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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 July 2002 (18.07.2002)

PCT

(10) International Publication Number
WO 02/055985 A2

(51) International Patent Classification[7]: G01N

(21) International Application Number: PCT/US01/45340

(22) International Filing Date:
1 November 2001 (01.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/248,882 15 November 2000 (15.11.2000) US

(71) Applicant: ROCHE DIAGNOSTICS CORPORATION [US/US]; 9115 Hague Road Bldg. D, P.O. Box 50457, Indianapolis, IN 46250-0457 (US).

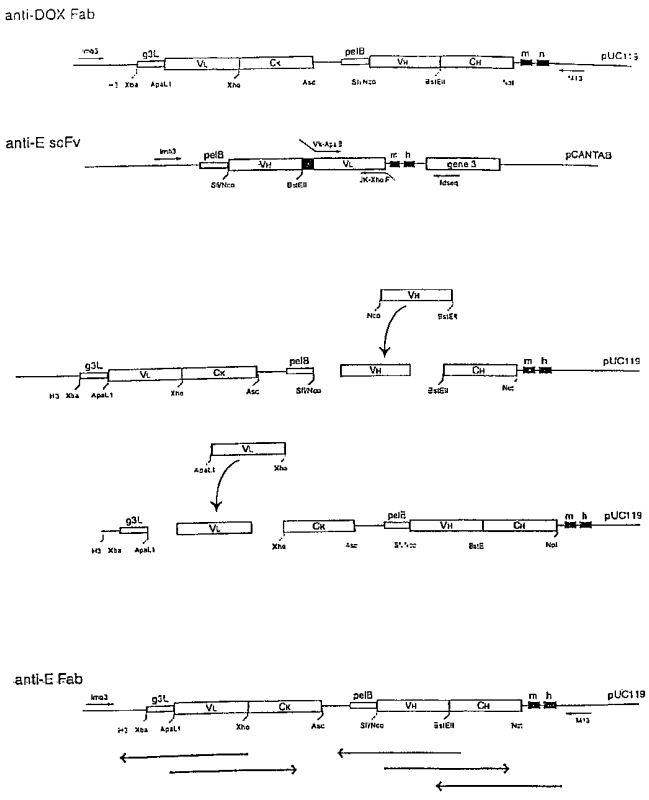
(72) Inventors: SCHUELER, Paula, A.; 9115 Hague Road Bldg. D, P.O. Box 50457, Indianapolis, IN 46250-0457 (US). XU, Hongxia; 9115 Hague Road Bldg. D, P.O. Box 50457, Indianapolis, IN 46250-0457 (US). FOLTZ, Lisa; 9115 Hague Road Bldg. D, P.O. Box 50457, Indianapolis, IN 46250-0457 (US). WU, Xingyong; 9115 Hague Road Bldg. D, P.O. Box 50457, Indianapolis, IN 46250-0457 (US). SHA, Yehsiung; 9115 Hague Road Bldg. D, P.O. Box 50457, Indianapolis, IN 46250-0457 (US). NAGY, Alexandra; 9115 Hague Road Bldg. D, P.O. Box 50457, Indianapolis, IN 46250-0457 (US). MAHONEY, Walter, C.; 9115 Hague Road Bldg. D, P.O. Box 50457, Indianapolis, IN 46250-0457 (US).

(74) Agents: WEILD, David, III et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,

[Continued on next page]

(54) Title: METHODS AND REAGENTS FOR IDENTIFYING RARE FETAL CELLS IN THE MATERIAL CIRCULATION



(57) Abstract: This invention provides methods and compositions useful for identifying and diagnosing rare fetal cells in a mixed cell population such as a maternal blood sample. The methods entail the use of specific nucleic acid probes that hybridize to fetal cell associated RNAs to identify the rate fetal cells or antibodies that bind to polypeptides encoded by the fetal cell associated RNAs for fetal cell detection. The cells detected by the methods of the present invention are useful for diagnosing the fetal cells for a genetic trait of interest, such as trisomy 21. Novel methods for simultaneous screening for fetal cells and diagnosing the fetal cells are also provide. Compositions comprising the fetal cell associated nucleic acids of the invention and their encoded proteins are also provided. The present invention further provides kits useful for practicing the present methods.

WO 02/055985 A2

MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

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

Published:

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

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

METHODS AND REAGENTS FOR IDENTIFYING RARE FETAL CELLS IN THE MATERNAL CIRCULATION

1. FIELD OF THE INVENTION

This invention relates generally to the fields of cell purification, cell identification, and prenatal genetic analysis. More particularly, the invention provides methods and compositions for identifying individual cells of fetal origin in samples of maternal blood. The methods encompass the use of specific nucleic acid probes to identify the rare fetal cells in the maternal blood sample and optionally further diagnosing the detected fetal cells for a genetic trait of interest. Compositions comprising the nucleic acid probes and kits useful in the present methods are also provided.

2. BACKGROUND OF THE INVENTION

Amniocentesis and chorionic villus sampling are the currently accepted methods for prenatal testing for genetic abnormalities. However, both of these procedures are invasive and are accompanied by a small risk (on the order of 1%) of fetal death. Obtaining and identifying fetal cells in the maternal circulation holds considerable promise for prenatal genetic testing. Particularly advantageous is the fact that the test sample is obtained by a relatively non-invasive procedure that poses essentially no risk to the fetus. In addition, since fetal cells peak in the maternal circulation at about 10-16 weeks of gestation, it is possible to perform the genetic analysis at an early stage in pregnancy. However, fetal cells are extremely rare in maternal blood, on the order of 1 to 50 cells per 10^7 nucleated blood cells. These low levels of fetal cells make even minimal levels of non-specific binding problematic for affinity separation of fetal cells from maternal cells.

A number of fetal cells are known to make their way into the maternal circulation, including leukocytes, trophoblast cells and nucleated red blood cells. Leukocytes have been generally excluded from consideration as targets for isolation from maternal blood for a number of reasons, including a lack of generic markers for use in isolation as well as the possibility of persistence in maternal blood of fetal leukocytes from previous pregnancies. Trophoblast cells have been considered undesirable due to concerns that these cells may be subject to confined placental mosaicism, rendering them unrepresentative of the fetus. Nucleated fetal erythroid cells, however, are considered an attractive target for prenatal genetic analysis.

