLPIIGNFFQVDMKDIRQSLTNFSKTYGPVYTLYVG
SQPTVVLHGYEALKEALVDHGEEFSGRGRLPICEKVAKGQGIAFSHGNVWKATRHFTVKTLRNLGMGKGT
IEDKVQEEAKWLVKELKKTNGSPCDPQFIMGCAPGNVICCIILQNRFDYEDKDFLNLIEKVNEAVKIISS
PGIQVFNIFPILLDYCPGNHNIYLKNYTWVKSYLLEKIKEHEESLDVSNPRDFIDYFLIEMNQENANQWM
NYTLEHLAIMVTDLFFAGIETVSSTMRFALLLLMKYPHVTAKVQEEIDHVIGRHRSPSMQDRSHMPYTNA
MVHEVQRYIDIGPNGLLHDVTCDTKFRNYFIPKGTAVLTSLTSVLHDSKEFPNPEMFDPGHFLDENGNFK
KSDYFIPFSAGKRMCLGESLARMELFLFLTTILQNFKLKSLVDPKDINTTPICSSLSSVPPTFQMRFIPL

SEQ ID 14 thyroid autoantigen

MSEWESYYKTEGEEEEEEEQSPDTNGEYKYSGRDSLIFLVDASRAMFESQQEDELTPFDMSIQCIQSVYT
SKIISSDRDLLAVVFYGTEKDKNSVNFKSIYVLQDLDNPGAKRVLELDRFKGQQGKKHFRDTIGHGSDYS
LSEVLWVCANLFSDVQFKMSHKRIMLFTNEDDPHGNDSAKASRARTKASDLRDTGIFLDLMHLKKRGGFD
VSLFYRDIISIAEDEDLGVHFEESSKLEDLLRKVRAKETKKRVLSRLKFKLGKDVALMVGVYNLVQKANK
PFPVRLYRETNEPVKTKTRTFNVNTGSLLLPSDTKRSLTFGTRQIVLEKEETEELKRFDEPGLILMGFKP
MVMLKNHHYLRPSLFLYPEESLVNGSSTLFSALLTKCVEKEVIAVCRYTARKNVSPYPVALVPQEEELDD
QNIQVTPAGFQLVFLPYADDKRKVPFTEKVMANPEQIDKMKAHVQKLRETYRSDSFENPVLQQHFRNLEA
LALDMMESEQVVDLTLPKVEAIKKRLGSLADEFKELVYPPGYNPEGKIAKRKADNEGSASKKPKVELSEE
ELKDLFAKGTLGKLTVPALRDICKAYGLKSGPKKQELLEALSRHLEKTD

SEQ ID 15 myosin, heavy polypeptide 9

MAQQAADKYLYVDKNFINNPLAQADCGAKKLVWVPSTKNGFEPASLKEEVGEEAIVELVENGKKVKVNKD
DIQKMNPPKFSKVEDMAELTCLNEASVLHNLKERYYSGLIYTYSGLFCVVINPYKNLPIYSEEIVDMYKG
KKRHEMPPHIYAITDTAYRSMMQDREDQSILCTGESGAGKTENTKKVIQYLAHVASSHKSKKDQGELERQ
LLQANPILEAFGNAKTVKNDNSSRFGKFIRINFDVNGYIVGANIETYLLEKSRAIRQAKEERTFHIFYYL
LSGAGEHLKTDLLLEPYNKYRFLSNGHVTIPGQQDKDMFQETMEAMRIMGIPEDEQMGLLRVISGVLQLG
NIVFKKERNTDQASMPDNTAAQKVSHLLGINVTDFTRGILTPRIKVGRDYVQKAQTKEQADFAIEALAKA
TYERMFRWLVLRINKALDKTKRQGASFIGILDIAGFEIFDLNSFEQLCINYTNEKLQQLFNHTMFILEQE
EYQREGIEWNFIDFGLDLQPCIDLIEKPAGPPGILALLDEECWFPKATDKSFVEKVVQEQGTHPKFQKPK
QLKDKADFCIIHYAGKVDYKADEWLMKNMDPLNDNIATLLHQSSDKFVSELWKDVDRIIGLDQVAGMSET
ALPGAFKTRKGMFRTVGQLYKEQLAKLMATLRNTNPNFVRCIIPNHEKKAGKLDPHLVLDQLRCNGVLEG
IRICRQGFPNRVVFQEFRQRYEILTPNSIPKGFMDGKQACVLMIKALELDSNLYRIGQSKVFFRSGVLAH
LEEERDLKITDVIIGFQACCRGYLARKAFAKRQQQLTAMKVLQRNCAAYLRLRNWQWWRLFTKVKPLLNS
IRHEDELLAKEAELTKVREKHLAAENRLTEMETMQSQLMAEKLQLQEQLQAKTELCAEAEELRARLTAKK
QELEEICHDLEARVEEEEERCQYLQAEKKKMQQNIQELEEQLEEEESARQKLQLEKVTTEAKLKKLEEDQ
IIMEDQNCKLAKEKKLLEDRVAEFTTDLMEEEEKSKSLAKLKNKHEAMITDLEERLRREEKQRQELEKTR
RKLEGDSTDLSDQIAELQAQIAELKMQLAKKEEELQAALARVEEEAAQKNMALKKIRELETQISELQEDL
ESERACRNKAEKQKRDLGEELEALKTELEDTLDSTAAQQELRSKREQEVSILKKTLEDEAKTHEAQIQEM
RQKHSQAVEELAEQLEQTKRVKATLEKAKQTLENERGELANEVKALLQGKGDSEHKRKKVEAQLQELQVK
FSEGERVRTELADKVSKLQVELDSVTGLLNQSDSKSSKLTKDFSALESQLQDTQELLQEENRQKLSLSTK
LKQMEDEKNSFREQLEEEEEEAKRNLEKQIATLHAQVTDMKKKMEDGVGCLETAEEAKRRLQKDLEGLSQ
RLEEKVAAYDKLEKTKTRLQQELDDLLVDLDHQRQSVSNLEKKQKKFDQLLAEEKTISAKYAEERDRAEA
EAREKETKALSLARALEEAMEQKAELERLNKQFRTEMEDLMSSKDDVGKSVHELEKSNRALEQQVEEMKT
QLEELEDELQATEDAKLRLEVNLQAMKAQFERDLQGRDEQSEEKKKQLVRQVREMEAELEDERKQRSIAM
AARKKLEMDLKDLEAHIDTANKNREEAIKQLRKLQAQMKDCMRDVDDTRASREEILAQAKENEKKLKSME
AEMIQLQEELAAAERAKRQAQQERDELADEIANSSGKGALALEEKRRLEALIALLEEELEEEQGNTELIN
DRLKKANLQIDQINTDLNLERSHAQKNENARQQLERQNKELKAKLQEMESAVKSKYKASIAALEAKIAQL
EEQLDNETKERQAASKQVRRAEKKLKDVLLQVEDERRNAEQFKDQADKASTRLKQLKRQLEEAEEEAQRA
NASRRKLQRELEDATETADAMNREVSSLKNKLRRGDMPFVVTRRIVRKGTGDCSDEEVDGKADGADAKAT
E

SEQ ID 16 Myosin light polypeptide 6

MCDFTEDQTAEFKEAFQLFDRTGDGKILYSQCGDVMRALGQNPTNAEVLKVLGNPKSDEMNVKVLDFEHF
LPMLQTVAKNKDQGTYEDYVEGLRVFDKEGNGTVMGAEIRHVLVTLGEKMTEEEVEMLVAGHEDSNGCIN
YEELLRMVLNG

SEQ ID 17 prosaposin

MYALALLASLLVTALTSPVQDPKICSGGSAVVCRDVKTAVDCRAVKHCQQMVWSKPTAKSLPCDICKTVV
TEAGNLLKDNATEEEILHYLEKTCAWIHDSSLSASCKEVVDSYLPVILDMIKGEMSNPGEVCSALNLCQS
LQEYLAEQNQRQLESNKIPEVDLARVVAPFMSNIPLLLYPQDRPRSQPQPKANEDVCQDCMKLVTDIQTA
VRTNSSFVQGLVDHVKEDCDRLGPGVSDICKNYVDQYSEVAVQMMMHMQPKEICVMVGFCDEVKRVPMRT
LVPATEAIKNILPALELTDPYEQDVIQAQNVIFCQVCQLVMRKLSELIINNATEELLIKGLSKACSLLPA
PASTKCQEVLVTFGPSLLDVLMHEVNPNFLCGVISLCSANPNLVGTLEQPAAAIVSALPKEPAPPKQPEE
PKQSALRAHVPPQKNGGFCEVCKKLVIYLEHNLEKNSTKEEILAALEKGCSFLPDPYQKQCDEFVAEYEP
LLLEILVEVMDPSFVCSKIGVCPSAYKLLLGTEKCVWGPGYWCQNSETAARCNAVDHCKRHVWN

SEQ ID 18 similar to Myosin regulatory light chain 2-A, smooth muscle isoform (Myosin RLC-A) (predicted)

MKPAQSSIRPCRLRPPNWAEGKDFWVSDPKKGQARDGTALAVSRAGALQPSGGAGLGSGAKASHVRPAAL
LSSISLNGLRRGSPPISTLWNAGRGELQDRRNHLPKRGEEAESQFKRRREKDEESGWPIEAAENGRQASF
RFASGGRKWPFLGGWKVRKFVRAAAKLRRDAHSWFSRLSRVSGAGRRAGVARPQEVSSPGGGERGGRREP
RLGSAPGQDLTATMSSKRAKTKTTKKRPQRATSNVFAMFDQSQIQEFKEAFNMIDQNRDGFIDKEDLHDM
LASMGKNPTDEYLDAMMNEAPGPINFTMFLTMFGEKLNGTDPEDVIRNAFACFDEEAIGTIQEDYLRELL

TTMGDRFTDEEVDELYREAPIDKKGNFNYIEFTRILKHGAKDKDD

SEQ ID 19 solute carrier family 4, member 1

MGDMQDHEKVLEIPDRDSEEELEHVIEQIAYRDLDIPVTEMQESEALPTEQTATDYIPTSTSTSHPSSSQ
VYVELQELMMDQRNQELQWVEAAHWIGLEENLREDGVWGRPHLSYLTFWSLLELQKVFSKGTFLLDLAET
SLAGVANKLLDSFIYEDQIRPQDRDELLRALLLKRSHAEDLKDLEGVKPAVLTRSGAPSEPLLPHQPSLE
TKLYCAQAEGGSEEPSPSGILKIPPNSETTLVLVGRASFLVKPVLGFVRLKEAVPLEDLVLPEPVSFLLV
LLGPEAPHIDYTQLGRAAATLMTERVFRVTASLAQSRGELLSSLDSFLDCSLVLPPTEAPSEKALLNLVP
VQKELLRKRYLPRPAKPDPNLYEALADGGKEGPGDEDDPLRRTGRIFGGLIRDIRRRYPYYLSDITDALS
PQVLAAVIFIYFAALSPAVTFGGLLGEKTRNLMGVSELLISTAVQGILFALLGAQPLLVLGFSGPLLVFE
EAFYSFCESNNLEYIVGRAWIGFWLILLVYLVVAFEGSFLVQYISRYTQEIFSFLISLIFIYETFSKLIK
IFQDYPLQESYAPVVMKPKPQGPVPNTALLSLVLMVGTFLLAMMLRKFKNSTYFPGKLRRVIGDFGVPIS
ILIMVLVDTFIKNTYTQKLSVPDGLKVSNSSARGWVIHPLGLYNHEPKWMMPASVLPALLVFILIFLESQ
ITTLIVSKPERKMIKGSGFHLDLLLVVGMGGVAALFGMPWLSATTVRSVTHANALTVMGKASGPGAAAQI
QEVKEQRISGLLVSVLVGLSILMEPILSRIPLAVLFGIFLYMGITSLSGIQLFDRILLLFKPPKYHPDVP
FVKRVKTWRMHLFTGIQIICLAVLWVVKSTPASLALPFVLILTVPLRRLLLPLIFRELELQCLDGDDAKV
TFDEAEGLDEYDEVPMPV

SEQ ID 20 transgelin 2 (predicted)

MANRGPSYGLSREVQQKIEKQYDPDLEQILIQWITTQCRKGVSQPQPGRENFQNWLKDGTVLCELINSLY
PEGQAPVKKIQASTMAFKQMEQISQFLQAAERYGINTTDIFQTVDLWEGKNMACVQRTLMNLGGLAVARD
DGLFSGDPNWFPKKSKENPRNFSDNQLQEGKNVIGLQMGTNRGASQAGMTGYGMPRQIL

SEQ ID 21 tropomyosin 3, gamma

MAGSTTIEAVKRKIQVLQQQADDAEERAERLQREVEGERRAREQAEAEVASLNRRIQLVEEELDRAQERL
ATALQKLEEAEKAADESERGMKVIENRALKDEEKMELQEIQLKEAKHIAEEADRKYEEVARKLVIIEGDL
ERTEERAELAESKCSELEEELKNVTNNLKSLEAQAEKYSQKEDKYEEEIKILTDKLKEAETRAEFAERSV
AKLEKTIDDLEERLYSQLERNRLLSNELKLTLHGLCD

SEQ ID 22 talin

MVALSLRISIGNVVKTMQFEPSTMVYDACRMIRERIPEALAGPPSDFGLFLSDDDPKKGIWLEAGKALDY
YMLRNGDTMEYRKKQRPLKIRMLDGTVKTIMVDDSKTVTDMLMTICARIGITNHDEYSLVRELMEEKKDE
GTGTLRKDKTLLRDEKKMEKLKQKLHTDDELNWLDHGRTLREQGVEEHETLLLRRKFFYSDQNVDSRDPV
QLNLLYVQARDDILNGSHPVSFDKACEFAGFQCQIQFGPHNEQKHKAGFLDLKDFLPKEYVKQKQERKIF
QAHKNCGQMSEIEAKVRYVKLARSLKTYGVSFFLVKEKMKGKNKLVPRLLGITKECVMRVDEKTKEVIQE
WSLTNIKRWAASPKSFTLDFGDYQDGYYSVQTTEGEQIAQLIAGYIDILKKKKSKDHFGLEGDEESTML
EDSVSPKKSTVLQQQYNRVGKAEHGSVALPAIMRSGASGPENFQVGSMPPAQQQVTSGQMHRGHMPPLTS
AQQALTGTINSSMQAVQAAQAAALDDFDALPPLGQDAASKAWRKNKMDESKHEIHSQVDAITAGTASVVNL
TAGDPAETDYTAVGCAVITISSNLTEMSRGVKLLAALLEDEGGNGRPLLQAAKGLAGAVSELLRSAQPAS
AEPRQNLLQAAGNVGQASGELLQQIGESDTDPHFQDILMQLAKAVASAAAALVLKAKSVAQRTEDSQLQT
QVIAAATQCALSTSQLVACTKVVAPTISSPVCQEQLVEAGRLVAKAVEGCVSASQAATEDGQLLRGVGAA
ATAVTQALNELLQHVKAHATGAGEAGRYDQATDTILTYTENIFSSMGDAGEMVRQARILAQATSDLVNAI
KADAEGESDLENSRKLLSAAKILADATAKMVEAAKGAAAHPDSEEQQQRLREAAEGLRMATNAAAQNAIK
KKLVQRLEHAAKQAAASATQTIAAAQHAASAPKASAGPQPLLVQSCKAVAEQIPLLVQGVRGSQAQPDSP
SAQLALIAASQSFLQPGGKMVAAAKASVPTIQDQASAMQLSQCAKNLGTALAELRTAAQKAQEACGPLEM
DSALSVVQNLEKDLQEIKAAAREGKLKPLPGETMEKCTQDLGNSTKAVSSAIAKLLGEIAQGNENYAGIA
ARDVAGGLRSLAQAARGVAALTTDPAVQAIVLDTASDALDKASSLIEEAKKASGHPGDPESQQRLAQVAK
AVTQALNRCVSCLPGQRDVDNALRAVQDASKRLLSDSLPPSTGTFQEAQSRLNEAAAGLNQAATELVQAS
RGTPQDLARASGRFGQDFSTFLEAGVEMAGQAPSQEDRAQVVSNLKGISMSSSKLLLAAKALSTDPAAPN
LKSQLAAAARAVTDSINQLITMCTQQAPGQKECDNALRQLETVRELLENPVQPINDMSYFGCLDSVMENS
KVLGEAMTGISQNAKNGNLPEFGDAIATASKALCGFTEAAAQAAYLVGVSDPNSQAGQQGLVEPTQFARA
NQAIQMACQSLGEPGCTQAQVLSAATIVAKHTSALCNSCRLASARTANPTAKRQFVQSAKEVANSTANLV
KTIKALDGAFTEENRAQCRAATAPLLEAVDNLSAFASNPEFSSVPAQISPEGRAAMEPIVISAKTMLESA