The isolation of nucleated fetal erythroid cells from maternal blood has, however, been fraught with difficulty. The low abundance of these cells in maternal blood renders separation extremely difficult, as even extremely low non-specific binding by a

separation reagent will result in large numbers of maternal cells if the reagent positively selects for the fetal cells, and unacceptably low yields if the antibody negatively selects for maternal cells.

Also, no fetal blood cell specific markers are known in the art. Enrichment of fetal blood cells has been performed using markers such as fetal hemoglobin (Hemoglobin F or HbF), which is estimated to be present in 0.1% to 0.7% of erythroid cells in normal adult blood. Immature erythroid cells (*i.e.*, cells of the erythroid lineage at the reticulocyte stage and earlier) express markers which have been used to enrich fetal blood cells (*e.g.*, glycophorin A, CD36, and the transferrin receptor, also known as TfR and CD71). However, these markers can also be found on cells in adult blood, and it has also been found that blood samples taken during pregnancy contain relatively high levels of maternal immature erythroid cells.

A variety of methods have been proposed for isolation or enrichment of fetal cells in maternal blood. These methods include centrifugation techniques, immunoaffinity techniques, and fluorescent *in situ* hybridization (FISH) methods. However, these methods suffer from a number of deficiencies.

Centrifugation methods generally rely on density gradients for separation of nucleated from non-nucleated cells and frequently include a lysis step to eliminate erythrocytes. See, for example, U.S. Patents Nos. 5,432,054, and 5,646,004, International Patent Application No. WO 95/09245 and Rao et al. (1994, Ann. NY Acad. Sci. 731:142-143). However, these techniques co-enrich large numbers of maternal nucleated erythroid cells, and so do not provide the level of enrichment required for reproducible genetic screening of fetal cells.

Immunoaffinity approaches have been described using a variety of different antibodies. However, most approaches rely on the use of antibodies directed to markers in the erythroid pathway. For example, Bianchi et al. (1993, Prenatal Diag. 13:293-300) describes a method utilizing CD71 (transferrin receptor) CD36 (thrombospondin receptor) and/or glycophorin A antibodies for flow sorting to enrich 'fetal' cells from maternal blood. The use of antibodies to erythroid cell markers such as CD71, CD36 and glycophorin co-enriches maternal erythroid cells, which substantially outnumber fetal erythroid cells in maternal blood samples.

Fluorescent *in situ* hybridization methods have been used to sort cells which express particular RNAs from maternal blood samples. These methods suffer from the same problems as immunoaffinity methods, due to the lack of fetal cell specific probes. WO 96//17085 teaches the use of probes specific for HLA-G, a non-classical Class I MHC molecule which is an oncofetal marker found on extravillous cytotrophoblast cells, for use in sorting HLA-G expressing cells from samples, such a maternal blood.

Fetal cells present in the maternal circulation are at various stages of development. As with other cell types, expression patterns of cellular markers change as a given cell proceeds down a developmental pathway. Reagents for identifying fetal cells must accommodate such variations. Accordingly, there is a need in the art for new reagents and methods for separation and identification of fetal cells in maternal blood.

Citation or identification of any reference herein shall not be construed as an admission that such reference is available as prior art to the present invention.

3. SUMMARY OF THE INVENTION

The present invention provides methods for detecting a fetal cell in a maternal blood sample, comprising the steps of: (a) contacting said maternal blood sample with a first probe comprising a nucleotide sequence corresponding to SEQ ID NO: 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, which first probe selectively or specifically hybridizes to mRNA in fetal cells if present in the maternal blood sample; and (b) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell.

The present invention further provides methods for detecting a fetal cell in a maternal blood sample, comprising the steps of: (a) contacting said maternal blood sample with a first probe comprising a nucleotide sequence having at least 90% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, which first probe selectively or specifically hybridizes to fetal cells if present in the maternal blood sample; and (b) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell.

The present invention yet further provides methods for detecting a fetal cell in a maternal blood sample, comprising the steps of: (a) contacting said maternal blood sample with a first probe comprising a nucleotide sequence having at least 80% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, which first probe selectively or specifically hybridizes to fetal cells if present in the maternal blood sample; and (b) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell.

The present invention yet further provides methods for detecting a fetal cell in a maternal blood sample, comprising the steps of: (a) performing differential expression

analysis on RNA or cDNA obtained from fetal liver myeloid cells relative to RNA or cDNA obtained from mature myeloid cells; (b) identifying an RNA or cDNA species that is selectively or specifically expressed in the fetal liver myeloid cells, thereby identifying an RNA or cDNA species that is useful as a probe for fetal cells in the maternal circulation; (c) contacting the maternal blood sample with a probe comprising a nucleotide sequence corresponding to all or a portion of the RNA or cDNA of step (b); and (d) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell.

The present invention further provides methods for diagnosing an abnormality in a fetal cell, comprising the steps of: (a) contacting a maternal blood sample with a first probe comprising a nucleotide sequence corresponding to SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, which first probe selectively or specifically hybridizes to mRNA in the fetal cell if present in the maternal blood sample; (b) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell; and (c) if the maternal blood sample comprises a fetal cell, determining whether the abnormality exists in said fetal cell, thereby diagnosing the abnormality. The fetal cell detection and diagnostic steps can be performed concurrently or successively (in either order).

The present invention further provides methods for diagnosing an abnormality in a fetal cell, comprising the steps of: (a) contacting a maternal blood sample comprising said fetal cell with a first probe comprising a nucleotide sequence having at least 90% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, which first probe selectively or specifically hybridizes to mRNA in the fetal cell if present in the maternal blood sample; (b) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell, and (c) if the maternal blood sample contains a fetal cell, determining whether the abnormality exists in said fetal cell, thereby diagnosing the abnormality. The fetal cell detection and diagnostic steps can be performed concurrently or successively (in either order).

The present invention further provides methods for diagnosing an abnormality in a fetal cell, comprising the steps of: (a) contacting a maternal blood sample comprising said fetal cell with a first probe comprising a nucleotide sequence having at least 80% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13,

14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, which first probe selectively or specifically hybridizes to mRNA in the fetal cell if present in the maternal blood sample; (b) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell; and (c) if the maternal blood sample contains a fetal cell, determining whether the abnormality exists in said fetal cell, thereby diagnosing the abnormality. The fetal cell detection and diagnostic steps can be performed concurrently or successively (in either order).

The present invention yet further provides methods for diagnosing an abnormality in a fetal cell, comprising the steps of: (a) performing differential expression analysis on RNA or cDNA obtained from fetal liver myeloid cells relative to RNA or cDNA obtained from mature myeloid cells; (b) identifying an RNA or cDNA species that is selectively or specifically expressed in the fetal liver myeloid cells, thereby identifying an RNA or cDNA species that is useful as a probe for fetal cells in the maternal circulation; (c) contacting the maternal blood sample with a probe comprising a nucleotide sequence corresponding to all or a portion of the RNA or cDNA of step (b); (d) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell; and (e) if the maternal blood sample comprises a fetal cell, determining whether the abnormality exists in said fetal cell, thereby diagnosing the abnormality.

The fetal cell detection and diagnosis methods of the present invention optionally further comprise contacting the maternal blood sample with a second probe which selectively or specifically hybridizes to fetal cells if present in the maternal blood sample prior to identifying whether a fetal cell is present in the maternal blood sample, and, optionally, detecting a cell in said maternal blood sample which comprises mRNA that hybridizes to the second probe. The second probe is preferably labeled with the same type of label as the first probe. The first and second probes can correspond to the same mRNA or to different mRNAs. In a preferred embodiment, the first probe corresponds to the J42-4d gene (SEQ ID NO:11) and the second probe corresponds to fetal epsilon globin. Such probes are preferably at least 25, most preferably at least 30, and most preferably 150-200 nucleotides in length. The probes are also preferably riboprobe prepared according to the method described in Section 8.1 below.

In certain embodiments of the fetal cell detection and diagnosis methods of the present, the maternal blood sample is immunoenriched for fetal cells prior to contacting the blood sample with the fetal cell specific or selective probe. The maternal blood sample can be positively immunoenriched by (a) contacting the maternal blood sample with an antibody that selectively or specifically binds to fetal cells in the maternal blood sample;

and (b) separating cells in the maternal blood sample that bind to the antibody from cells that do not bind to the antibody, thereby immunoenriching the maternal blood sample for fetal cells. Alternatively, the maternal blood sample can be negatively immunoenriched by (a) contacting the maternal blood sample with an antibody that selectively or specifically binds to maternal cells in the maternal blood sample; and (b) separating cells in the maternal blood sample that do not bind to the antibody from cells that bind to the antibody, thereby immunoenriching the maternal blood sample for fetal cells.

The diagnostic methods of the invention can be used to detect chromosomal abnormalities. In certain specific embodiments, the chromosomal abnormalities are aneuploidies, including but not limited to trisomy 13, trisomy 21, or Klinefelter or other sex chromosome syndromes. In other specific embodiments, the chromosomal abnormalities are single gene disorders. The single gene disorder can be a deletion, insertion or substitution disorder. In exemplary embodiments, the single gene disorder is spina bifida, sickle-cell anemia, a thalassemia, Marfan Syndrome, Duchenne Muscular Dystrophy, or cystic fibrosis. In yet other embodiments, the single gene disorders detected by the methods of the present invention are nucleoeotide triplet expansions in one or more genes. Such genes include but are not limited to the Fragile X Syndrome gene, the Friedreich's ataxia gene, the myotonic dystrophy gene, or the Huntington's disease genes. In yet other embodiments, the chromosomal abnormalities are viral sequences, *e.g.*, HIV sequences, inserted in the fetal cell genome.

The present invention yet further provides methods for identifying a nucleic acid useful as a probe for fetal cells in the maternal circulation, comprising the steps of: (a) performing differential expression analysis on RNA or cDNA obtained from fetal liver myeloid cells relative to RNA or cDNA obtained from mature myeloid cells; and (b) identifying an RNA or cDNA species that is selectively or specifically expressed in the fetal liver myeloid cells, wherein the RNA or cDNA species that is selectively or specifically expressed in the fetal liver myeloid cells is useful as a probe for fetal cells in the maternal circulation. In a preferred embodiment, the fetal liver is human fetal liver. In another preferred embodiment, the fetal liver myeloid cells are obtained before 20 weeks of gestation. In preferred modes of the embodiment, the fetal liver myeloid cells are obtained between 10 and 15 weeks of gestation, *e.g.*, at 10, 10.5, 11, 11.5, 12, 12.5, 13, 13.5, 14, 14.5, or 15 weeks of gestation. The mature myeloid cells can be fetal cord blood cells obtained after 20 weeks of gestation, fetal peripheral blood cells obtained after 20 weeks of gestation, fetal liver myeloid cells obtained after about 20 weeks of gestation, adult bone marrow cells or adult peripheral blood cells. In yet another preferred embodiment, the differential expression analysis comprises subtraction suppression hybridization.

The present invention yet further provides kits comprising in one or more containers (a) a first probe comprising a nucleotide sequence corresponding to SEQ ID

NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, which first probe selectively or specifically hybridizes to mRNA in fetal cells if present in a maternal blood sample and (b) instructions for diagnostic use or a label indicating regulatory approval for diagnostic use. In other embodiments, the present invention provides kits comprising in one or more containers a first probe comprising a nucleotide sequence having at least 90% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, which first probe selectively or specifically hybridizes to fetal cells if present in a maternal blood sample. In yet other embodiments, the present invention provides kits comprising in one or more containers a first probe comprising a nucleotide sequence having at least 80% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, which first probe selectively or specifically hybridizes to fetal cells if present in a maternal blood sample. The kits can further comprise one or more antibodies for immunoenriching for fetal cells in a maternal blood sample, for example an antibody that selectively or specifically binds to fetal cells in a maternal blood sample. The kits can also optional comprise a second probe that selectively or specifically hybridizes to mRNA in fetal cells if present in a maternal blood sample. The second prove can comprise (i) a nucleotide sequence corresponding to SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42; (ii) a nucleotide sequence having at least 90% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42; or (iii) a nucleotide sequence having at least 80% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42. The second probe can correspond to the same or a different fetal cell specific or selective mRNA as the first probe. The kits of the invention can further include diagnostic reagents for determining the gender of the fetal cells or for identifying abnormalities associated with the fetal cells.