GGLIQTARALAVNPRDPPRWSVLAGHSRTVSDSIKKLITSMRDKAPGQLECETAIAALNSCLRDLDQASL
AAVSQQLAPREGISQEALHTQMLTAVQEISHLIEPLASAARAEASQLGHKVSQMAQYFEPLTLAAVGAAS
KTLSHPQQMALLDQTKTLAESALQLLYTAKEAGGNPKQAAHTQEALEEAVQMMTEAVEDLTTTLNEAASA
AGVVGGMVDSITQAINQLDEGPMGEPEGSFVDYQTTMVRTAKAIAVTVQEMVTKSNTSPEELGPLANQLT
SDYGRLASQAKPAAVAAENEEIGAHIKHRVQELGHGCSALVTKAGALQCSPSDVYTKKELIECARRVSEK
VSHVLAALQAGNRGTQACITAASAVSGIIADLDTTIMFATAGTLNREGAETFADHREGILKTAKVLVEDT
KVLVQNAAGSQEKLAQAAQSSVATITRLADVVKLGAASLGAEDPETQVVLINAVKDVAKALGDLISATKA
AAGKVGDDPAVWQLKNSAKVMVTNVTSLLKTVKAVEDEATKGTRALEATTEHIRQELAVFCSPEPPAKTS
TPEDFIRMTKGITMATAKAVAAGNSCRQEDVIATANLSRRAIADMLRACKEAAFHPEVAPDVRLRALHFG
RECANGYLELLDHVLLTLQKPNPELKQQLTGHSKRVAGSVTELIQAAEAMKGTEWVDPEDPTVIAENELL
GAAAAIEAAAKKLEQLKPRAKPKEADESLNFEEQILEAAKSIAAATSALVKAASAAQRELVAQGKVGAIP
ANALDDGQWSQGLISAARMVAAATNNLCEAANAAVQGHASQEKLISSAKQVAASTAQLLVACKVKADQDS
EAMKRLQAAGNAVKRASDNLVKAAQKAAAPEDQENETVVVKEKMVGGIAQIIAAQEEMLRKERELEEARK
KLAQIRQQQYKFLPSELRDEH

SEQ ID 23 similar to hypothetical protein D4Ertd765e (predicted)

MSAAEVGTFVQRLRGLMSEIAAFPAPTIAAMDGFALGGGLELALACDLRIAASSAVMGLIETTRGLLPGA
GGTQRLPRCLGVALAKELIFTGRRLNGVQAHELGLVNHAVAQNEEGDAAYHRALALAQEILPQAPIAVRL
GKVAIDRGMEVDIASGMAIEHMCYAQNIPTQDRLEGMAAFREKRPPKFVGK

SEQ ID 24 aminoadipate aminotransferase

MNYSRFLTATSLARKTSPIRATVEIMSRAPKDIISLAPGSPNPKVFPFKSAVFTVENGSTIRFEGEMFQR
ALQYSSSYGIPELLSWLKQLQIKLHNPPTVNYSPNEGQMDLCITSGCQDGLCKVFEMLINPGDTVLVNEP
LYSGALFAMKPLGCNFISVPSDDCGIIPEQLKKVLSQWKPEDSKDPTKRTPKFLYTIPNGNNPTGNSLTG
DRKKEIYELARKYDFLIIEDDPYYFLQFTKPWEPTFLSMDVDGRVIRADSLSKVISSGLRVGFITGPKSL
IQRIVLHTQISSLHPCTLSQLMISELLYQWGEEGFLAHVDRAIDFYKNQRDFILAAADKWLRGLAEWHVP
KAGMFLWIKVNGISDAKKLIEEKAIEREILLVPGNSFEVDNSAPSSFFRASFSQVTPAQMDLVFQRLAQL
IKDVS

SEQ ID 25 acyl-CoA synthetase long-chain family member 1

MEVHELFRYFRMPELIDIRQYVRTLPTNTLMGFGAFAALTTFWYATRPKALKPPCDLSMQSVEVTGTTEG
VRRSAVLEDDKLLLYYYDDVRTMYDGFQRGIQVSNDGPCLGSRKPNQPYEWISYKQVAEMAECIGSALIQ
KGFKPCSEQFIGIFSQNRPEWVTIEQGCFTYSMVVVPLYDTLGTDAITYIVNKAELSVIFADKPEKAKLL
LEGVENKLTPCLKIIVIMDSYDNDLVERGQKCGVEIIGLKALEDLGRVNRTKPKPPEPEDLAIICFTSGT
TGNPKGAMVTHQNIMNDCSGFIKATESAFIASPEDVLISFLPLAHMFETVVECVMLCHGAKIGFFQGDIR
LLMDDLKVLQPTIFPVVPRLLNRMFDRIFGQANTSVKRWLLDFASKRKEAELRSGIVRNNSLWDKLIFHK
IQSSLGGKVRLMITGAAPVSATVLTFLRAALGCQFYEGYGQTECTAGCCLSLPGDWTAGHVGAPMPCNYI
KLVDVEDMNYQAAKGEGEVCVKGANVFKGYLKDPARTAEALDKDGWLHTGDIGKWLPNGTLKIIDRKKHI
FKLAQGEYIAPEKIENIYLRSEAVAQVFVHGESLQAFLIAIVVPDVEILPSWAQKRGFQGSFEELCRNKD
INKAILEDMVKLGKNAGLKPFEQVKGIAVHPELFSIDNGLLTPTLKAKRPELRNYFRSQIDELYSTIKI

SEQ ID 26 aldehyde dehydrogenase 1 family, member L1

MKIAVIGQSLFGQEVYCQLRKEGHEVVGVFTIPDKDGKADPLGLEAEKDGRAVFKFPRWRARGQALPEVV
AKYQALGAELNVLPFCSQFIPMEVINAPRHGSIIYHPSLLPRHRGASAINWTLIHGDKKGGFTIFWADDG
LDTGDLLLQKECEVLPDDTVSTLYNRFLFPEGIKGMVQAVRLIAEGTAPRCPQSEEGATYEGIQKKETAK
INWDQPAEAIHNWIRGNDKVPGAWTEACGQKLTFFNSTLNTSGLSTQGEALPIPGAHRPGVVTKAGLILF
GNEHRMLLVKNIQLEDGKMMPASQFFKGSASSDLELTEAELATAEAVRSSWMRILPNVPEVEDSTDFFKS
GAASVDVVRLVEEVKELCDGLELENEDVYMATTFREFIQLLVRKLRGEDDESECVINYVERAVNKLTLQM
PYQLFIGGEFVDAEGSKTYNTINPTDGSVICQVSLAQVSDVDKAVAAAKEAFENGLWGKINARDRGRLLY
RLADVMEQHQEELATIEALDRGAVYTLALKTHVGMSIQTFRYFAGWCDKIQGATIPINQARPNRNLTLTK
KEPVGVCGIVIPWNYPLMMLSWKTAACLAAGNTVVIKPAQVTPLTALKFAELTLKAGIPKGVVNILPGSG
SLVGQRLSDHPDVRKIGFTGSTEVGKHIMKSCALSNVKKVSLELGGKSPLIIFADCDLNKAVQMGMSSVF
FNKGENCIAAGRLFVEESIHNQFVQKVVEEVEKMKIGNPLERDTNHGPQNHEAHLRKLVEYCQRGVKEGA

TLVCGGNQVPRPGFFFQPTVFTDVEDHMYIAKEESFGPIMIISRFADGDVDAVLSRANATEFGLASGVFT
RDINKALYVSDKLQAGTVFINTYNKTDVAAPFGGFKQSGFGKDLGEAALNEYLRIKTVIFEY

SEQ ID 27 arginosuccinate synthetase

MSSKGSVVLAYSGGLDTSCILVWLKEQGYDVIAYLANIGQKEDFEEARKKALKLGAKKVFIEDVSKEFVE
EFIWPAVQSSALYEDRYLLGTSLARPCIARKQVEIAQREGAKYVSHGATGKGNDQVRFELTCYSLAPQIK
VIAPWRMPEFYNRFKGRNDLMEYAKQHGIPIPVTPKSPWSMDENLMHISYEAGILENPKNQAPPGLYTKT
QDPAKAPNTPDVLEIEFKKGVPVKVTNVKDGTTHSTSLDLFMYLNEVAGKHGVGRIDIVENRFIGMKSRG
IYETPAGTILYHAHLDIEAFTMDREVRKIKQGLGLKFAELVYTGFWHSPECEFVRHCIDKSQERVEGKVQ
VSVFKGQVYILGRESPLSLYNEELVSMNVQGDYEPIDATGFININSLRLKEYHRLQSKVTAK

SEQ ID 28 dimethylglycine dehydrogenase

MLRLGALRLRGLALRSSQGRPSSAGLREGQESPPSPPEWKDRAETVIIGGGCVGVSLAYHLAKAGMRDVV
LLEKSELTAGSTWHAAGLTTYFHPGINLKKIHYDSIKLYERLEEETGQVVGFHQPGSIRLATTPERVDEF
KYQMTRTNWHATEQYIIEPEKIHELFPLLNMDKILAGLYNPGDGHIDPYSLTMALATGARKYGALLKYPA
PVTSLKPRPDGTWDVETPQGSVRANRIVNAAGFWAREVGKMIGLDHPLIPVQHQYVITSTIPEVKALKRE
LPVLRDLEGSYYLRQERDGLLFGPYESQEKMKLQASWVAHGVPPGFGKELFESDLDRITEHVEAAMEMVP
VLKKADIINIVNGPITYSPDILPMVGPHQGVRNYWVAIGFGYGIIHAGGVGTYLSDWILHGEPPFDLIEL
DPNRYGKWTTTQYTEAKARESYGFNNIVGYPKEERFAGRPTQRVSGLYKILESKCSMGFHAGWEQPHWFY
KPGQDTQYRPSFRRTNWFEPVGSEYKQVMQRVGVIDLSPFGKFNIKGQDSTQLLDHLCANVIPKVGFTNI
SHMLTPRGRVYAELTVSHQSPGEFLLITGSGSELHDLRWIEEAAVRGGYDVEIRNITDELGVLGVAGPYA
RRVLQKLTSEDLSDDVFRFLQTKSLKISDIPVTAIRISYTGELGWELYHRRBDSAALYERIMNAGQEEGI
DNFGTYALNALRLEKAFRAWGSEMNCDTNPLEAGLDYFIKLNKPANFTGKQALKQIKAKGLKRRLVCLTL
ATDNVDPEGNESVWYKGKVIGNTTSGSYSYSIQKSLAPAYVPVELSEVGQQVEVELLGKNYPATIIQEPL
VLTEPTRTRLQKDGRKS

SEQ ID 29 dihydropyrimidinase

MAPQERLLIRGGRVVNDDFSQVADVLVEDGVVRALGRDLLPPGDTSRGLRILDAAGKLVLPGGIDTHTHM
QFPFMGSQSVDDFHQGTKAALAGGTTMIIDFAIPQKGSSLIEAFETWRNWADPKVCCDYSLHVAVTWWSD
KVKEEMKTLAQDKGVNSFKMFMAYKDLYMVQDQQMYAAFSQCKELGAIAQVHAENGDLIAEGAKKMLALG
ITGPEGHELCRPEAVEAEATLRAITIASAVNCPLYIVHVMSKSAAKVIADAKREGKVVYGEPIAAGLGTD
GTQYWNKEWRHAAHHVMGPPLRPDPSTPGFLMNLLANGDLTTTGSDNCTFNTCQKALGKDDFTKIPNGVN
GVEDRMSVIWEKGVHSGKMDENRFVAVTSTNAAKIFNLYPKKGRIAVGSDADMVIWDPEATRTISAKTHH
QAVNFNIFEGMVCHGVPLVTISRGRVVYEAGVFDVTAGHGKFIPRQPFAEFIYKRVKQRDQTCTPIVKR
APYKGEVITLKPRETKEDDTAGTRMQGHS

SEQ ID 30 phosphodiesterase I/nucleotide pyrophosphatase 2

MARQGCLGSFQVISLFTFAISVNICLGFTASRIKRAEWDEGPPTVLSDSPWTNTSGSCKGRCFELQEVGP
PDCRCDNLCKSYSSCCHDFDELCLKTVRGWECTKDRSGEVRNEENACHCPEDCLSRGDCCTNYQVVCKGE
SHWVDDAARNQSSECLQVCPPPLIIFSVDGFRASYMKKGSKVMPNIEKLRSCGTHVPYTRPVYPTKTFPN
LYTLATGLYPESHGIVGNSMYDPVFDASFHLRGREKFNHRWWGGQPLWITATKQGVRAGTFFWSVSIPHE
RRILTILQWLSLPDNERPSVYAFYSEQPDFSGHKYGPFGPEMTNPLREIDKTVGQLMDGLKQLRLHRCVN
VIFVGDHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRAKSINNSKYDPKTIIANLTCKKPDQHFKP
YMKQHLPKRLHYANNRRIEDIHLLVDRRWHVARKPLDVYKKPSGKCFFQGDHGFDNKVNSMQTVFVGYGP
TFKYRTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLLRTNTFRPTMPDEVSRPNYPGIMYLQSEF
DLGCTCDDKVEPKNKLEELNKRLHTKGSTEAETGKFRGSKHENKKNLNGSVEPRKERHLLYGRPAVLYRT
SYDILYHTDFESGYSEIFLMPLWTSYTISKQAEVSSIPEHLTNCVRPDVRVSPGFSQNCLAYKNDKQMSY
GFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWAYFQRVLVKKYASERNGVNVISGPIFDYNYDGLRD
TEDEIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNDESCNSSEDESKWVEE
LMKMHTARVRDIEHLTGLDFYRKTSRSYSEILTLKTYLHTYESEI

SEQ ID 31 fatty acid amide hydrolase

MVLSEVWTTLSGVSGVCLACSLLSAAVVLRWTGRQKARGAATRARQKQRASLETMDKAVQRFRLQNPDLD

SEALLTLPLLQLVQKLQSGELSPEAVFFTYLGKAWEVNKGTNCVTSYLTDCETQLSQAPRQGLLYGVPVS
LKECFSYKGHDSTLGLSLNEGMPSESDCVVVQVLKLQGAVPFVHTNVPQSMLSFDCSNPLFGQTMNPWKS
SKSPGGSSGGEGALIGSGGSPLGLGTDIGGSIRFPSAFCGICGLKPTGNRLSKSGLKGCVYGQTAVQLSL
GPMARDVESLALCLKALLCEHLFTLDPTVPPLPFREEVYRSSRPLRVGYYETDNYTMPSPAMRRALIETK
QRLEAAGHTLIPFLPNNIPYALEVLSAGGLFSDGGRSFLQNFKGDFVDPCLGDLILILRLPSWFKRLLSL
LLKPLFPRLAAFLNSMRPRSAEKLWKLQHEIEMYRQSVIAQWKAMNLDVLLTPMLGPALDLNTPGRATGA
ISYTVLYNCLDFPAGVVPVTTVTAEDDAQMELYKGYFGDIWDIILKKAMKNSVGLPVAVQCVALPWQEEL
CLRFMREVEQLMTPQKQPS