In the foregoing fetal cell detection and diagnosis methods and related kits of the present invention, the first probe can be designed to either specifically or selectively hybridize to fetal cells. The first probe is preferably labeled, for example by a radioactive or fluorescent label, a colorimetric reagent, or an enzyme.

The fetal cell detection and diagnosis methods and related kits of the present invention utilize probes having sequences that hybridize to RNAs in fetal cells to a greater extent than RNAs found in non-fetal, *e.g.*, maternal cells in a mixed cell population. Such probes comprise a nucleotide sequence having 20-30 nucleotides with at least 80%

sequence identity to a corresponding portion of a fetal cell specific or selective transcript, *e.g.*, SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42. In other embodiments, the nucleotide sequence has at 30-40, 40-60, 60-80, 80-100, 100-150, 150-200, or greater than 200 nucleotides with at least 80% sequence identity to corresponding portion of a fetal cell specific or selective transcript. In various embodiments, the nucleotide sequence has at least 65%, more preferably at least 85%, yet more preferably at least 95% sequence identity to a corresponding portion, *e.g.*, a 30-40, 40-60, 60-80, 80-100, 100-150, 150-200 or greater than 200 nucleotide portion, of a fetal cell specific or selective transcript. In one specific embodiment, the nucleotide sequence has 100% identity to a corresponding portion of a fetal cell specific or selective transcript.

In certain embodiments, a first probe of the invention is less than 40, 50, 100, 200, 300, 400, 500, 1000 or 1500 nucleotides in length. The probe can be an RNA probe, a DNA probe, or a chimeric probe. The probe is preferably single stranded, but can also be partially double stranded.

The fetal cell sought to be detected or diagnosed by the methods and compositions of the present invention is preferably an erythroblast or a trophoblast.

The present invention further provides isolated nucleic acid molecules selected from the group consisting of: (a) a nucleic acid molecule having a nucleotide sequence which is at least 90% identical to the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or a complement thereof; (b) a nucleic acid molecule comprising at least 15 nucleotide residues and having a nucleotide sequence identical to at least 15 consecutive nucleotide residues of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or a complement thereof, (c) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78; (d) a nucleic acid molecule which encodes a fragment at least 10 consecutive amino acid residues of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78; (e) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of any SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78; wherein the fragment comprises consecutive amino acid residues corresponding to at least half of the full length of any of said SEQ ID NOs; (f) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of any of SEQ ID

NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78, wherein the nucleic acid molecule hybridizes with a nucleic acid molecule consisting of the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42 under stringent conditions, or a complement thereof. In certain preferred embodiments, the nucleic acid molecule is selected from the group consisting of: (a) a nucleic acid having the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or a complement thereof; and (b) a nucleic acid molecule which encodes a polypeptide having the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78, or a complement thereof. The isolated nucleic acids of the invention can further optionally comprise vector nucleic acid sequences and/or nucleic acid nucleic acid sequences encoding a heterologous polypeptide. The present invention also encompasses prokaryotic and eukaryotic host cells, including but not limited to mammalian and non-mammalian, *e.g.*, bacterial, host cells, which contain the nucleic acid molecules of the invention.

The present invention further provides isolated polypeptides selected from the group consisting of: (a) a fragment of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78; (b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes with a nucleic acid molecule consisting of the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42 under stringent conditions, or a complement thereof; (c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 90% identical to a nucleic acid consisting of the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or a complement thereof. In certain embodiment of the invention, the isolated polypeptides have the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78. The polypeptides of the invention can further comprises heterologous amino acid residues. The present invention further encompasses methods for producing a polypeptide selected from the group consisting of: (a) a polypeptide comprising the amino acid sequence of any of SEQ ID NOs: 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48,

49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78; (b) a polypeptide comprising a fragment of at least 10 contiguous amino acids of the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78; and (c) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes with a nucleic acid molecule consisting of the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or a complement thereof under stringent conditions; the methods comprising culturing a host comprising a nucleic acid of the invention under conditions in which the nucleic acid molecule is expressed.

3.1. DEFINITIONS

SPECIFIC MARKER: a marker (protein, nucleic acid, carbohydrate or other compound) which is found only in or on the target cell type among other cell types in a biological sample of interest. For example, a fetal erythroid cell phenotype specific marker is a marker that is found only on fetal cells of the erythroid lineage, but cannot be detected in/on other cells from the fetus or mother.

SELECTIVE MARKER: a marker that is found predominantly on or in the target cell type, but may be found in other cells as well. For example, fetal hemoglobin is found in fetal blood cells, as well as in a small percentage of maternal blood cells. The selective marker is preferably at least five times more abundant in a target cell relative to a non-target cell in the biological sample of interest, more preferably at least 10, 15, 20, 25, 30, 35, 40, 45 or 50 times more abundant in the target cell relative to a non-target cell in the biological sample of interest, *e.g.*, a maternal blood sample.

SPECIFIC: a nucleic acid used in a reaction, such as a probe used in a hybridization reaction, a primer used in a PCR, or a nucleic acid present in a pharmaceutical preparation, is referred to as "specific" if it hybridizes or reacts only with the intended target. Similarly, a polypeptide is referred to as "specific" if it binds only to its intended target, such as a ligand, hapten, substrate, antibody, or other polypeptide. An antibody is referred to as "specific" if it binds only to the intended target.

SELECTIVE: a nucleic acid used in a reaction, such as a probe used in a hybridization reaction, a primer used in a PCR, or a nucleic acid present in a pharmaceutical

preparation, is referred to as "selective" if it hybridizes or reacts with the intended target more frequently, more rapidly, or with greater duration than it does with alternative substances. Similarly, a polypeptide is referred to as "selective" if it binds an intended target, such as a ligand, hapten, substrate, antibody, or other polypeptide more frequently, more rapidly, or with greater duration than it does to alternative substances. An antibody is referred to as "selective" if it binds via at least one antigen recognition site to the intended target more frequently, more rapidly, or with greater duration than it does to alternative substances.

ASSOCIATED: specific or selective.

CORRESPOND OR CORRESPONDING: Between nucleic acids, "corresponding" means homologous to or complementary to a particular sequence or portion of the sequence of a nucleic acid. As between nucleic acids and polypeptides, "corresponding" refers to amino acids of a peptide in an order derived from the sequence or portion of the sequence of a nucleic acid or its complement.

ERYTHROID: an immature cell of the erythroid lineage (*i.e.*, a cell of the erythroid lineage which is not a mature erythrocyte). Erythroid cells include reticulocytes, orthochromatic erythroblasts, polychromatophilic erythroblasts, basophilic erythroblasts, proerythroblasts, colony forming unit-erythroid (CFU-E) and burst forming unit-erythroid (BFU-E).

NUCLEIC ACID OF THE INVENTION: A nucleic acid comprising a nucleotide sequence corresponding to all or a portion of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 and 42, or a variant or derivative thereof.

POLYPEPTIDE OF THE INVENTION: A polypeptide comprising an amino acid sequence corresponding to all or a potion of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, and 78, or a variant or derivative thereof.

FETAL CELL PROBE: A nucleic acid that specifically or selectively hybridizes with a fetal cell RNA or its complement relative to RNAs in other cells in a sample of interest, *e.g.*, non-fetal cells in a maternal blood. A fetal cell probe can be labeled and used for detection of fetal cell RNA. A fetal cell probe can also be in the form of an

oligonucleotide useful for PCR amplification of a cDNA corresponding to said fetal cell RNA.

TARGET CELL: a cell of fetal origin in a mixed cell population.

REFERENCE CELLS: a cell that is not a target cell in a mixed cell population.

4. BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a genetic map showing the strategy for converting specific scFv antibody into Fab antibody. Phagemids from a naive scFv library are cloned and selected for the correct antigen binding characteristics. The immunoglobulin V_H and V_L regions encoded in the scFv insert are then excised and substituted for the separately translated V_H and V_L encoding regions in the anti-DOX Fab vector.

FIG. 2 is a three-panel figure showing half-tone reproductions of the analysis of erythroblast antigen. The antigen was immunopurified from a cell extract using Clone 1 anti-erythroblast antibody, subsequently captured on a Nickel absorbant. Upper Panel: Silver-stained polyacrylamide gel; Lower Panels: Western blot at two different exposures. The arrow indicates the position of a specifically identified antigen with an apparent molecular weight of ~90 kDa.

FIG. 3 is a two-panel figure showing half-tone reproductions of the analysis of erythroblast antigen. The gels were stained with Coomassie Brilliant Blue. Upper Panel: antigen purified by capturing Clone 1 antibody using Nickel absorbant; Lower Panel: antigen purified by capturing biotinylated Clone 1 antibody using streptavidin Dynabeads™. The arrows indicate two bands with apparent molecular weights of ~90 kDa and ~78 kDa.

FIG. 4 is a three-panel figure showing half-tone reproductions of the analysis of erythroblast antigen. Upper Panel: Silver-stained polyacrylamide gel; Lower Panels: Western blot from two separate experiments. The analysis compares antigen immunopurified using different erythroblast-specific antibody clones. Clones 18 and 28 appear to recognize ~90 kDa and ~78 kDa bands comigrating with those recognized by Clone 1. Clones 17, 22, and 23 appear to identify different bands.

FIG. 5 is a four-panel half-tone figure showing test results for Clone 95 nucleic acid as a probe for fetal cells. The top two panels are a "Southern" (virtual Northern) blot of cDNA from various tissue sources, probed with Clone 95, and shown at

two different exposures. The middle panel is an extended Southern analysis using cDNA from a larger panel of tissue samples. The lower panel is a Northern blot of mRNA from different tissues probed with Clone 95.

FIG. 6 is a two-panel half-tone figure showing test results for Clone 369 nucleic acid as a probe for fetal cells. The top panel is a Southern blot; the lower panel is a Northern blot.

FIG. 7 is a six-panel half-tone figure showing the testing of a probe for *in situ* hybridization. The cells were obtained from human fetal liver blood collected at 16 weeks gestation. The cells were overlaid with the probe at concentrations of 0, 10, and 40 μg/mL (left to right). The staining pattern is consistent with hybridization of the probe with a complementary mRNA sequence present in the cytoplasm.

FIG. 8 is a three-panel half-tone figure showing the results of a genetic analysis according to the invention. Upper left panel shows the phase contrast photomicrograph of cells enriched from a maternal blood sample. The cells were obtained by affinity enrichment using the anti-erythrocyte antibody from Clone 1 in a magnetic activated cell sorting technique. The panel to the right shows the cytoplasmic staining pattern obtained by *in situ* hybridization using the nucleic acid probe from Clone 369. This specifies which cells in the field are fetal in origin. The lower panel shows the nuclear staining pattern obtained by *in situ* hybridization using a nucleic acid probe specific for a repeat sequence of Chromosome Y, developed using a different fluorescent marker. The single dot appears in the same cell as the probe, and characterizes the genotype of the fetal cell as having a single Y chromosome.

FIG. 9A and 9B shows Northern blot data of various tissues obtained from the experiments described in Section 7.2. The probes used in to probe the blots shown in this figure correspond to SEQ ID NO:34 and related clones.

FIG. 10A and 10B shows Northern blot data of various tissues obtained from the experiments described in Section 7.2. The probes used in to probe the blots shown in this figure correspond to SEQ ID NOs:24 and 27 and related clones.

FIG. 11A and 11B shows Northern blot data of various tissues obtained from the experiments described in Section 7.2. The probes used in to probe the blots shown in this figure correspond to SEQ ID NO:21 and related clones.

FIG. 12A and 12B shows Northern blot data of various tissues obtained from the experiments described in Section 7.2. The probes used in to probe the blots shown in this figure correspond to SEQ ID NO:36 and related clones.

FIG. 13A and 13B shows Northern blot data of various tissues obtained from the experiments described in Section 7.2. The probes used in to probe the blots shown in this figure correspond to SEQ ID NOs:15 and 31 and related clones.

FIG. 14A and 14B shows Northern blot data of various tissues obtained from the experiments described in Section 7.2. The probes used in to probe the blots shown in this figure correspond to SEQ ID NO:10 and related clones.