SEQ ID 32 guanidinoacetate methyltransferase

MSSSAASPLFAPGEDCGPAWRAAPAAYDTSDTHLQILGKPVMERWETPYMHSLAAAAASRGGRVLEVGFG
MAIAASRVQQAPIKEHWIIECNDGVFQRLQNWALKQPHKVVPLKGLWEEEAPTLPDGHFDGILYDTYPLS
EETWHTHQFNFIKTHAFRLLKPGGILTYCNLTSWGELMKSKYTDITAMFEETQVPALLEAGFQRENICTE
VMALVPPADCRYYAFPQMITPLVTKH

SEQ ID 33 L-gulono-gamma-lactone oxidase

MVHGYKGVQFQNWAKTYGCSPEVYYQPTSVEEVREVLALAREQKKKVKVVGGGHSPSDIACTDGFMIHMG
KMNRVLQVDKEKKQITVEAGILLADLHPQLDEHGLAMSNLGAVSDVTVAGVIGSGTHNTGIKHGILATQV
VALTLMTADGEVLECSESRNADVFQAARVHLGCLGIILTVTLQCVPQFHLQETSFPSTLKEVLDNLDSHL
KRSEYFRFLWFPHTENVSIIYQDHTNKAPSSASNWFWDYAIGFYLLEFLLWTSTYLPCLVGWINRFFFWM
LFNCKKESSNLSHKIFTYECRFKQHVQDWAIPREKTKEALLELKAMLEAHPKVVAHYPVEVRFTRGDDIL
LSPCFQRDSCYMNIIMYRPYGKDVPRLDYWLAYETIMKKFGGRPHWAKAHNCTRKDFEEMYPTFHKFCDI
REKLDPTGMFLNSYLEKVFY

SEQ ID 34 3-hydroxyisobutyrate dehydrogenase

MAASLGFRGAASGLRYWSGRRRPVGSLAAVCSRSMASKTPVGFIGLGNMGNPMAKNLIKHGYPLILYDVF
PDVCKEFKEAGEQVASSPADVAEKADRIITMLPSSMNSIEVYSGANGILKKVKKGSLLIDSSTIDPSVSK
ELAKEVEKMGAVFMDAPVSGGVGAARSGNLTFMVGGVEEEFAAAQELLGCMGSNVLYCGAVGSGQSAKIC
NNMLLAISMIGTAEAMNLGIRSGLDPKLLAKILNMSSGRCWSSDTYNPVPGVMDGVPSSNNYQGGFGTTL
MAKDLGLAQDSATSTKTPILLGSVAHQIYRMMCSKGYSKKDFSSVFQYLREEETF

SEQ ID 35 ketohexokinase

MEEKQILCVGLVVLDIINVVDKYPEEDTDRRCLSQRWQRGGNASNSCTVLSLLGARCAFMGSLAHGHVAD
FLVADFRRRGVDVSQVAWQSQGDTPCSCCIVNNSNGSRTILYDTNLPDVSAKDFEKVDLTRFKWIHIEG
RNASEQVKMLQRIEQYNATQPLQQKVRVSVEIEKPREELFQLFGYGEVVFVSKDVAKHLGFRSAGEALKG
LYSRVKKGATLICAWAEEGADALGPDGQLLHSDAFPPPRVVDTLGAGDTFNASVIFSLSKGNSMQEALRF
GCQVAGKKCGLQGFDGIV

SEQ ID 36 kynureninase (L-kynurenine hydrolase)

MEPSPLELPVDAVRRIAAELNCDPTDERVALRLDEEDKLKRFKDCFYIPKMRDLPSIDLSLVNEDDNAIY
FLGNSLGLQPKMVKTYLEEELDKWAKIGAYGHEVGKRPWIIGDESIVTLMKDIVGAHEKEIALMNALTVN
LHLLLLSFFKPTPKRHKILLEAKAFPSDHYAIESQIQLHGLDVEKSMRMIKPREGEETLRMEDILEVIEK
EGDSIAVVLFSGLHFYTGQLFNIPATTQAGHAKGCFVGFDLAHAVGNVELHLHDWDVDFACWCSYKYLNS
GAGGLAGAFIHEKHAHTIKPALVGWFGHELSTRFNMDNKLQLIPGVNGFRISNPPILLVCSLHASLEIFQ
QATMTALRRKSILLTGYLEYLLKHYHGGNDTENKRPVVNIITPSRAEERGCQLTLTFSISKKGVFKELEK
RGVVCDKREPEGIRVAPVPLYNSFHDVYKFIRLLTAILDSTERN

SEQ ID 37 Pyruvate carboxylase

MLKFQTVRGGLRLLGVRRSSTAPVASPNVRRLEYKPIKKVMVANRGEIAIRVFRACTELGIRTVAVYSEQ
DTGQMHRQKADEAYLIGRGLAPVQAYLHIPDIIKVAKENGVDAVHPGYGFLSERADFAQACQDAGVRFIG
PSPEVVRKMGDKVEARAIAIAAGVPVVPGTNSPINSLHEAHEFSNTYGFPIIFKAAYGGGGRGMRVVHSY
EELEENYTRAYSEALAAFGNGALFVEKFIEKPRHIEVQILGDQYGNILHLYERDCSIQRRHQKVVEIAPA

THLDPQLRSRLTSDSVKLAKQVGYENAGTVEFLVDKHGKHYFIEVNSRLQVEHTVTEEITDVDLVHAQIH
VSEGRSLPDLGLRQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISP
HYDSLLVKVIAHGKDHPTAAATKMSRALAEFRVRGVKTNIPFLQNVLNNQQFLAGIVDTQFIDENPELFQL
RPAQNRAQKLLHYLGHVMVNGPTTPIPVKVSPSPVDPIVPVVPIGPPPAGFRDILLREGPEGFARAVRNH
QGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFNNLFSIENWGGATFDVAMRFLYECPWRRLQEL
RELIPNIPFQMLLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRIFDSLNYLPNMLLGMEAAGSAGGVV
EAAISYTGDVADPSRTKYSLEYYMGLAEELVRAGTHILCIKDMAGLLKPAAATMLVSSLRDRFPDLPLHI
HTHDTSGSGVAAMLACAQAGADVVDVAVDSMSGMTSQPSMGALVACTKGTPLDTEVPLERVFDYSEYWEG
ARGLYAAFDCTATMKSGNSDVYENEIPGGQYTNLHFQAHSMGLGSKFKEVKKAYVEANQMLGDLIKVTPS
SKIVGDLAQFMVQNGLSRAEAEAQAEELSFPRSVVEFLQGYIGIPHGGFPEPFRSKVLKDLPRIEGRPGA
SLPPLNLKELEKDLIDRHGEEVTPEDVLSAAMYPDVFAQFKDFTATFGPLDSLNTRLFLQGPKIAEEFEV
ELERGKTLHIKALAVSDLNRAGQRQVFFELNGQLRSILVKDTQAMKEMHFHPKALKDVKGQIGAPMPGKV
IDVKVAAGAKVVKGQPLCVLSAMKMETVVTSPMEGTIRKVHVTKDMTLEGDDLILEIE

SEQ ID 38 pyruvate dehydrogenase kinase 2 subunit p45 (PDK2)

MRWFRALLKNASLAGAPKYIEHFSKFSPSPLSMKQFLDFGSSNACEKTSFTFLRQELPVRLANIMKEINL
LPDRVLSTPSVQLVQSWYVQSLLDIMEFLDKDPEDHRTLSQFTDALVTIRNRHNDVVPTMAQGVLEYKDT
YGDDPVSNQNIQYFLDRFYLSRISIRMLINQHTLIFDGSTNPAHFKHIGSIDPNCSVSDVVKDAYDMAKL
LCDKYYMASPDLEIQEVNATNATQPIHMVYVPSHLYHMLFELFKNAMRATVESHESSLTLPPIKIMVALG
EEDLSIKMSDRGGGVPLRKIERLFSYMYSTAPTPQPGTGGTPLAGFGYGLPISRLYAKYFQGDLQLFSME
GFGTDAVIYLKALSTDSVERLPVYNKSAWRHYQTIQEAGDWCVPSTEPKNTSTYRVS

SEQ ID 39 solute carrier family 10, member 1

MEVHNVSAPFNFSLPPGFGHRATDKALSIILVLMLLIMLSLGCTMEFSKIKAHLWKPKGVIVALVAQFG
IMPLAAFLLGKIPHLSNIEALAILICGCSPGGNLSNLFTLAMKGDMNLSIVMTTCSSFSALGMMPLLLYV
YSKGIYDGDLKDKVPYKGIMISLVIVLIPCTIGIVLKSKRPHYVPYILKGGMIITFLLSVAVTALSVINV
GNSIMFVMTPHLLATSSLMPFSGFLMGYILSALFQLNPSCRRTISMETGFQNIQLCSTILNVTFPPEVIG
PLFFFPLLYMIFQLAEGLLIIIIFRCYEKIKPPKDQTKITYKAAATEDATPAALEKGTHNGNIPPLQPGP
SPNGLNSGQMAN

SEQ ID 40 solute carrier family 37 (glycerol-6-phosphate transporter), member 4

MAAQGYGYYRTVIFTAMFGGYSLYYFNRKTFSFVMPSLVDEIALDKDDLGLITSSQSAAYAISKFVSGVL
SDQMSARWLFSSGLLLVGLVNVVFSWSSTVTAFAALWFLNGLAQGLGWPPCGKILRKWFEPSQFGTWWAV
LSTSMNLAGSLGPILATILAQSYSWRSTLALSGSLCVVVSFFCLLLIHNEPADVGLRNLDPAPSKGKKGS
SKEESTLQDLLLSPYLWVLSTGYLVVFGVKTCCTDWGQFPLIQERGQSALVGSSYISALEVGGLVGSIAA
GYLSDRAMAKAGLSVYGNPRHSLLLLMMAGMAASMFLFRVTVTSDSPKIWILVLGAVFGFSSYGPIALFG
VIANESAPPNLCGTSHAIVGLMANVGGFLAGLPFSTIAKHYSWSTAFWVAEVICIASTVVFFLLRNIRTK
MGRVSKKAE

SEQ ID 41 similar to Sulfotransferase K2 (rSULT1C2A)

MLCSELETGAWGALAAELGASEAASVSVDQGAGSALLGLESEAAWAASLFASRPAAGASVFLASPPSWGA
RSGARVPSALGARLAGLGGLVAHRGPGPGRLLLRHWPGLGLLLRGALALLGLGLGRLLLTGRLLVGLLFL
SPCLPIWSVLGLLCALLDLSGVGGRLGFSLLWGPLFLGARAFSLLVFVLGLLLIHIVTLLAQLASHHGVP
LLRGADEKSSLHFFKLYLLPLPHNKLWATNQTPINERTEMVYSFRQLKQAARIQGKRWTHPEGQLGSSNI
CVPLTHLSSSPKYLYVARNAKDCMVSYYHFYRMSQVLPNPGTWNEYFETFNGKVSCGSWFDHVKGWWEI
RDRYQILFLFYEDMKRDPKREIQKVMQFMGKNLDEEVVDKIVLETSFEKMKDNPLTNFSTIPKTIMDQSI
SPFMRKGIVGDWKNHFTVAQNERFDEIYEQKMDGTSLNFCMEL

SEQ ID 42 arachidonate 5-lipoxygenase activating protein

MDQEAVGNVVLLAIVTLISVVQNAFFAHKVELESKAQSGRSFQRTGTLAFERVYTANQNCVDAYPTFLVV
LWTAGLLCSQVPAAFAGLMYLFVRQKYFVGYLGERTQSTPGYIFGKRIILFLFLMSLAGILNHYLIFFFG
SDFENYIRTTFTISPLLLIP

SEQ ID 43 adenine phosphoribosyl transferase (predicted)

MSESELQLVARRIRSFPDFPIPGVLFRDISPLLKDPDSFRASIRLLAGHLKSTHGGKIDYIAGLDSRGFL
FGPSLAQELGVGCVLIRKRGKLPGPTVSASYSLEYGKAELEIQKDALEPGQKVVIVDDLLATGGTMCAAC
ELLSQLRAEVVECVSLVELTSLKGREKLGPVPFFSLLQYE

SEQ ID 44 brain abundant, membrane attached signal protein 1

MGSKLSKKKKDGYNVNDEKAKDKDKKAEGAGTEEEGTQKESEPQAAADATEVKESAEEKPKDAADGEAKAE
EKEADKAAAKEEAPKAEPEKSEGAAEEQPEPAPAPEQEAAAPGPAAGGEAPKAGEASAESTGAADGAPQE
EGEAKKTEAPAAGPEAKSDAAPAASDSKPSTEPAPSSKETPAASEAPSSAAKAPAPAAPAAEPQAEAPVA
SSEQSVAVKE

SEQ ID 45 benzodiazepin receptor

MSQSWVPAVGLTLVPSLGGFMGAYFVRGEGLRWYASLQKPSWHPPRWTLAPIWGTLYSAMGYGSYIIWKE
LGGFTEEAMVPLGLYTGQLALNWAWPPIFFGARQMGWALVDLMLVSGVATATTLAWHRVSPPAARLLYPY
LAWLAFATMLNYYVWRDNSGRRGGSRLTE

SEQ ID 46 carbonyl reductase

MSSDRPVALVTGANKGIGFAIVRDLCRKFLGDVVLTARDESRGHEAVKQLQTEGLSPRFHQLDIDNPQSI
RALRDFLLQEYGGLNVLVNNAGIAFKVVDPTPFHIQAEVTMKTNFFGTQDVCKELLPIIKPQGRVVNVSS
SVSLRALKSCSPELQQKFRSETITEEELVGLMNKFIEDAKKGVHAKEGWPNSAYGVTKIGVTVLSRIYAR
KLNEERREDKILLNACCPGWVRTDMAGPKATKSPEEGAETPVYLALLPPGAEGPHGQFVQDKKVEPW

SEQ ID 47 CD47 antigen (Rh-related antigen, integrin-associated signal transducer)

MWPLAAALLLGSCCCGSAQLLLSKVKSVEFTSCNDTVVIPCKVLNVEAQSTDEMFVKWKLNKSYIFIYDG
NKNSTTREQNFTSAKISVSDLLKGIASLTMDTHEAVVGNYTCEVTELSREGKTVIELKNRPVSWFSTNEK
ILIVIFPILAILLFWGKFGILTLKYKSSHTNKRIILLLVAGLALTLIVVVGAILFIPGEKPVKNASGLGL
IVISTGILILLQYNVFMTAFGMTSFTIAILITQVLGYVLAVVGMCLCIMACHPVHGPLLISGLGIIALAE
LLGLVYMKFVASNQRTIQPPRNN

SEQ ID 48 embigin

MRSHTGLRALVAPGCSLLLLYLLAATRPDRAVGDPADSAFTSLPVREEMMAKYANLSLETYNISLTEQTR
VSEQNITLERPSHLELECTFTATEDVMSMNVTWKKDDALLETTDGFNTTKMGDTLYSQYRFTVFNSKQMG
KYSCFLGEELRGTFNIRVPKVHGKNKPLITYVGDSTVLKCECQNCLPLNWTWYMSNGTAQVPIDVHVNDK
FDINGSYANETKLKVKHLLEEDGGSYWCRAAFPLGESEEHIKLVVLSFMVPLKPFLAIIAEVILLVAIIL
LCEVYTQKKKNDPDDGKEFEQIEQLKSDDSNGIENNVPRYRKTDSGDQ