FIG. 15A and 15B shows Northern blot data of various tissues obtained from the experiments described in Section 7.2. The probes used in to probe the blots shown in this figure correspond to SEQ ID NO:41 and related clones.

FIG. 16A and 16B shows two schematics of the probe signal amplification method that is preferred for detecting fetal cell associated RNAs, as described in Section 8, *infra*.

FIG. 17 Cord blood cells stained for DAPI (nuclear stain) and L15-1A (cytoplasmic localization) as described in Section 8 below.

FIG. 18 Cord blood cells stained for DAPI (nuclear stain) and L15-1A (cytoplasmic localization) as described in Section 8 below.

FIG. 19A and 19B Cord blood cells (CB) or bone marrow cells (BM) stained for DAPI (nuclear stain), with or without one or both of J42-4d (cytoplasmic localization), fetal globin epsilon (cytoplasmic localization).

FIG. 20 Cord blood cells stained for DAPI (nuclear, diffuse), gamma and epsilon globins (cytoplasmic) and X and Y chromosomes (subnuclear) using the simultaneous detection methodology of Section 5.9.5.

5. DETAILED DESCRIPTION OF THE INVENTION

The present invention provides nucleic acids probes that are useful for identifying a blood cell of fetal origin in a mixed cell population, *e.g.*, a maternal blood sample. The nucleic acid probes are adapted to hybridize with RNA (typically mRNA)

present in the fetal cell, or, in some instances, to cDNA reverse transcribed from the RNA. Thus, the fetal cell can be distinguished from maternal cells or other cells that may be present in the mixed population (the "reference cells"), and separated or analyzed *in situ*. These are referred to in this disclosure as "probes."

5.1. IDENTIFICATION OF FETAL CELL ASSOCIATED RNAS

The present invention provides a methods for discovering nucleic acids the are preferentially or uniquely expressed in fetal cells relative to other cells, *e.g.*, cells of maternal origin, in a mixed cell population. Such nucleic acids are useful for designing probes for identifying fetal cells in a mixed cell population. Polypeptide translation products of such RNAs can be used to prepare antibodies that can be used select for fetal cells in a mixed cell population.

The methods described herein take advantage of the key role the liver plays in the production of fetal cells during gestation. At about 8 weeks, the fetal liver takes over from the yolk sac as the main source of fetal blood cells of all types, including erythroid cells and their precursors. Peak production occurs from about 10-20 weeks of gestation, after which the bone marrow begins to take over. Production of erythroid cells by the liver drops to about 20% of peak levels by week 30, and is virtually absent at term. Fetal liver is therefore an excellent source for RNA species that are more highly expressed in fetal blood cells compared with maternal blood cells. Erythroid cells are easily obtained from fetal liver samples collected between 9-20 weeks of gestation. A preferred collection period is at 16-20 weeks, which corresponds to the highest concentration of nucleated erythroid cells, as a percentage of total cells present. A preferred source is human fetal liver, although other species can be used as a substitute, adjusting gestation times as appropriate. Once a fetal cell associated RNA is identified in a non-human species, the corresponding human homolog can be identified and its expression analyzed to confirm that its expression is associated with cells of fetal rather than maternal origin.

The methods provided herein entail the use of differential expression analysis to identify RNAs that are associated with fetal cells. Generally, the differential expression methods provided herein entail manipulating RNAs obtained fetal cell to either (a) eliminate or reduce RNAs found in cells which are likely to contaminate the test sample or (b) amplify those RNAs which are not found (or found at reduced levels) in cells likely to contaminate the test sample. The differential expression analysis methods identify on a molecular level RNA or cDNA molecules ("tags") absent from or present at relatively lower amounts in "driver RNA" or "driver cDNA" prepared from "reference cells" (cells which should not be identified by the probe sequence, *e.g.*, maternal blood cells), and present (at relatively higher amounts) in "tester RNA" or "tester cDNA" prepared from target cells

(cells which should be identified by the probe sequence, *e.g.*, fetal cells). Such differential expression analysis techniques are discussed in Section 5.1.1, *infra*.

With respect to the fetal liver as a source of tester nucleic acid, it is preferable that the tissue be chilled very shortly after harvesting, and that mRNA be prepared from the tissue as soon as possible. The erythroid cells are easily separated from hepatic parenchymal cells by gentle manipulation followed by low-speed or gradient centrifugation.

The success of the present methods in identifying probes that are specific to fetal cells or immature erythroid cells is demonstrated in FIG. 17-19, which show examples of nucleated cells in cord blood cells or human bone marrow cells detected through the presence of DAPI stained nuclei using the DAPI channel (blue), and peroxidase-antibody cascade complexes were detected using TSA-green through the FITC channel (green).

Following differential expression analysis, it is preferably to "validate" the tags as fetal cell specific or fetal cell selective. "Validation" of the specificity or selectivity generally involves clonally expanding each candidate tag, and then evaluating its characteristics by further analysis. Exemplary validation steps to ensure the specificity or selectivity of the fetal cell tags are discussed in Section 5.1.2, below.

Having identified tag sequences with desirable specificity characteristics, further characterization of the RNA or corresponding which is the source of the probe sequence can be performed. Expression patterns can be determined by *in situ* hybridization using various tissue sections. The full length sequence of the cloned DNA insert can be obtained, and modified probes and primers can be designed. The sequence can be used to pull out overlapping inserts from a cDNA library obtained by SSH or by reverse transcription of fetal mRNA, for example, by the CapFinder™ technique, and the sequence of the entire transcript can be determined. Probe sequences can then be obtained that hybridize anywhere along the transcript. The encoding region can be identified, and the amino acid sequence of the translation product can be predicted. The encoded polypeptides can be recombinantly expressed and used for making antibodies, which antibodies can be used in the fetal cell detection methods of the present invention.

5.1.1. DIFFERENTIAL EXPRESSION METHODS

A variety of methods are known in the art for identifying differentially expressed RNAs that can be used to identify fetal cell associated tags, which can then be used to screen for the corresponding full length cDNAs. Additonally, proteome methods may be used to identify polypeptides and their corresponding RNAs or cDNAs that are differentially expressed among maternal and fetal cells. "Differential expression," as the term is used herein, is understood to refer to both quantitative as well as qualitative

differences in expression patterns, *e.g.*, of a gene or genes, between target cells (*e.g.*, fetal cells) and reference cells (*e.g.*, maternal cells).

Methods of differential expression are well-known to one skilled in the art, and include but are not limited to differential display, serial analysis of gene expression (SAGE), nucleic acid array technology, subtractive hybridization, proteome analysis and mass-spectrometry of two-dimensional protein gels. The methods of gene expression profiling are exemplified by the following references describing differential display (Liang and Pardee, 1992, Science 257:967-971), proteome analysis (Humphery-Smith *et al.*, 1997, Electrophoresis 18:1217-1242; Dainese *et al.*, 1997, Electrophoresis 18:432-442), SAGE (Velculescu *et al.*, 1995, Science 270:484-487), subtractive hybridization (Wang and Brown, 1991, Proc. Natl. Acad. Sci. U.S.A. 88:11505-11509), and hybridization-based methods of using nucleic acid arrays (Heller *et al.*, 1997, Proc. Natl. Acad. Sci. U.S.A. 94:2150-2155; Lashkari *et al.*, 1997, Proc. Natl. Acad. Sci. U.S.A. 94:13057-13062; Wodicka *et al.*, 1997, Nature Biotechnol. 15:1259-1267). All such methods are encompassed by the present invention.

In one embodiment, subtractive hybridization is used to identify fetal cell tags. The principle of subtractive hybridization is that cDNAs common to both the target (*e.g.*, fetal) cells and reference (*e.g.*, maternal) cells are selected out by hybridizing to each other, leaving differentially expressed cDNA clones. See Wang *et al.*, 1991, Proc. Nat'l Acad. Sci USA 11505-11509. The subtractive hybridization method of Wang *et al.* removes commonly expressed cDNA from the experimental and control cDNA pools and thereby enriches for differentially expressed genes.

In another embodiment, one of a number of variations of differential display is used to identify fetal cell tags. See Liang *et al.*, 1992, Science 257:967; Liang *et al.*, 1995; Methods Enzymol. 254:304; U.S. Patent 5,262,311; U.S. Patent 5,599,672. Generally, Liang *et al.* describe a protocol which involves the reverse transcription of a messenger ribonucleic acid ("mRNA") population, in independent reactions, with each of twelve anchor primers (T_{12} MN), where M can be G (guanine), A (adenine) or C (cytosine) and N can be G, A, C or T (thymidine). The resulting single-stranded cDNAs are then amplified by the polymerase chain reaction (hereinafter, "PCR") using the same anchor primer used for reverse transcription together with an upstream or 5' decamer of arbitrary sequence. The PCR products, which are labeled by incorporation of tracer amounts of a radioactive nucleotide, are resolved for analysis by denaturating polyacrylamide gel electrophoresis (PAGE). This technique permits the simultaneous visualization of transcripts associated with the reference and target cells, *e.g.*, maternal and fetal cells. Liang *et al.* postulated that each two-primer combination could amplify only a limited subpopulation of cDNAs, and that the twelve anchor primers together with twenty arbitrary decamers (*i.e.*, 240 PCR reactions) should result in the display of the 3' termini of all

distinct mRNAs that are theoretically expressed in any given cell type (Liang and Pardee, 1992, Science 257:967-971). However, some of the genes identified, although useful for PCR-based identification of fetal cells, are below the limit of detection for *in situ* hybridization, which is a preferred method for identifying fetal cells according to this invention.

In yet another embodiment, fetal cell tags can be identified by combining subtractive hybridization and differential display. The combined methods involves subtractive hybridization followed by a differential display applied to the subtracted libraries.

In yet another embodiment, fetal cell tags can be identified by representational difference analysis of cDNA, which enriches for differences through rounds of subtraction and selective amplification.

In a preferred embodiment, suppression subtractive hybridization (SSH) is used to amplify candidate probe sequences. Suppression subtractive hybridization, which utilizes a combination of subtractive hybridization and polymerase chain reaction technology, is well known in the art and may even be performed using commercially available kits (Diatchenko *et al.*, 1996, Proc. Natl. Acad. Sci USA 93(12):6025-6030; PCR-select cDNA Subtraction Kit (Clontech), which is based on methods described in U.S. Patent No. 5,565,340). Generally, mRNA is isolated from the tissue or cell type which produces the tag sequences (*e.g.*, cell/tissue specific or selected mRNA's), then converted into cDNA using any convenient method for production of double-stranded cDNA. cDNA (or a portion of the cDNA) from the tissue or cell type which produces the probe sequences ("tester cDNA") is digested with a restriction endonuclease to produce appropriate 'sticky ends' (single stranded overhangs to which other nucleic acids, such as adaptors, may be annealed), split into two portions (a "first portion" and a "second portion"), and the two portions are modified by the addition of different adaptors of known sequence (*e.g.*, the first portion is modified by addition of a first adaptor and the second portion is modified by the addition of a second, different adaptor). "Driver cDNA" prepared from "reference cells" (*e.g.*, cells which should not be identified by the probe sequence, such as maternal cells in a maternal blood sample) is separately mixed with the modified first and second portions of tester cDNA, and each mixture is denatured and allowed to anneal. Each resulting mixture contains single-stranded tester cDNA, homoduplex tester cDNA, heteroduplex tester/driver cDNA, single stranded driver cDNA, and homoduplex driver cDNA. These mixtures are combined, along with an additional portion of denatured driver cDNA, and allowed to anneal, creating a complex mixture comprising single stranded tester cDNA and portion 1 and portion 2 tester cDNA, homoduplex driver cDNA and portion 1 and portion 2 tester cDNA, heteroduplex portion 1 tester/driver cDNA, heteroduplex portion 2 tester /driver cDNA and heteroduplex portion 1/portion 2 tester cDNA. The ends of the duplex cDNA's

are "filled in" using a template-driven reaction (*e.g.*, using DNA polymerase), then amplified using a template-driven amplification process such as the polymerase chain reaction and two primers, a first primer which will anneal to the first adaptor, and a second primer which will anneal to the second adaptor. Only heteroduplex portion 1/portion 2 tester cDNA will be geometrically amplified by the amplification reaction. The end result of SSH is a population of amplified sequences which are derived from RNAs more prevalent in the tester sample than the driver sample.