SEQ ID 49 myeloid differentiation primary response gene 116

MAPSPRPQHVLHWKEAHSFYLLSPLMGFLSRAWSRLRGPEVSEAWLAETVAGANQIEADALLTPPPVSEN
HLPLRETEGNGTPEWSKAAQRLCLDVEAQSSPPKTWGLSDIDEHNGKPGQDGLREQEVEETAGLPTLQPL
HLQGADKKVGEVVAREEGVSELAYPTSHWEGGPAEDEEDTETVKKAHQASAASIAPGYKPSTSVYCPGEA
EHRATEEKGTDNKAEPSGSHSRVWEYHTRERPKQEGETKPEQHRAGQSHPCQNAEAEEGGPETSVCSGSA
PLKAWVYRPGEDTEEEEDSDLDSAEEDTAHTCTTPHTSAFLKAWVYRPGEDTEEEDDGDWDSAEEDTAQS
CTTPHTSAFLKAWVYRPGEDTEEEDDSENVAPVDSETVDSCQSTQHCLPVEKTKGCGEAEPPPFQVAFYL
PGQKPAPPWAAPKLPLRLQKRLRSFKAPARNQDPEIPLKGRKVHESEKVTVHFLAVWAGPAQAARRGPWE
QFARDRSRFARRIAQAEEQLGPYLTPAFRARAWTRLRNLPLPLSSSSLPLPEPCSSTEATPLSQDVTTPS
PLPSEIPPPSLDLGGRRG

SEQ ID 50 N-myc downstream regulated gene 2

MAELQEVQITEEKPLLPGQTPEAAKTHSVETPYGSVTFTVYGTPKPKRPAIFTYHDVGLNYKSCFQPLFQ
FGDMQEIIQNFVRVHVDAPGMEEGAPVFPLGYQYPSQDQLADMIPCILQYLNFSTIIGVGVGAGAYILSR
YALNHPDTVEGLVLINIDPNAKGWMDWAAHKLTGLTSSIPEMILGHLFSQEELSGNSELIQKYRSLITHA

PNLENIELYWNSYNNRRDLNFERGGEMTLKCPVMLVVGDQAPHEDAVVECNSKLDPTQTSFLKMADSGGQ
PQLTQPGKLTEAFKYFVQGMGYMASSCMTRLSRSRTASLTSAASIDGSRSRSRTLSQSSESGTLPSGPPG
HTMEVSC

SEQ ID 51 expressed in non-metastatic cells 1, protein (NM23A) (nucleoside diphosphate kinase)

MANSERTFIAIKPDGVQRGLVGEIIKRFEQKGFRLVGLKFIQASEDLLKEHYIDLKDRPFFSGLVKYMHS
GPVVAMVWEGLNVVKTGRVMLGETNPADSKPGTIRGDFCIQVGRNIIHGSDSVESAEKEISLWFQPEELV
DYKSCAQNWIYE

SEQ ID 52 neuregulin 1

MSERKEGRGKGKGKKKDRGSRGKPGPAEGDPSPALPPRLKEMKSQESAAGSKLVLRCETSSEYSSLRFKW
FKNGNELNRKNKPENIKIQKKPGKSELRINKASLADSGEYMCKVISKLGNDSASANITIVESNEFITGMP
ASTETAYVSSESPIRISVSTEGANTSSSTSTSTTGTSHLIKCAEKEKTFCVNGGECFTVKDLSNPSRYLC
KCPNEFTGDRCQNYVMASFYKAEELYQKRVLTITGICIALLVVGIMCVVAYCKTKKQRQKLHDRLRQSLR
SERSNLVNIANGPHHPNPPPENVQLVNQYVSKNVISSEHIVEREVETSFSTSHYTSTAHHSTTVTQTPSH
SWSNGHTESVISESNSVIMMSSVENSRHSSPAGGPRGRLHGLGQPRDNSFLRHARETPDSYRDSPHSERY
VSAMTTPARMSPVDFHTPSSPKSPPSEMSPPVSSMTVSMPSVAVSPFVEEERPLLLVTPPRLREKKYDHH
PQQLNSFHHNPAHQSTSLPPSPLRIVEDEEYETTQEYESVQEPVKEVTNSRRAKRTKPNGHIANRLEMDS
NTSSVSSNSESETEDERVGEDTPFLGIQNPLAASLEVAPAFRLAESRTNPAGRFSTQEELQARLSSVIAN
QDPIAV

SEQ ID 53 oxidized low density lipoprotein (lectin-like) receptor 1

MAFDDKMKPVNGQPDQKSCGKKPKGLHLLSSTWWCPAAVTLAILCLVLSVTLIVQQTQLLQVSDLLKQYQ
ANLTQQDHILEGQMSAQKKAENASQESKRELKEQIDTLTWKLNEKSKEQEKLLQQNQNLQEALQRAVNAS
EESKWELKEQIDILNWKLNGISKEQKELLQQNQNLQEALQKAEKYSEESQRELKEQIDTLSWKLNEKSKE
QEELLQQNQNLQEALQRAANSSGPCPQDWIWHKENCYLFHGPFNWEKSRENCLSLDAQLLQISTTDDLNF
VLQATSHSTSPFWMGLHRKNPNHPWLWENGSPLSFQFFRTRGVSLQMYSSGTCAYIQGGVVFAENCILTA
FSICQKKANLLLTQ

SEQ ID 54 protein tyrosine kinase 9

MSHQTGIQASEDVKEIFARARNGKYRLLKISIENEQLVVGSCSPPSDSWEQDYDPFVLPLLEDKQPCYVL
FRLDSQNAQGYEWIFIAWSPDHSHVRQKMLYAATRATLKKEFGGGHIKDEVFGTVKEDVSLHGYRKYLLS
QSSPAPLTAAEEELRQIKISEVQTDVSVDTKHQTLQGVAFPISRDAFQALEKLSKRQLNYVQLEIDIKNE
TILANTENTELKDLPKRIPKDSARYHFFLYKHSHEGDYLESIVFIYSMPGYTCSIRERMLYSSCKSPLL
DIVERQLQMDVIRKIEIDNGDELTADFLYDEVHPKQHAHKQSFAKPKGPAGKRGIRRLIRGPAEAEATTD

SEQ ID 55 syndecan binding protein

MSLYPSLEDLKVDKVIQAQTASSANPASQAFFLNVVDSAAFPPGGNLYPKLYPELSQYMGLSLSEAEICE
NMPVVSGAPTQGQLVARPSSVNYMVAPVTGNDAGIRRAEIKQGIREVILCKDQDGKIGLRLKSVDNGIFV
QLVQANSPASLVGLRFGDQVLQINGENCAGWSSDKAHKVLKQAFGEKITMTIRDRPFERTVTMHKDSSGH
VGFIFKSGKITSIVKDSSAARNGLLTDHHICEINGQNVIGLKDAQIADILSTAGTVVTITIMPAFIFEHI
IKRMAPSIMKSLMDHTIPEV

SEQ ID 56 serine protease inhibitor, Kazal type 1

MKVAIIFLLSALALLSLAGNPPAEVNGKTPNCPKQIMGCPRIYDPVCGTNGITYPSECSLCFENRKFGTS
IHIQRRGTC

SEQ ID 57 tyrosine 3/tryptophan 5 -monooxygenase activation protein, eta polypeptide

MGDREQLLQRARLAEQAERYDDMASAMKAVTELNEPLSNEDRNLLSVAYKNVVGARRSSWRVISSIEQKT
MADGNEKKLEKVKAYREKIEKELETVCNDVLALLDKFLIKNCNDFQYESKVFYLKMKGDYYRYLAEVASG
EKKNSVVEASEAAYKEAFEISKEHMQPTHPIRLGLALNFSVFYYEIQNAPEQACLLAKQAFDDAIAELDT
LNEDSYKDSTLIMQLLRDNLTLWTSDQQDEEAGEGN

SEQ ID 58 S100 calcium binding protein A10

MPSQMEHAMETMMLTFHRFAGEKNYLTKEDLRVLMEREFPGFLENQKDPLAVDKIMKDLDQCRDGKVGFQ
SFLSLVAGLIIACNDYFVVHMKQKK

SEQ ID 59 B-cell translocation gene 2

MSHGKRTDMLPEIAAAVGFLTSLLRTRGCVSEQRLKVFSRALQDALTDHYKHHWFPEKPSKGSGYRCIRI
NHKMDPIISKVASQIGLSQPQLHQLLPSELTLWVDPYEVSYRIGEDGSICVLYEEAPVATSYGLLTCKNQ
MMLGRSSPSKNYVMTVSS

SEQ ID 60 3'-phosphoadenosine 5'-phosphosulfate synthase 2 (predicted)

MKSTDIVLKANGLLLDLKQGKSLTLPVNDNDERHKIQFERCQGSSQEQLTNRGRMIRVSDKLIAFKRRVK
CVRMTPPRMISSLTTRLELTVKYPVGSNRTKGERMALRFFKARVVAEKSVRRTLPCSQEHVPAREQQECG
EIPVPVKFKSQWVVFQGCCGVRLHCAGTANGRKGSIRRGTGRLGCDQQKSTNVVYQAHHVSRNKRGQVVG
TRGGFRGCLSGAGKTTISFALEEYLVSHAIPCYSLDGDNVRHGLNKNLGFSAGDREENIRRIAEVAKLFA
DAGLVCITSFISPFAKQPLISSIGYVKRVASVPEASAPVDRKLLWDRENARKIHESAGLPFFEIFVDAPL
NICESRDVKGLYKRARAGEIKGFTGIDSNYEKPETPECVLKTNLSSVSDCVQQVVELLQEQSIVPHTTIK
GIHELFVPENKIDQIRAELETLPSLPITKLDLQWVQILSEGWATPLKGFMREKEYLQTLHFDTLLDDGVI
NMSIPIVLPVSGDDKARLEGCSKFALMYEGRRVALLQDPEFYEHRKEERCSRVWGTASAKHPHIKADSCS
MPYVSTGIVPLVDIRIAMNTGVQISLQDMVMEGGDWLVGGDLQVLERIRWNDGLDQYRLTPLELKQKCKD
MDAGSHKGEPAAGTREQGQAILLTPNLGSPFLDAVFAFQLRNPVHNGHALLMQDTRRRLLERGYKHPVLL
LHPLGGWTKDDDVPLDWRMKQHAAVLEEGILDPKSTIVAIFPSPMLYAGPTEVQWHCRCRMIAGANFYIV
GRDPAGMPHPETKKDLYEPTHGGKVLSMAPGLTSVEIPFRVAAYNKIKKAMDFYDPARHDEFDFISGTR
MRKLAREGEDPPDGFMAPKAWKVLTDYYRSLEKIN

SEQ ID 61 biogenesis of lysosome-related organelles complex-1, subunit 2

MAAAAEGVPATRREEPPRDDAAVETAEEAKEPAEADINELCRDMFSKMATYLTGELTATSEDYKLLENMN
KLTSLKYLEMKDIAINISRNLKDLNQKYAGLQPYLDQINVEEQVAALEQAAYKLDAYSKKLEAKYKKLE
KR

SEQ ID 62 Similar to RIKEN cDNA 6720467C03 (predicted)

MPSGLPCPSLDRRGARRVATETPDRSGGAPGALSAPGGSGGVSCSRAPWRPGAALVGLPSRVPGGAKGKM
LRRNLDDRDAQTKQLQDAVTNVEKHFGELCQIFAAYVRKTARLRDKADLLVNEINLYASTETPNLKQGLK
DFADEFAKLQDYRQAEVERLEAKVVEPLKAYGTIVKMKRDDLKATLTARNREAKQLSQLERTRQRNPSDR
HAETELQRATIDATRTTRHLEETIDNFEKQKIKDIKNILSEFITIEMLFHGKALEVLTAAYQNIQKIDED
EDLEVFRNSLYLSDYPSRLDIVRANSKSPLQRPLSTKCTSGTGQVATCRTKKDQHVEDEDDEELDVTEDE
N

SEQ ID 63 transformed mouse 3T3 cell double minute 2 (predicted)

MNLFGFASCQLILVPSQGRTAGGHDHSRSNLDEKDSGSAGSGFSPSPTSLVPTGWFRSPNTVLLLSNVFH
SFRDPIMTVVLLLPLRFECKHNRAARYRDPSPLPPRSSTARPLSPAQGPPPSPIGSGGGAGRARSDARMA
LGFRPRRLAGSRGDGLLAPGHDALIHQAVCRRAETDTPGTSPTTVLLLRASRPMCNTNMSVSTEQAAGTS
QIPASEQETLVRPKPLLLKLLKSVGAQKDIYTMKEIIFYIGQYIMTKRLYDEKQQHIVYCSNDLLGDVFG
VPSFSVKEHRKIYAMIYRNLVVVSQQDSGTSPSESRCQPEGGSDLKDPVQASQEEKPSSSDVVSRPSTSS
RRRAISETEENTDELPGERQRKRHRALSFDESLGLCVLREICCERSSSSEATDTPSHQDLDDGVSDHSAD
CLDQDSVSDQFSVEFEVESLDSEDYSLSDEGHELSDEDDEVYRVTVYQAGESDADSFEGDPEISLADYWK
CTSCNEMNPPLPSHCNRCWTLRENWLPDDKGKDKVEISEKAKLESSDQAEEGLDVPDGKKVTEDDAKESS
AEDSEEKVAQMLLSQESDDYSQPSTSSSIVYSSQESGKELKEDTQDKEESMESSFSLNAIEPCVICQGRP
KNGCIVHGKTGHLMSCFTCAKKLKKRNKPCPVCRQPIQMIVLTYFN

SEQ ID 64 fatty acid synthase

MEEVVIAGMSGKLPESENLQEFWDNLIGGVDMVTDDDRRWKAGLYGLPRRSGKLKDLSRFDASFFGVHPK

QAHTMDPQLRLLLEVTYEAIVDGGINPDSLRGTHTGVWVGVSGSETSEALSRDPETLVGYSMVGCQRAMM
ANRLSFFFDFRGPSIALDTACSSSLMALQNAYQAIHSGQCPAAIVGGINVLLKPNTSVQFLRLGMLSPEG
TCKAFDTAGNGYCRSEGVVAVLLTKKSLARRVYATILNAGTNTDGFKEQGVTFPSGDIQEQLIRSLYQSA
GVAPESFEYIEAHGTGTKVGDPQELNGITRALCATRQEPLLIGSTKSNMGHPEPASGLAALAKVLLSLEH
GLWAPNLHFHSPNPEIPALLDGRLQVVDQPLPVRGGNVGINSFGFGGSNVHIILRPNTQPPPAPAPHATL
PRLLRASGRTPEAVQKLLEQGLRHSQDLAFLSMLNDIAAVPATAMPFRGYAVLGGERGGPEVQQVPAGER
PLWFICSGMGTQWRGMGLSLMRLDRFRDSILRSDEAVKPFGLKVSQLLLSTDESTFDDIVHSFVSLTAIQ
IGLIDLLSCMGLRPDGIVGHSLGEVACGYADGCLSQEEAVLAAYWRGQCIKEAHLPPGAMAAVGLSWEEC
KQRCPPGVVPACHNSKDTVTISGPQAPVFEFVEQLRKEGVFAKEVRTGGMAFHSYFMEAIAPPLLQELKK
VIREPKPRSARWLSTSIPEAQWHSSLARTSSAEYNVNNLVSPVLFQEALWHVPEHAVVLEIAPHALLQAV
LKRGLKPSCTIIPLMKKDHRDNLEFFLAQIGRLHLSGIDANPNALFPPVEFPAPRGTPLISPLIKWDHSL
AWDVPAAEDFPNGSGSPSAAIYNIDTSSESPDHYLVDHTLDGRVLFPATGYLSIVWKTLARALGLGVEQL
PVVFEDVVLHQATILPKTGTVSLEVRLLEASRAFEVSENGNLVVSGKVYQWDDPDPRLFDHPESPTPNPT
EPLFLAQAEVYKELRLRGYDYGPHFQGILEASLEGDSGRLLWKDNWVSFMDTMLQMSILGSAKHGLYLPT
RVTAIHIDPATHRQKLYTLQDKAQVADVVVSRWLRVTVAGGVHISGLHTESAPRRQQEQQVPILEKFCFT
PHTEEGCLSERAALQEELQLCKGLVQALQTKVTQQGLKMVVPGLDGAQIPRDPSQQELPRLLSAACRLQL
NGNLQLELAQVLAQERPKLPEDPLLSGLLDSPALKACLDTAVENMPSLKMKVVEVLAGHGHLYSRIPGLL
SPHPLLQLSYTATDRHPQALEAAQAELQQHDVAQGQWDPADPAPSALGSADLLVCNCAVAALGDPASALS
NMVAALREGGFLLLFTLLRGHPLGDIVAFLTSTEPQYGQGILSQDAWESLFSRVSLRLVGLKKSFYGSTL
PLCRRPTFQDSPIFLPVDDTSFRWVESLKGILADEDSSRPVWLKAINCATSGVVGLVNCLRREPGGNRLR
CVLLSNLSSTSHVPEVDPGSAELQKVLQGDLVMNVYRDGAWGAFRHFLEEDKPEEPTAHAFVSTLTRGD
LSSIRWVCSSLRHAQPTCPGAQLCTVYYASLNFRDIMLATGKLSPDAIPGKWTSQDSLLGMEFSGRDASG
KRVMGLVPAKGLATSVLLSPDFLWDVPSNWTLEEAASVPVVYSTAYYALVVRGRVRPGETLLIHSGSGGV
GQAAIAIALSLGCRVFTTVGSAEKRAYLQARFPQLDSTSFANSRDTSFEQHVLWHTGGKGVDLVLNSLAE
EKLQASVRCLATHGRFLEIGKFDLSQNHPLGMAIFLKNVTFHGVLLDAFFNESSADWREVWALVQAGIRD
GVVRPLKCTVFHGAQVEDAFRYMAQGKHIGKVVVQVLAEEPEAVLKGAKPKLMSAISKTFCPAHKSYIIA
GGLGGFGLELAQWLIQRGVQKLVLTSRSGIRTGYQAKQVRRWRRQGVQVQVSTSNISSLEGARGLIAEAA
QLGPVGGVFNLAVVLRDGLLENQTPEFFQDVCKPKYSGTLNLDRVTREACPELDYFVVFSSVSCGRGNAG
QSNYGFANSAMERICEKRRHEGLPGLAVQWGAIGDVGILVETMSTNDTIVSGTLPQRMASCLEVLDLFLN
QPHMVLSSFVLAEKAAAYRDRDSQRDLVEAVAHILGIRDLAAVNLDSSLADLGLDSLMSVEVRQTLEREL
NLVLSVREVRQLTLRKLQELSSKADEASELACPTPKEDGLAQQQTQLNLRSLLVNPEGPTLMRLNSVQSS
ERPLFLVHPIEGSTTVFHSLASRLSIPTYGLQCTRAAPLDSIHSLAAYYIDCIRQVQPEGPYRVAGYSYG
ACVAFEMCSQLQAQQSPAPTHNSLFLFDGSPTYVLAYTQSYRAKLTPGCEAEAETEAICFFVQQFTDMEH
NRVLEALLPLKGLEERVAAAYDLIIKSHQGLDRQELSFAARSFYYKLRAAEQYTPKAKYHGNVMLLRAKT
GGAYGEDLGADYNLSQVCDGKVSVHVIEGDHRTLLEGSGLESISIIHSSLAEPRVSVREG

SEQ ID 65 hydroxysterold (17-beta) dehydrogenase 2

MSTFFSDTAWICLAVPTVLCGTVFCKYKKSSGQLWSWMVCLAGLCAVCLLILSPFWGLILFSVSCFLMYT
YLSGQELLPVDQKAVLVTGGDCGLGHALCKYLDELGFTVFAGVLNENGPGAEELRRTCSPRLSVLQMDIT
KPVQIKDAYSKVAAMLQDRGLWAVINNAGYLGFPTDGELLLMTDYKQCMAVNFFGTVEVTKTFLPLLRKS
KGRLVNVSSMGGGAPMERLASYGSSKAAVTMFSSVMRLELSKWGIKVASIQPGGFLTNIAGTSDKWEKLE
KDILDHLPAEVQEDYGQDYILAQRNFLLLINSLASKDFSPVLRDIQHAILAKSPFAYYTPGKGAYLWICL
AHYLPIGIYDYFAKRHFGQDKPMPRALRMPNYKKKAT

SEQ ID 66 sulfotransferase, estrogen-preferring

MNSELDYYEKFEEVHQILMYKDFVKYWDNVEAFQARPDDLVIATYPKSGTTWVSEIVYMIYKEGDVEKCK
EDVIFNRIPFLECRKENLMNGVKQLDEMNSPRIVKTHLPPELLPASFWEKDCKIIYLCRNAKDVAVSFYY
FFLMVAGHPNPGSFPEFVEKFMQGQVPYGSWYKHVKSWWEKGKSPRVLFLFYEDLKEDIRKEVIKLIHFL
ERKPSEELVDRIIHHTSFQEMKNNPSTNYTTLPDEIMNQKLSPFMRKGITGDWKNHFTVALNEKFDKHYE
QQMKESTLKFRTEI

SEQ ID 67 ZXD family zinc finger C

MSKLLRHRRKHDDDRRFTCPVEGCGKSFTRAEHLKGHSITHLGTKPFECPVEGCCARFSARSSLYIHSKK
HVQDVGAPKSRCPVSTCNRLFTSKHSMKAHMVRQHSRRQDLLPQLEAPSSLTPSSELSSPGQSELTNMDL

AALFSDTPANASGSAGGSDEALNSGILTIDVTSVSSSLGGNLPANNSSLGPMEPLVLVAHSDIPPSLDSP
LVLGTAATVLQQGSFSVDDVQTVSAGALGCLVALPMKNLSDDPLALTSNSNLAAHITTPTSSSTPRENAS
VPELLAPIKVEPDSPSRPGAVGQQEGSHGLPQSTLPSPAEQHGAQDTELSAGTGNPYLESGGSARTDYRA
IQLAKEKKQRGAGSNAGASQSTQRKIKEGKMSPPHEHASQNSWLCGSLVVPSGGRPGPAPAAGVQCGAQG
VQVQLVQDDPSGEGVLPSARGPATFLPFLTVDLPVYVLQEVLPSSGGPAGREATQFPGSTINLQDLQ

SEQ ID 68 S100 calcium-binding protein A9

MTCKMSQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIMEDLDTNA
DKQLSFEEFIMLMARLTWASHEKMHEGDEGPGHHHKPGLGEGTP

SEQ ID 69 aldehyde dehydrogenase 1A1

MSSSGTPDLPVLLTDLKIQYTKIFINNEWHDSVSGKKFPVFNPATEEELCQVEEGDKEDVDKAVKAARQA
FQIGSPWRTMDASERGRLLYKLADLIERDRLLLATMESMNGGKLYSNAYLNDLAGCIKTLRYCAGWADKI
QGRTIPIDGNFFTYTRHEPIGVCGQIIPWNFPLVMLIWKIGPALSCGNTVVVKPAEQTPLTALHVASLIK
EAGFPPGVVNIVPGYGPTAGAAISSHMDIDKVAFTGSTEVGKLIKEAAGKSNLKRVTLELGGKSPCIVLA
DADLDNAVEFAHHGVFYHQGQCCIAASRIFVEESIYDEFVRRSVERAKKYILGNPLTPGVTQGPQIDKEQ
YDKILDLIESGKKEGAKLECGGGPWGNKGYFVQPTVFSNVTDEMRIAKEEIFGPVQQIMKFKSLDDVIKR
ANNTFYGLSAGVFTKDIDKAITISSALQAGTVWVNCYGVVSAQCPFGGFKMSGNGRELGEYGFHEYTEVK
TVTVKISQKNS

SEQ ID 70 Rho GDP dissociation inhibitor (GDI) alpha

MAEQEPTAEQLAQIAAENEEDEHSVNYKPPAQKSIQEIQELDKDDESLRKYKEALLGRVAVSADPNVPNV
VVTGLTLVCSSAPGPLELDLTGDLESFKKQSFVLKEGVEYRIKISFRVNREIVSGMKYIQHTYRKGVKID
KTDYMVGSYGPRAEEYEFLTPVEEAPKGMLARGSYSIKSRFTDDDKTDHLSWEWNLTIKKDWKD

SEQ ID 71 carbonic anhydrase I

MASPDWGYDDKNGPEQWSKLYPIANGNNQSPVDIKTSETKHDTSLKPISVSYNPATAKEIINVGHSFHVN
FEDNDNRSVLKGGPFSDSYRLFQFHFHWGSTNEHGSEHTVDGVKYSAELHVAHWNSAKYSSLAEAASKAD
GLAVIGVLMKVGEANPKLQKVLDALQAIKTKGKRAPFTNFDPSTLLPSSLDFWTYPGSLTHPPLYESVTW
IICKESISVSSEQLAQFRSLLSNVEGDNAVPMQHNNRPTQPLKGRTVRASF

SEQ ID 72 carbonic anhydrase II

MSHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDTHTAKYDPSLKPLSVSYDQATSLRILNNGHAFNVEF
DDSQDKAVLKGGPLDGTYRLIQFHFHWGSLDGQGSEHTVDKKKYAAELHLVHWNTKYGDFGKAVQQPDGL
AVLGIFLKVGSAKPGLQKVVDVLDSIKTKGKSADFTNFDPRGLLPESLDYWTYPGSLTTPPLLECVTWIV
LKEPISVSSEQVLKFRKLNFNGEGEPEELMVDNWRPAQPLKNRQIKASFK

SEQ ID73 ATP-dependent DNA helicase II, 70 kDa subunit

MSQWESYYKTEGDEEAEEQEENLEASGDYKYSGRDSLIFLVDASKAMFESQSEDELTPFDMSIQCIQSV
YISKIISSDRDLLAVVFYGTEKDKNSVNFKNIYVLQELDNPGAKRILELDQFKGQQPQKRFQDMMGHGSD
YSLSEVLWVCANLFSDVQFKMSHKRIMLFTNEDNPHGNDSAKASRARTKAGDLRDTGIFLDLMHLKKPGG
FDISLFYRDIISIAEDEDLRVHFEESSKLEDLLRKVRAKETRKRALSRLKLKLNKDIVISVGIYNLVQKA
LKPPPIKLYRETNEPVKTKTRTFNTSTGGLLLPSDTKRSQIYGSRQIILEKEETEELKRFDDPGLMLMGF
KPLVLLKKHHYLRPSLFVYPEESLVIGSSTLFSALLIKCLEKEVAALCRYTPRRNIPPYFVALVPQEEEL
DDQKIQVTPPGFQLVFLPFADDKRKMPFTEKIMATPEQVGKMKAIVEKLRFTYRSDSFENPVLQQHFRNL
EALALDLMEPEQAVDLTLPKVEAMNKRLGSLVDEFKELVYPPDYNPEGKVTKRKHDNEGSGSKRPKVEYS
EEELKTHISKGTLGKFTVPMLKEACRAYGLKSGLKKQELLEALTKHFQD

SEQ ID 74 myosin, heavy polypeptide 9, non-muscle

MAQQAADKYLYVDKNFINNPLAQADWAAKKLVWVPSDKSGFEPASLKEEVGEEAIVELVENGKKVKVNKD
DIQKMNPPKFSKVEDMAELTCLNEASVLHNLKERYYSGLIYTYSGLFCVVINPYKNLPIYSEEIVEMYKG
KKRHEMPPHIYAITDTAYRSMMQDREDQSILCTGESGAGKTENTKKVIQYLAYVASSHKSKKDQGELERQ

LLQANPILEAFGNAKTVKNDNSSRFGKFIRINFDVNGYIVGANIETYLLEKSRAIRQAKEERTFHIFYYL
LSGAGEHLKTDLLLEPYNKYRFLSNGHVTIPGQQDKDMFQETMEAMRIMGIPEEEQMGLLRVISGVLQLG
NIVFKKERNTDQASMPDNTAAQKVSHLLGINVTDFTRGILTPRIKVGRDYVQKAQTKEQADFAIEALAKA
TYERMFRWLVLRINKALDKTKRQGASFIGILDIAGFEIFDLNSFEQLCINYTNEKLQQLFNHTMFILEQE
EYQREGIEWNFIDFGLDLQPCIDLIEKPAGPPGILALLDEECWFPKATDKSFVEKVMQEQGTHPKFQKPK
QLKDKADFCIIHYAGKVDYKADEWLMKNMDPLNDNIATLLHQSSDKFVSELWKDVDRIIGLDQVAGMSET
ALPGAFKTRKGMFRTVGQLYKEQLAKLMATLRNTNPNFVRCIIPNHEKKAGKLDPHLVLDQLRCNGVLEG
IRICRQGFPNRVVFQEFRQRYEILTPNSIPKGFMDGKQACVLMIKALELDSNLYRIGQSKVFFRAGVLAH
LEEERDLKITDVIIGFQACCRGYLARKAFAKRQQQLTAMKVLQRNCAAYLKLRNWQWWRLFTKVKPLLQV
SRQEEEMMAKEEELVKVREKQLAAENRLTEMETLQSQLMAEKLQLQEQLQAETELCAEAEELRARLTAKK
QELEEICHDLEARVEEEEERCQHLQAEKKKMQQNIQELEEQLEEEESARQKLQLEKVTTEAKLKKLEEEQ
IILEDQNCKLAKEKKLLEDRIAEFTTNLTEEEEKSKSLAKLKNKHEAMITDLEERLRREEKQRQELEKTR
RKLEGDSTDLSDQIAELQAQIAELKMQLAKKEEELQAALARVEEEAAQKNMALKKIRELESQISELQEDL
ESERASRNKAEKQKRDLGEELEALKTELEDTLDSTAAQQELRSKREQEVNILKKTLEEEAKTHEAQIQEM
RQKHSQAVEELAEQLEQTKRVKANLEKAKQTLENERGELANEVKVLLQGKGDSEHKRKKVEAQLQELQVK
FNEGERVRTELADKVTKLQVELDNVTGLLSQSDSKSSKLTKDFSALESQLQDTQELLQEENRQKLSLSTK
LKQVEDEKNSFREQLEEEEEAKHNLEKQIATLHAQVADMKKKMEDSVGCLETAEEVKRKLQKDLEGLSQR
HEEKVAAYDKLEKTKTRLQQELDDLLVDLDHQRQSACNLEKKQKKFDQLLAEEKTISAKYAEERDRAEAE
AREKETKALSLARALEEAMEQKAELERLNKQFRTEMEDLMSSKDDVGKSVHELEKSKRALEQQVEEMKTQ
LEELEDELQATEDAKLRLEVNLQAMKAQFERDLQGRDEQSEEKKKQLVRQVREMEAELEDERKQRSMAVA
ARKKLEMDLKDLEAHIDSANKNRDEAIKQLRKLQAQMKDCMRELDDTRASREEILAQAKENEKKLKSMEA
EMIQLQEELAAAERAKRQAQQERDELADEIANSSGKGALALEEKRRLEARIAQLEEELEEEQGNTELIND
RLKKANLQIDQINTDLNLERSHAQKNENARQQLERQNKELKVKLQEMEGTVKSKYKASITALEAKIAQLE
EQLDNETKERQAACKQVRRTEKKLKDVLLQVDDERRNAEQYKDQADKASTRLKQLKRQLEEAEEEAQRAN
ASRRKLQRELEDATETADAMNREVSSLKNKLRRGDLPFVVPRRMARKGAGDGSDEEVDGKADGAEAKPAE

SEQ ID 75 similar to Myosin light polypeptide 6 (Myosin light chain alkali 3) (Myosin light chain 3) (MLC-3) (LC17)

MMCDFTEDQTTEFKEAFQLFDRTGDGKILYNQCGDVMRPLGQNPTNTEVVKVLRNPKSNEMNVKLLDFEH
FLPMLQMVAKNKDQGTYEDYVEGLQVFDKEGNGTIMGVEFWHVLVTLGEKITEEEVEVLVAGNEGSNGCI
DYEAFVRHILSG

SEQ ID 76 prosaposin

MYALFLLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLPCDICKDVV
TAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIIKGEMSRPGEVCSALNLCES
LQKHLAELNHQKQLESNKIPELDMTEVVAPFMANIPLLLYPQDGPRSKPQPKDNGDVCQDCIQMVTDIQT
AVRTNSTFVQALVEHVKEECDRLGPGMADICKNYISQYSEIAIQMMMHMQPKEICALVGFCDEVKEMPMQ
TLVPAKVASKNVIPALELVEPIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCSKLP
KSLSEECQEVVDTYGSSILSILLEEVSPELVCSMLHLCSGTRLPALTVHVTQPKDGGFCEVCKKLVGYLD
RNLEKNSTKQEILAALEKGCSFLPDPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAHKPLL
GTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN

SEQ ID 77 myosin regulatory light chain MRCL3

MSSKRTKTKTKKRPQRATSNVFAMFDQSQIQEFKEAFNMIDQNRDGFIDKEDLHDMLASLGKNPTDEYLD
AMMNEAPGPINFTMFLTMFGEKLNGTDPEDVIRNAFACFDEEATGTIQEDYLRELLTTMGDRFTDEEVDE
LYREAPIDKKGNFNYIEFTRILKHGAKDKDD

SEQ ID 78 solute carrier family 4, anion exchanger, member 1 (erythrocyte membrane protein band 3, Diego blood group)

MEELQDDYEDMMEENLEQEEYEDPDIPESQMEEPAAHDTEATATDYHTTSHPGTHKVYVELQELVMDEKN
QELRWMEAARWVQLEENLGENGAWGRPHLSHLTFWSLLELRRVFTKGTVLLDLQETSLAGVANQLLDRFI
FEDQIRPQDREELLRALLLKHSHAGELEALQGVKPAVLTRSGDPSQPLLPQHSSLETQLFCEQGDGGTEG
HSPSGILEKIPPDSEATLVLVGRADFLEQPVLGFVRLQEAAELEAVELPVPIRFLFVLLGPEAPHIDYTQ

LGRAAATLMSERVFRIDAYMAQSRGELLHSLEGFLDCSLVLFPTDAPSEQALLSLVPVQRELLRRRYQSS
PAKPDSSFYKGLDLNGGPDDPLQQTGQLFGGLVRDIRRRYPYYLSDITDAFSPQVLAAVIFIYFAALSPA
ITFGGLLGEKTRNQMGVSELLISTAVQGILFALLGAQPLLVVGFSGPLLVFEEAFFSFCETNGLEYIVGR
VWIGFWLILLVVLVVAFEGSFLVRFISRYTQEIFSFLISLIFIYETFSKLIKIFQDHPLQKTYNYNVLMV
PKPQGPLPNTALLSLVLMAGTFFFAMMLRKFKNSSYFPGKLRRVIGDFGVPISILIMVLVDFFIQDTYTQ
KLSVPDGFKVSNSSARGWVIHPLGLRSEFPIWMMFASALPALLVFILIFLESQITTLIVSKPERKMVKGS
GFHLDLLLVVGMGGVAALFGMPWLSATTVRSVTHANALTVMGKASTPGAAAQIQEVKEQRISGLLVAVLV
GLSILMEPILSRIPLAVLFGIFLYMGVTSLSGIQLFDRILLLFKPPKYHPDVPYVKRVKTWRMHLFTGIQ
IICLAVLWVVKSTPASLALPFVLILTVPLRRVLLPLIFRNVELQCLDADDAKATEDEEEGRDEYDEVAMP
V

SEQ ID 79 transgelin 2

MANRGPAYGLSREVQQKIEKQYDADLEQILIQWITTQCRKDVGRPQPGRENFQNWLKDGTVLCELINALY
PEGQAPVKKIQASTMAFKQMEQISQFLQAAERYGINTTDIFQTVDLWEGKNMACVQRTLMNLGGLAVARD
DGLPSGDPNWFPKKSKENPRNFSDNQLQEGKNVIGLQMGTNRGASQAGMTGYGMPRQIL

SEQ ID 80 tropomyosin 3

MMEAIKKKMQMLKLDKENALDRAEQAEAEQKQAEERSKQLEDELAAMQKKLKGTEDELDRAQERLATALQ
KLEEAEKAADESERGMKVIENRALKDEEKMELQEIQLKEAKHIAEEADRKYEEVARKLVIIEGDLERTEE
RABLAESKCSELEEELKNVTNNLKSLEAQAEKYSQKEDKYEEEIKILTDKLKEAETRAEFAERSVAKLEK
TIDDLEDELYAQKLKYKAISEELDHALNDMTSI

SEQ ID 81 talin 1

MVALSLKISIGNVVKTMQFEPSTMVYDACRIIRERIPEAPAGPPSDFGLFLSDDDPKKGIWLEAGKALDY
YMLRNGDTMEYRKKQRPLKIRMLDGTVKTIMVDDSKTVTDMLMTICARIGITNHDEYSLVRELMEEKKEE
ITGTLRKDKTLLRDEKKMEKLKQKLHTDDELNWLDHGRTLREQGVEEHETLLLRRKFFYSDQNVDSRDPV
QLNLLYVQARDDILNGSHPVSFDKACEFAGFQCQIQPGPHNEQKHKAGFLDLKDFLPKEYVKQKGERKIP
QAHKNCGQMSEIEAKVRYVKLARSLKTYGVSFFLVKEKMKGKNKLVPRLLGITKECVMRVDEKTKEVIQE
WNLTNIKRWAASPKSFILDFGDYQDGYYSVQITEGEQIAQLIAGYIDIILKKKKSKDHFGLEGDEESTML
EDSVSPKKSTVLQQQYNRVGKVEHGSVAEPAIMRSGASGPENFQVGSMPPAQQQITSGQMHRGHMPPLTS
AQQALTGTINSSMQAVQAAQATLDDFDTLPPLGQDAASKAWRKNKMDESKHEIHSQVDAITAGTASVVNL
TAGDPAETDYTAVGCAVTTISSNLTEMSRGVKLLAALLEDEGGSGRPLLQAAKGLAGAVSELLRSAQPAS
AEPRQNLLQAAGNVGQASGELLQQIGESDTDPHFQDALMQLAKAVASAAAALVLKAKSVAQRTEDSGLQT
QVIAAATQCALSTSQLVACTKVVAPTISSPVCQEQLVEAGRLVAKAVEGCVSASQAAATEDGQELRGVGAA
ATAVTQALNELLQHVKAHATGAGPAGRYDQATDTILTVTENIFSSMGDAGEMVGQARILAQATSDLVNAI
KADAEGESDLENSRKLLSAAKILADATAKMVEAAKGAAAHPDSEEQQQRLREAAEGLRMATNAAAQNAIK
KKLVQRLEHAAKQAAASATQTIAAAQHAASTPKASAGRQPLLVQSCKAVAEQIPLLVQQVRGSQAQPDSP
SAQLALIAASQSFLQPGGKMVAAAKASVPTIQDQASAMQLSQCAKNLGTALAELRTAAQKAQEACGPLEM
DSALSVVQNLEKDLQEVKAAARDGKLKPLPGETMEKCTQDLGNSTKAVSSAIAQLLGEVAQGNENYAGIA
ARDVAGGLRSLAQAARGVAALTSDPAVQAIVLDTASDVLDKASSLIEEAKKAAGHPGDPESQQRLAQVAK
AVTQALNRCVSCLPGQRDVDNALRAVGDASKRLLSDSLPPSTGTFQEAQSRLNEAAAGLNQAATELVQAS
RGTPQDLARASGRFGQDFSTFLEAGVEMAGQAPSQEDRAQVVSNLKGISMSSSKLLLAAKALSTDPAAPN
LKSQLAAAARAVTDSINQLITMCTQQAPGQKECDPALRELETVRELLENPVQPINDMSYFGCLDSVMENS
KVLGEAMTGISQNAKNGNLPEFGDAISTASKALCQFTEAAAQAAYLVGVSDPNSQAGQQGLVEPTQFARA
NQAIQMACQSLGEPGCTQAQVLSAATIVAKHTSALCNSCRLASARTTNPTAKRQFVQSAKEVANSTANLV
KTIKALDGAPTEENRAQCRAATAPLLEAVDNLSAFASNPEFSSIPAQISPEGRAAMEPIVISAKTMLESA
GGLIQTARALAVNPRDPPSWSVLAGHSRTVSDSIKKLITSMRDKAPGQLECETAIAAENSCLRDLDQASL
AAVSQQLAPREGISQEALHTQMLTAVQEISHLIEPLANAARAEASQLGHKVSQMAQYFEPLTLAAVGAAS
KTLSHPQQMALLDQTKTLAESALQLLYTAKEAGGNPKQAAHTQEALEEAVQMMTEAVEDLTITLNEAASA
AGVVGGMVDSITQAINQLDEGPMGEPEGSFVDYQTTMVRTAKAIAVTVQEMVTKSNTSPEELGPLANQLT
SDYGRLASEAKPAAVAAENEEIGSHIKHRVQELGHGCAALVTKAGALQCSPSDAYTKKELIECARRVSEK
VSHVLAALQAGNRGTQACITAASAVSGIIADLDTTIMFATAGTLNREGTETFADHREGILKTAKVLVEDT
KVLVQNAAGSQEKLAQAAQSSVATITRLADVVKLGAASLGAEDPETQVVLINAVKDVAKALGDLISATKA
AAGKVGDDPAVWQLKNSAKVMVTNVTSLLKTVKAVEDEATKGTRALEATTEHIRQELAVFCSPEPPAKTS

TPEDFIRMTKGITMATAKAVAAGNSCRQEDVIATANLSRRAIADMLRACKEAAYHPEVAPDVRLRALHYG
RECANGYLELLDHVLLTLQKPSPELKQQLTQHSKRVAGSVTELIQAAEAMKGTEWVDPEDPTVIAENELL
QAAAAIEAAAKKLEQLKPRAKPKEADESLNFEEQILEAAKSIAAATSALVKAASAAQRELVAQGKVGAIP
ANALDDGQWSQGLISAARMVAAATNNLCEAANAAVQGHASQEKLISSAKQVAASTAQLLVACKVKADQDS
EAMKRLQAAGNAVKRASDNLVKAAQKAAAFEEQENETVVVKEKMVGGIAQIIAAQEEMLRKERELEEARK
KLAQIRQQQYKFLPSELRDEH

SEQ ID 82 enoyl Coenzyme A hydratase domain containing 2

MLRVLCLLRPWRPLRARGCASDGAAGGSEIQVRALAGPDQGITEILMNRPSARNALGNVFVSELLETLAQ
LREDRQVRVLLFRSGVKGAFCAGADLKEREQMSEAEVGVFVQRLRGLMNDIASSAVMGLIETTRGLLPGA
GGTQRLPRCLGVALAKELIFTGRRLSGTEAHVLGLVNHAVAQNEEGDAAYQRARALAQEILPQAPIAVRL
GKVAINRGTEVDIASGMAIEGMCYAQNIPTRDRLEGMAAFREKRTPKFVGK

SEQ ID 83 alpha-aminoadipate aminotransferase

MNYARFITAASAARNPSPIRTMTDILSRGPKSMISLAGGLPNPNMFPPKTAVITVENGKTIQFGEEMMKR
ALQYSPSAGIPELLSWLKQLQIKLHNPPTIHYPPSQGQMDLCVTSGSQQGLCKVFEMIINPGDNVLLDEP
AYSGTLQSLHPLGCNIINVASDESGIVPDSLRDILSRWKPEDAKNPQKNTPKFLYTVPNGNNPTGNSLTS
ERKKEIYELARKYDFLIIEDDPYYFLQFNKFRVPTFLSMDVDGRVIRADSFSKIISSGLRIGFLTGPKPL
IERVILHIQVSTLHPSTFNQLMISQLLHEWGEEGFMAHVDRVIDFYSNQKDAILAAADKWLTGLAEWHVP
AAGMFLWIKVKGINDVKELIEEKAVKMGVLMLPGNAFYVDSSAPSPYLRASFSSASPEQMDVAFQVLAQL
IKESL

SEQ ID 84 acyl-CoA synthetase long-chain family member 1

MQAHELFRYFRMPELVDFRQYVRTLPTNTLMGFGAFAALTTFWYATRPKPLKPPCDLSMQSVEVAGSGGA
RRSALLDSDEPLVYFYDDVTTLYEGPQRGIQVSNNGPCLGSRKPDQPYEWLSYKQVAELSECIGSALIQK
GFKTAPDQFIGIFAQNRPEWVIIEQGCFAYSMVIVPLYDTLGNEAITYIVNKAELSLVFVDKPEKAKLLL
EGVENKLIPGLKIIVVMDAYGSELVERGQRCGVEVTSMKAMEDLGRANRRKPKPPAPEDLAVICFTSGTT
GNPKGAMVTHRNIVSDCSAFVKATENTVNPCPDDTLISFLPLAHMFERVVECVMLCHGAKIGFFQGDIRL
LMDDLKVLQPTVFPVVPRLLNRMFDRIFGQANTTLKRWLLDPASKRKEAELRSGIIRNNSLWDRLIFHKV
QSSLGGRVRLMVTGAAPVSATVLTFLRAALGCQFYEGYGQTECTAGCCLTMPGDWTAGHVGAPMPCNLIK
LVDVEEMNYMAAEGEGEVCVKGPNVFQGYLKDPAKTAEALDKDGWLHTGDIGKWLPNGTLKIIDRKKHIF
KLAQGEYIAPEKIENIYMRSEPVAQVFVHGESLQAFLIAIVVPDVETLCSWAQKRGFEGSFEELCRNKDV
KKAILEDMVRLGKDSGLKPFEQVKGITLHPELFSIDNGLLTPTMKAKRPELRNYFRSQIDDLYSTIKV

SEQ ID 85 aldehyde dehydrogenase 1 family, member L1

MKIAVIGQSLFGQEVYCHLRKEGHEVVGVFTVPDKDGKADPLGLEAEKDGVPVFKYSRWRAKGQALPDVV
AKYQALGAELNVLPFCSQFIPMEIISAPRHGSIIYHPSLLPRHRGASAINWTLIHGDKKGGFSIFWADDG
LDTGDLLLQKECEVLPDDTVSTLYNRFLFPEGIKGMVQAVRLIAEGKAPRLPQPEEGATYEGIQKKETAK
INWDQPAEAIHNWIRGNDKVPGAWTEACEQKLTFFNSTLNTSGLVPEGDALPIPGAHRPGVVTKAGLILF
GNDDKMLLVKNIQLEDGKMILASNFFKGAASSVLELTEAELVTAEAVRSVWQRILPKVLEVEDSTDFFKS
GAASVDVVRLVEEVKELCDGLELENEDVYMASTFGDFIQLLVRKLRGDDEEGECSIDYVEMAVNKRTVRM
PHQLFIGGEFVDAEGAKTSETINPTDGSVICQVSLAQVTDVDKAVAAAKDAFENGRWGKISARDRGRLMY
RLADLMEQHQEELATIEALDAGAVYTLALKTHVGMSIQTFRYFAGWCDKIQGSTIPINQARPNRNLTLTR
KEPVGVCGIIPWNYPLMMLSWKTAACLAAGNTVVIKPAQVTPLTALKFAELTLKAGIPKGVVNVLPGSG
SLVGQRLSDHPDVRKIGFTGSTEVGKHIMKSCAISNVKKVSLELGGKSPLIFADCDLNKAVQMGMSSVF
FNKGENCIAAGRLFVEDSIHDEFVRRVVEEVRKMKVGNPLDRDTDHGPQNHHAHLVKLMEYCQHGVKEGA
TLVCGGNQVPRPGFFFEPTVFTDVEDHMFIAKEESFGPVMIISRFADGDLDAVLSRANATEFGLASGVFT
RDINKALYVSDKLQAGTVFVNTYNKTDVAAPFGGFKQSGFGKDLGEAALNEYLRVKTVTFEY

SEQ ID 86 argininosuccinate synthetase

MSSKGSVVLAYSGGLDTSCILVWLKEQGYDVIAYLANIGQKEDFEEARKKALKLGAKKVFIEDVSREFVE
EFIWPAIQSSALYEDRYLLGTSLARPCIARKQVEIAQREGAKYVSHGATGKGNDQVRFELSCYSLAPQIK
VIAPWRMPEFYNRFKGRNDLMEYAKQHGIPIPVTPKNPWSMDENLMHISYEAGILENPKNQAPPGLYTKT

QDPAKAPNTPDILEIEPKKGVPVKVTNVKDGTTHQTSLELFMYLNEVAQKHGVGRIDIVENRFIGMKSRG
IYETPAGTILYHAHLDIEAFTMDREVRKIKQGLGLKFAELVYTGFWHSPECEPVRHCIAKSQERVEGKVQ
VSVLKGQVYILGRESPLSLYNEELVSMNVQGDYEPTDATGFINISLRLKEYHRLQSKVTAK

SEQ ID 87 dimethylglycine dehydrogenase

MLRPGAQLLRGLLLRSCPLQGSPGRPRSVCGREGEEKPPLSAETQWKDRAETVIIGGGCVGVSLAYHLAK
AGMKDVVLLEKSELTAGSTWHAAGLTTYFHPGINLKKIHYDSIKLYEKLEEETGQVVGFHQPGSIRLATT
PVRVDEFKYQMTRTGWHATEQYLIEPEKIQEMFPLLNMNKVLAGLYNPGDGHIDPYSLTMALAAGARKCG
ALLKYPAPVTSLKARSDGTWDVETPQGSMRANRIVNAAGFWAREVGKMIGLEHPLIPVQHQYVVTSTISE
VKALKRELPVLRDLEQSYYLRQERDGLLFGPYESQEKMKVQDSWVTNGVPPGFGKELFESDLDRIMEHIK
AAMEMVPVLKKADIINVVNGPITYSPDILPMVGPHQGVRNYWVAIGFGYGIIHAGGVGKYLSDWILHGEP
PFDLIELDPNRYGKWTTTQYTEAKARESYGFNNIVGYPKEERFAGRPTQRVSGLYQRLESKCSMGFHAGW
EQPHWFYKPGQDTQYRPSFRRTNWFEPVGSEYKQVMQRVAVTDLSPFGKFNIKGQDSIRLLDHLFANVIP
KVGFTNISHMLTPKGRVYAELTVSHQSPGEFLLITGSGSELHDLRWIEEEAVKGGYDVEIKNITDELGVL
GVAGPQARKVLQKLTSEDLSDDVFKFLQTKSLKVSNIPVTAIRISYTGELGWELYHRREDSVALYDAIMN
AGQEEGIDNFGTYAMNALRLEKAFRAWGLEMNCDTNPLEAGLEYFVKLNKPADFIGKQALKQIKAKGLKR
RLVCLTLATDDVDPEGNESIWYNGKVVGNTTSGSYSYSIQKSLAFAYVPVQLSEVGQQVEVELLGKNYPA
VIIQEPLVLTEPTRNRLQKKGGKDKT

SEQ ID 88 dihydropyrimidinase

MAAPSRLLIRGGRVVNDDFSEVADVLVEDGVVRALGHDLLPPGGAPAGLRVLDAAGKLVLPGGIDTHTHM
QFPFMGSRSIDDFHQGTKAALSGGTTMIIDFAIPQKGGSLIEAFETWRSWADPKVCCDYSLHVAVTWWSD
QVKEEMKILVQDKGVNSFKMFMAYKDLYMVTDLELYEAFSRCKEIGAIAQVHAENGDLIAEGAKKMLALG
ITGPEGHELCRPEAVEAEATLRAITIASAVNCPLYIVHVMSKSAAKVIADARRDGKVVYGEPIAASLGTD
GTHYWNKEWHHAAHHVMGPPLRPDPSTPDFLMNLLANDDLTTTGTDNCTFNTCQKALGKDDFTKIPNGVN
GVEDRMSVIWEKGVHSGKMDENRFVAVTSTNAAKIFNLYPRKGRIAVGSDADIVIWDPKGTRTISAKTHH
QAVNFNIFEGMVCHGVPLVTISRGKVVYEAGVFSVTAGDGKFIPRKPFAEYIYKRIKQRDRTCTPTPVER
APYKGEVATLKSRVTKEDATAGTRKQAHP

SEQ ID 89 ectonucleotide pyrophosphatase/phosphodiesterase 2 (autotaxin)

MARRSSFQSCQIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNISGSCKGRCFELQEAG
PPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARGDCCTNYQVVCKG
ESHWVDDDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPNIEKLRSCGTHSPYMRPVYPTKT
FPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGREKFNHRWWGGQPLWITATKQGVKAGTFFWSVVI
PHERRILTILRWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSPFTPAKRFKRKVAPKRRQER
PVAPPKKRRRKIHRMDHYAAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCD
RTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYDPKAIIANLTCKKPDQHFKPYLKQHLPKRLHYANN
RRIEDIHLLVERRWHVARKPLDVYKKPSGKCFFQGDHGFDNKVNSMQTVFVGYGPTFKYKTKVPPFENIE
LYNVMCDLLGLKPAPNNGTHGSLNHLLRTNTFRPTMPEEVTRPNYPGIMYLQSDFDLGCTCDDKVEPKNK
LDELNKRLHTKGSTEERHLLYGRPAVLYRTRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDH
LTSCVRPDVRVSPSFSQNCLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQR
VLVKKYASERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGP
LSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHT
YESEI

SEQ ID 90 fatty acid amide hydrolase

MVQYELWAALPGASGVALACCFVAAAVALRWSGRRTARGAVVRARQKQRAGLENMDRAAQRFRLQNPDLD
SEALLALPLPQLVQKLHSRELAPEAVLFTYVGKAWEVNKGTNCVTSYLADCETQLSQAPRQGLLYGVPVS
LKECFTYKGQDSTLGLSLNEGVPAECDSVVVHVLKLQGAVPFVHTNVPQSMFSYDCSNPLFGQTVNPWKS
SKSPGGSSGGEGALIGSGGSPLGLGTDIGGSIRFPSSFCGICGLKPTGNRLSKSGLKGCVYGQEAVRLSV
GPMARDVESLALCLRALLCEDMFRLDPTVPPLPFREEVYTSSQPLRVGYYETDNYTMPSPAMRRAVLETK
QSLEAAGHTLVPFLPSNIPHALETLSTGGLFSDGGHTFLQNFKGDFVDPCLGDLVSILKLPQWLKGLLAF
LVKPLLPRLSAFLSNMKSRSAGKLWELQHEIEVYRKTVIAQWRALDLDVVLTPMLAPALDLNAPGRATGA

VSYTMLYNCLDFPAGVVPVTTVTAEDEAQMEHYRGYFGDIWDKMLQKGMKKSVGLPVAVQCVALPWQEEL
CLRFMREVERLMTPEKQSS

SEQ ID 91 guanidinoacetate N-methyltransferase

MSAPSATPIFAPGENCSPAWGAAPAAYDAADTHLRILGKPVMERWETPYMHALAAAASSKGGRVLEVGFG
MAIAASKVQEAPIDEHWIIECNDGVFQRLRDWAPRQTHKVIPLKGLWEDVAPTLPDGHFDGILYDTYPLS
EETWHTHQFNFIKNHAFRLLKPGGVLTYCNLTSWGELMKSKYSDITMFEETQVPALLEAGFRRENIRTE
VMALVPPADCRYYAFPQMITPLVTKG

SEQ ID 92 3-hydroxyisobutyrate dehydrogenase

MAASLRLLGAASGLRYWSRRLRPAAGSFAAVCSRSVASKTPVGFIGLGNMGNPMAKNLMKHGYPLIIYDV
FPDACKEFQDAGEQVVSSPADVAEKADRIITMLPTSINAIEAYSGANGILKKVKKGSLLIDSSTIDPAVS
KELAKEVEKMGAVFMDAPVSGGVGAARSGNLTFMVGGVEDEFAAAQELLGCMGSNVVYCGAVGTGQAAKI
CNNMLLAISMIGTAEAMNLGIRLGLDPKLLAKILNMSSGRCWSSDTYNPVPGVMDGVPSANNYQGGFGTT
LMAKDLGLAQDSATSTKSPILLGSLAHQIYRMMCAKGYSKKDFSSVFQFLREEETF

SEQ ID 93 ketohexokinase

MEEKQILCVGLVVLDVISLVDKYPKEDSEIRCLSQRWQRGGNASNSCTVLSLLGAPCAFMGSMAPGHVAD
FVLDDLRRYSVDLRYTVFQTTGSVPIATVIINEASGSRTILYYDRSLPDVSATDFEKVDLTQFKWIHIEG
RNASEQVKMLQRIDAHNTRQPPEQKIRVSVEVEKPREELFQLFGYGDVVFVSKDVAKHLGFQSAEEALRG
LYGRVRKGAVLVCAWAEEGADALGPDGKLLHSDAFPPPRVVDTLGAGDTFNASVIFSLSQGRSVQEALRF
GCQVAGKKCGLQGFDGIV

SEQ ID 94 kynureninase (L-kynurenine hydrolase)

MEFSSLELPADTVQRIAAELKCHPTDERVALHLDEEDKLRHFRECFYIPKIQDLPPVDLSLVNKDENAIY
FLGNSLGLQPKMVKTYLEEELDKWAKIAAYGHEVGKRPWITGDESIVGLMKDIVGANEKEIALMNALTVN
LHLLMLSFFKPTPKRYKILLEAKAFPSDHYAIESQLQLHGLNIEESMRMIKPREGEETLRIEDILEVIEK
EGDSIAVILFSGVHFYTGQHFNIPAITKAGQAKGCYVGFDLAHAVGNVELYLHDWGVDFACWCSYKYLNA
GAGGIAGAFIHEKHAHTIKPARSEFFN

SEQ ID 95 pyruvate carboxylase

MLKFRTVHGGLRLLGIRRTSTAPAASPNVRRLEYKPIKKVMVANRGEIAIRVFRACTELGIRTVAIYSEQ
DTGQMHRQKADEAYLIGRQLAPVQAYLHIPDIIKVAKENNVDAVHPGYGFLSERADFAQACQDAGVRFIG
PSPEVVRKMGDKVEARAIAIAAGVPVVPGTDAPITSLHEAHEFSNTYGFPIIFKAAYGGGGRGMRVVHSY
EELEENYTRAYSEALAAFGNGALFVEKFIEKPRHIEVQILGDQYGNILHLYERDCSIQRRHQKVVEIAPA
AHLDPQLRTRLTSDSVKLAKQVGYENAGTVEFLVDRHGKHYFIEVNSRLQVEHTVTEEITDVDLVHAQIH
VSEGRSLPDLGLRQENIRINGCAIQCRVTTEDPARSFQPDTGRIEVFRSGEGMGIRLDNASAFQGAVISP
HYDSLLVKVIAHGKDHPTAATKMSRALAEFRVRGVKTNIAFLQNVLNNQQFLAGTVDTQFIDENPELFQL
RPAQNRAQKLLHYLGHVMVNGPTTPIPVKASPSPTDPVVPAVPIGPPPAGFRDILLREGPEGFARAVRNH
PGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFSKLFSMENWGGATFDVAMRFLYECPWRRLQEL
RELIPNIPFQMLLRGANAVGYTNYPDNVVFKFCEVAKENGMDVFRVFDSLNYLPNMLLGMEAAGSAGGVV
EAAISYTGDVADPSRTKYSLQYYMGLAEELVRAGTHILCIKDMAGLLKPTACTMLVSSLRDRFPDLPLHI
HTHDTSGAGVAAMLACAQAGADVVDVAADSMSGMTSQPSMGALVACTRGTPLDTEVPMERVFDYSEYWEG
ARGLYAAFDCTATMKSGNSDVYENEIPGGQYTNLHFQAHSMGLGSKFKEVKKAYVEANQMLGDLIKVTPS
SKIVGDLAQFMVQNGLSRAEAEAQAEELSFPRSVVEFLQGYIGVPHGGFPEPFRSKVLKDLPRVEGRPGA
SLPPLDLQALEKELVDRHGEEVTPEDVLSAAMYPDVFAHFKDFTATFGPLDSLNTRLFLQGPKIAEEFEV
ELERGKTLHIKALAVSDLNRAGQRQVFFELNGQLRSILVKDTQAMKEMHFHPKALKDVKGQIGAPMPGKV
IDIKVVAGAKVAKGQPLCVLSAMKMETVVTSPMEGTVRKVHVTKDMTLEGDDLILEIE

SEQ ID 96 pyruvate dehydrogenase kinase, isoenzyme 2

MRWVWALLKNASLAGAPKYIEHFSKFSPSPLSMKQFLDFGSSNACEKTSFTFLRQELPVRLANIMKEINL

LPDRVLSTPSVQLVQSWYVQSLLDIMEFLDKDPEDHRTLSQFTDALVTIRNRHNDVVPTMAQGVLEYKDT
YGDDPVSNQNIQYFLDRFYLSRISIRMLINQHTLIFDGSTNPAHPKHIGSIDPNCNVSEVVKDAYDMAKL
LCDKYYMASPDLEIQEINAANSKQPIHMVYVPSHLYHMLFELFKNAMRATVESHESSLILPPIKVMVALG
EEDLSIKMSDRGGGVPLRKIERLFSYMYSTAPTPQPGTGGTPLAGFGYGLPISRLYAKYFQGDLQLFSME
GFGTDAVIYLKALSTDSVERLPVYNKSAWRHYQTIQEAGDWCVPSTEPKNTSTYRVS

SEQ ID 97 solute carrier family 10 (sodium/bile acid cotransporter family), member 1

MEAHNASAPFNFTLPPNFGKRPTDLALSVILVFMLFFIMLSLGCTMEFSKIKAHLWKPKGLAIALVAQYG
IMPLTAFVLGKVFRLKNIEALAILVCGCSPGGNLSNVFSLAMKGDMNLSIVMTTCSTFCALGMMPLLLYI
YSRGIYDGDLKDKVPYKGIVISLVLVLIPCTIGIVLKSKRPQYMRYVIKGGMIILLCSVAVTVLSAINV
GKSIMFAMTPLLIATSSLMPFIGFLLGYVLSALFCLNGRCRRTVSMETGCQNVQLCSTILNVAFPPEVIG
PLFFFPLLYMIFQLGEGLLLIAIFWCYEKFKTPKDKTKMIYTAATTEEIPGALGNGTYKGEDCSPCTA

SEQ ID 98 solute carrier family 37 (glycerol-6-phosphate transporter), member 4

MAAQGYGYYRTVIFSAMFGGYSLYYFNRKTFSFVMPSLVEEIPLDKDDLGFITSSQSAAYAISKFVSGVL
SDQMSARWLFSSGLLLVGLVNIFFAWSSTVPVFAALWFLNGLAQGLGWPPCGKVLRKWFEPSQFGTWWAI
LSTSMNLAGGLGPILATILAQSYSWRSTLALSGALCVVYSFLCLLLIHNEPADVGLRNLDPMPSEGKKGS
LKEESTLQELLLSPYLWVLSTGYLVVFGVKTCCTDWGQFFLIQEKGQSALVGSSYMSALEVGGLVGSIAA
GYLSDRAMAKAGLSNYGNPRHGLLLFMMAGMTVSMYLFRVTVTSDSPKLWILVLGAVFGFSSYGPIALFG
VIANESAPPNLCGTSHAIVGLMANVGGFLAGLPPSTIAKHYSWSTAFWVAEVICAASTAAFFLLRNIRTK
MGRVSKKAE

SEQ ID 99 sulfotransferase family, cytosolic, 1C, member 1

MALTSDLGKQIKLKEVEGTLLQPATVDNWSQIQSFEAKPDDLLICTYPKAGTTWIQEIVDMIEQNGDVEK
CQRAIIQHRHPFIEWARPPQPSGVEKAKAMPSPRILKTHLSTQLLPFSFWENNCKFLYVARNAKDCMVSY
YHFQRMNHMLPDPGTWEEYFETFINGKVVWGSWFDHVKGWWEMKDRHQILFLFYEDIKRDPKHEIRKVMQ
FMGKKVDETVLDKIVQETSFEKMKENPMTNRSTVSKSILDQSISSFMRKGTVGDWKNHFTVAQNERFDEI
YRRKMEGTSDNFCMEL

SEQ ID100 arachidonate 5-lipoxygenase-activating protein

MDQETVGNVVLLAIVTLISVVQNGFFAHKVEHESRTQNGRSFQRTGTLAFERVYTANQNCVDAYPTFLAV
LWSAGLLCSQVPAAFAGLMYLFVRQKYFVGYLGERTQSTPGYIFGKRIILFLFLMSVAGIFNYYLIFPFG
SDFENYIKTISTTISPLLLIP

SEQ ID 101 adenine phosphoribosyltransferase

MADSELQLVEQRIRSFPDFPTPGVVFRDISPVLKDPASFRAAIGLLARHLKATHQGRIDYIAGLDSRGFL
FGPSLAQELGLGCVLIRKRGKLPGPTLWASYSLEYGKAELEIQKDALEPGQRVVVVDDLLATGGTMNAAC
ELLGRLQAEVLECVSLVELTSLKGREKLAPVPFFSLLQYE

SEQ ID 102 brain abundant, membrane attached signal protein 1

MGGKLSKKKKGYNVNDEKAKEKDKKAEGAATEEEGTPKESEPQAAAEPAEAKEGKEKPDQDAEGKAEEKE
GEKDAAAAKEEAPKAEPEKTEGAAEAKAEPPKAPEQEQAAPGPAAGGEAPKAAEAAAAPAESAAPAAGEE
PSKEEGEPKKTEAPAAPAAQETKSDGAPASDSKPGSSEAAPSSKETPAATEAPSSTPKAQGPAASAEEPK
PVEAPAANSDQTVTVKE

SEQ ID 103 peripheral benzodiazapine receptor

MAPPWVPAMGFTLAPSLGCFVGSRFVHGEGLRWYAGLQKPSWHPPHWVLGPVWGTLYSAMGYGSYLVWKE
LGGFTEKAVVPLGLYTGQLALNWAWPPIFFGARQMGWALVDLLLVSGAAAATTVAWYQVSPLAARLLYPY
LAWLAFITTLNYCVWRDNHGWRGGRRLPE

SEQ ID 104 carbonyl reductase 1

MSSGIHVALVTGGNKGIGLAIVRDLCRLFSGDVVLTARDVTRGQAAVQQLQAEGLSPRFHQLDIDDLQSI
RALRDFLRKEYGGLDVLVNNAGIAFKVADPTPFHIQAEVTMKTNFFGTRDVCTELLPLIKPQGRVVNVSS
IMSVRALKSCSPELQQKFRSETITEEELVGLMNKFVEDTKKGVHQKEGWPSSAYGVTKIGVTVLSRIHAR
KLSEQRKGDKILLNACCPGWVRTDMAGPKATKSPEEGAETPVYLALLPPDAEGPHGQFVSEKRVEQW

SEQ ID 105 CD47 antigen

MWPLVAALLLGSACCGSAQLLFNKTKSVEFTFCNDTVVIPCFVTNMEAQNTTEVYVKWKFKGRDIYTFDG
ALNKSTVPTDFSSAKIEVSQLLKGDASLKMDKSDAVSHTGNYTCEVTELTREGETIIELKYRVVSWFSPN
ENILIVIFPIFAILLFWGQFGIKTLKYRSGGMDEKTIALLVAGLVITVIVIVGAILFVPGEYSLKNATGL
GLIVTSTGILILLHYYVFSTAIGLTSFVIAILVIQVIAYILAVVGLSLCIAACIPMHGPLLISGLSILAL
AQLLGLVYMKFVASNQKTIQPPRKAVEEPLNE

SEQ ID 106 embigin homolog

MRALPGLLEARARTPRLLLLQCLLAAARPSSADGSAPDSPFTSPPLREEIMANNESLESHNISLTEHSSM
PVEKNITLERPSNVNLTCQFTTSGDLNAVNVTWKKDGEQLENNYLVSATGSTLYTQYRFTIINSKQMGSY
SCFFREEKEQRGTFNFKVPELHGKNKPLISYVGDSTVLTCKCQNCFPLNWTWYSSNGSVKVPVGVQMNKY
VINGTYANETKLKITQLLEEDGESYWCRALFQLGESEEHIELVVLSYLVPLKPFLVIVAEVILLVATILL
CEKYTQKKKKHSDEGKEFEQIEQLKSDDSNGIENNVPRHRKNESLGQ

SEQ ID 107 protein phosphatase 1, regulatory subunit 15A

MAPGQAPHQATPWRDAHPFFLLSPVMGLLSRAWSRLRGLGPLEPWLVEAVKGAALVEAGLEGEARTPLAI
PHTPWGRRPGEEAEDSGGPGEDRETLGLKTSSSLPEAWGLLDDDDGMYGEREATSVPRGQGSQFADGQRA
PLSPSLLIRTLQQSDKNPGEEKAEEEGVAEEEGVNKFSYPPSHRECCPAVEEEDDEEAVKKEAHRTSTSA
LSPGSKPSTWVSCPGEEENQATEDKRTERSKGARKTSVSPRSSQSDPRSWEYRSGEASEEKEEKAHEETG
KGEAAPGPQSSAPAQRPQLKSWWCQPSDEEESEVKPLGAAEKDGEAECPPCIPPPSAFLKAWVYWPGEDT
EEEEDEEEDEDSDSGSDEEEGEAEASSSTPATGVFLKSWVYQPGEDTEEEEDEDSDTGSAEDEREAETSA
STPPASAFLKAWVYRPGEDTEEEEDEDVDSEDKEDDSEAALGEAESDPHPSHPDQSAHFRGWGYRPQKET
EEEEAAEDWGEAEPCPFRVAIYVPGEKPPPPWAPPRLPLRLQRRLKRPETPTHDPDPETPLKARKVRFSE
KVTVHFLAVWAGPAQAARQGPWEQLARDRSRFARRIAQAQEELSPCLTPAARARAWARLRNPPLAPIPAL
TQTLPSSSVPSSPVQTTPLSQAVATPSRSSAAAAAALDLSGRRG

SEQ ID 108 N-myc downstream-regulated gene 2

MAELQEVQITEEKPLLPGQTPEAAKTHSVETPYGSVTFTVYGTPKPKRPAILTYHDVGLNYKSCFQPLFQ
FEDMQEIIQNFVRVHVDAPGMEEGAPVFPLGYQYPSLDQLADMIPCVLQYLNFSTIIGVGVGAGAYILAR
YALNHPDTVEGLVLINIDPNAKGWMDWAAHKLTGLTSSIPEMILGHLFSQEELSGNSELIQKYRNITHA
PNLDNIELYWNSYNNRRDLNFERGQDITLRCPVMLVVGDQAPHEDAVVECNSKLDPTQTSFLKMADSGGQ
PQLTQPGKLTEAFKYFLQGMGYMASSCMTRLSRSRTASLTSAASVDGNRSRSRTLSQSSESGTLSSGPPG
HTMEVSC

SEQ ID 109 non-metastatic cells 1, protein (NM23A) expressed in

MANCERTFIAIKPDGVQRGLVGEIIKRFEQKGFRLVGLKFMQASEDLLKEHYVDLKDRPFFAGLVKYMHS
GPVVAMVWEGLNVVKTGRVMLGETNPADSKPGTIRGDFCIQVGRNIIHGSDSVESAEKEIGLWFHPEELV
DYTSCAQNWIYE

SEQ ID 110 neuregulin 1

MSERKEGRGKGKGKKKERGSGKKPESAAGSQSPALPPQLKEMKSQESAAGSKLVLRCETSSEYSSLRFKW
FKNGNELNRKNKPQNIKIQKKPGKSELRINKASLADSGEYMCKVISKLGNDSASANITIVESNEIITGMP
ASTEGAYVSSESPIRISVSTEGANTSSSTSTSTTGTSHLVKCAEKEKTFCVNGGECFMVKDLSNPSRYLC
K

SEQ ID 111 oxidised low density lipoprotein (lectin-like) receptor 1

MTFDDLKIQTVKDQPDEKSNGKKAKGLQFLYSPWWCLAAATLGVLCLGLVVTIMVLGMQLSQVSDLLTQE
QANLTHQKKKLEGQISARQQAEEASQESENELKEMIETLARKLNEKSKEQMELHHQNLNLQETLKRVANC
SAPCPQDWIWHGENCYLFSSGSFNWEKSQEKCLSLDAKLLKINSTADLDFIQQAISYSSFPFWMGLSRRN
PSYPWLWEDGSPLMPHLFRVRGAVSQTYPSGTCAYIQRGAVYAENCILAAFSICQKKANLRAQ

SEQ ID 112 twinfilin-like protein

MAHQTGIHATEELKEFFAKARAGSVRLIKVVIEDEQLVLGASQEPVGRWDQDYDRAVLPLLDAQQPCYLL
YRLDSQNAQGFEWLFLAWSPDNSPVRLKMLYAATRATVKKEFGGGHIKDELFQTVKDDLSFAGYQKHLSS
CAAPAPLTSAERELQQIRINEVKTEISVESKHQTLQGLAFPLQPEAQRALQQLKQKMVNYIQMKLDLERE
TIELVHTEPTDVAQLPSRVPRDAARYHFFLYKHTHEGDELESVVFIYSMPGYKCSIKERMLYSSCKSRLL
DSVEQDFHLEIAKKIEIGDGAELTAEFLYDEVHPKQHAFKQAFAKPKGPGGKRGHKRLIRGPGENGDDS

SEQ ID 113 syntenin

MSLYPSLEDLKVDKVIQAQTAFSANPANPAILSEASAPIPHDGNLYPRLYPELSQYMGLSLNEEEIRANV
AVVSGAPLQGQLVARPSSINYMVAPVTGNDVGIRRAEIKQGIREVILCKDQDGKIGLRLKSIDNGIFVQL
VQANSPASLVGLRFGDQVLQINGENCAGWSSDKAHKVLKQAFGEKITMTIRDRPFERTITMHKDSTGHVG
FIFKNGKITSIVKDSSAARNGLLTEHNICEINGQNVIGLKDSQIADILSTSGTVVTITIMPAFIFEHIK
RMAPSIMKSLMDHTIPEV

SEQ ID 114 serine protease inhibitor, Kazal type 1

MKVTGIFLLSALALLSLSGNTGADSLGREAKCYNELNGCTKIYDPVCGTDGNTYPNECVLCFENRKRQTS
ILIQKSGPC

SEQ ID 115 tyrosine 3/tryptophan 5 -monooxygenase activation protein, eta polypeptide

MGDREQLLQRARLAEQAERYDDMASAMKAVTELNEPLSNEDRNLLSVAYKNVVGARRSSWRVISSIEQKT
MADGNEKKLEKVKAYREKIEKELETVCNDVLSLLDKFLIKNCNDFQYESKVFYLKMKGDYYRYLAEVASG
EKKNSVVEASEAAYKEAFEISKEQMQPTHPIRLGLALNFSVFYYEIQNAPEQACLLAKQAFDDAIAELDT
LNEDSYKDSTLIMQLLRDNLTLWTSDQQDEEAGEGN

SEQ ID 116 S100 calcium-binding protein A10

MPSQMEHAMETMMFTFHKFAGDKGYLTKEDLRVLMEKEFPGFLENQKDPLAVDKIMKDLDQCRDGKVGFQ
SFFSLIAGLTIACNDYFVVHMKQKGKK

SEQ ID 117 B-cell translocation gene 2

MSHGKGTDMLPEIAAAVGFLSSLLRTRGCVSEQRLKVFSGALQEALTEHYKHHWFPEKPSKGSGYRCIRI
NHKMDPIISRVASQIGLSQPQLHQLLPSELTLWVDPYEVSYRIGEDGSICVLYEEAPLAASCGLLTCKNQ
VLLGRSSPSKNYVMAVSS

SEQ ID 118 3'-phosphoadenosine 5'-phosphosulfate synthase 1

MEIPGSLCKKVKLSNNAQNWGMQRATNVTYQAHHVSRNKRGQVVGTRGGFRGCTVWLTGLSGAGKTTVSM
ALEEYLVCHGIPCYTLDGDNIRQGLNKNLGFSPEDREENVRRIAEVAKLFADAGLVCITSFISPYTQDRN
NARQIHEGASLPFFEVFVDAPLHVCEQRDVKGLYKKARAGEIKGFTGIDSEYEKPEAPELVLKTDSCDVN
DCVQQVVELLQERDIVPVDASYEVKELYVPENKLHLAKTDAETLPALKINKVDMQWVQVLAEGWATPLNG
FMRERFYLQCLHFDCLLDGGVINLSVPIVLTATHEDKERLDGCTARALMYEGRRVAILRNPEFFEHRKEE
RCARQWGTTCKNHPYIKMVMEQGDWLIGGDLQVLDRVYWNDGLDQYRLTPTELKQKFKDMNADAVFAFQL
RNPVHNGHALLMQDTHKQLLERGYRRPVLLLHPLGGWTKDDDVPLMWRMKQHAAVLEEGVLNPETTVVAI
FPSPMMYAGPTEVQWHCRARMVAGANFYIVGRDPAGMPHPETGKDLYEPSHGAKVLTMAPGLITLEIVPF
RVAAYNKKKKRMDYYDSEHHEDFFISGTRMRKLAREGQKPPEGFMAPKAWTVLTEYYKSLEKA

SEQ ID 119 biogenesis of lysosome-related organelles complex-1, subunit 2

MFSKMATYLTQELTATSEDYKLLENMNKLTSLKYLEMKDIAINISRNLKDLNQKYAGLQPYLDQINVIEE
QVAALEQAAYKLDAYSKKLEAKYKKLEKR

SEQ ID 120 mouse double minute 2 homolog

MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYLGQYIMTKRLYDEK
QQHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIYRNLVVVNQQESSDSGTSVSENRCHLEGGSDQKDLVQ
ELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGERQRKRHKSDSISLSFDESLALCVIREICCERS
SSSESTGTPSNPDLDAGVSEHSGDWLDQDSVSDQFSVEFEVESLDSEDYSLSEEGQELSDEDDEVYQVTV
YQAGESDTDSFEEDPEISLADYWKCTSCNEMNPPLPSHCNRCWALRENWLPEDKGKDKGEISEKAKLENS
TQAEEGFDVPDCKKTIVNDSRESCVEENDDKITQASQSQESEDYSQFSTSSSIIYSSQEDVKEFEREETQ
DKEESVESSLPLNAIEPCVICQGRPKNGCIVHGKTGHLMACFTCAKKLKKRNKPCPVCRQPIQMIVLTYF