The SSH process may be reiterated using a different driver cDNA. Reiteration of the SSH process simply requires that the amplified product from the previous round of SSH be digested with a restriction enzyme to produce appropriate sticky ends to the amplified double stranded DNA, preferably using the same restriction endonuclease as in the previous round(s) of SSH. The digested DNA is then split into two portions, modified by the separate addition of different adaptors, and processed. SSH may be reiterated as many times as desired, with any number of driver cDNA samples.

Driver cDNA may be prepared from a variety of sources, including, but not limited to, samples of adult myeloid cells likely to contain a proportion of nucleated erythroid cells, such as adult bone marrow, nucleated cells from adult peripheral blood, and the like. Tumor cells may also be used to prepare driver cDNA. Alternatively, driver cDNA can be prepared from fetal tissues more mature than the source of the tester cDNA, such as fetal liver from later stages gestation.

Preferably, dispersed cells are used for preparation of tester and driver cDNA, although tissues may be used as well. More preferably, cells which have been subfractionated using physicochemical separation techniques or immunoadsorption techniques are utilized for cDNA preparation, particularly for tester cDNA preparation. The cells may also optionally be cultured, although this is generally not recommended, since the expression pattern of the cells may change as a result.

The product of the SSH reaction can be cloned into a suitable vector, thereby constituting a subtracted library from which individual candidate cDNA can be regenerated and validated. The use of the SSH technique permits the preparation of libraries corresponding to a number of fetal cell/reference cell (tester/driver) combinations to determine which combinations of cell types and collection times yield the richest proportion of valid clones.

5.1.2. VALIDATION

Described herein are non-limiting examples of validation steps that can be performed on fetal cell associated tags identified by the differential expression analysis methods of the invention. While each of these steps is optional, it is recommended that candidate sequences be evaluated by as many of these criteria as possible. The steps can be

performed in any order desired. They are generally listed in order of increasing difficulty or rarity of reagents, and it is generally convenient to perform the steps roughly in the order indicated.

1. *PRELIMINARY SEQUENCING.* The insert from each randomly selected cDNA clone is PCR amplified, and single-run sequencing of 50-200 nucleotides is performed. The sequence is then compaired against those available in public databases such as GenBank. It is recommended that this be done early in the validation process, to eliminate housekeeping genes, mRNA known to be generously expressed in adult blood cells, and redundant clones. If the sequence contains a single, clear open reading frame, then the orientation of the clone can also be predicted.

2. *INITIAL EXPRESSION SCREENING.* The cloned DNA is tested in blot analysis of expression patterns in an initial screening panel of fetal and adult cells. It is recommended that this be done using a Southern hybridization technique, using whole cDNA prepared either from mRNA or total RNA of the cells (a "virtual Northern"). The use of cDNA provides a renewable source of material for screening a number of clones. For example, the initial screening panel could include as positives: several fetal blood and fetal erythroid cell samples from fetal liver; and as negatives, adult and fetal liver parenchymal cells and several adult bone marrow cells.

3. *EXPANDED EXPRESSION SCREENING.* The cloned DNA is then tested for expression patterns using an expanded cell panel: for example, at least five fetal blood and erythroid cells taken at different stages of gestation, and at least three bone marrow and three peripheral blood samples from different adults. Preferred clones show at least 5 times and preferably at least 25 times the expression in the positive samples compared with the negative samples.

4. *NORTHERN ANALYSIS.* Cloned DNA that pass the preceding expression pattern analysis are preferably retested using mRNA from selected cell populations to verify that the DNA-RNA hybrids form with sufficient specificity to distinguish between the cell populations as a whole. Preferred clones show at least 5 times and preferably at least 25 times the expression in the positive samples compared with the negative samples. Information as to the size of the message and possible alternative splicing may also be obtained. Blots can be stripped and reused for testing of subsequent DNA clones. The blots can also be probed using DNA for housekeeping genes such as GADPH and ß-actin, or previously characterized sequences such as transferrin and γ-globin. This permits early elimination of DNA clones hybridizing to transcripts with the same size profile.

5. *ORIENTATION AND ABUNDANCE ANALYSIS.* Where the DNA is intended to specify fetal cells by hybridizing with mRNA *in situ*, the correct hybridizing strand should be identified. Orientation analysis is performed by Northern analysis using DNA from the cloned insert prepared as an asymmetric single-stranded probe. Abundance is determined by titration experiments using suitable standards, as are known in the art. The transcript should not only be specific for the desired cell type, it should be sufficiently abundant to provide ready detection of the specified cell according to the intended method.

6. *IN SITU mRNA HYBRIDIZATION.* Testing for probe sequences intended for *in situ* hybridization typically includes positive and negative screening using defined cell populations. Positive cell populations for fetal erythroid probes include nucleated erythroid cells from fetal liver, and cultured or uncultured cord blood cells. Positive cells for trophoblast probes are included in cell populations obtained from term placenta and chorionic villae. Negative cell populations include adult peripheral blood myeloid cells and bone marrow cells. Cells of interest in both positive and negative populations are either enriched or counterstained using specific antibody for an important phenotypic marker, such as those described earlier. Preferred DNA probe sequences have a relative rate of true positive to false positive identification of individual cells (estimated from the degree of enrichment of the cell population or the counterstaining) of about 10, 30, or 100 in order of increasing preference. The *in situ* hybridization analysis will also provide data on the intracellular distribution of the hybridizing transcript. A broad and abundant distribution facilitates most types of subsequent testing. At this stage, technical aspects of hybridization can also be refined; such as the agents used for cell attachment, fixation, and permeabilization; and the labeling, detection or signal amplification methods. Probe specificity can be confirmed by pre-treating the cells with RNAse, and by parallel probing with control sequences.

7. *INDEPENDENT CONFIRMATION BY RT-PCR.* Wherever possible, it is important to confirm that fetal cells express the probe sequence using a method other than a solid-phase hybridization assay (such as Northern blotting or *in situ* hybridization). In this test, PCR amplification is conducted using erythroid cells immunoaffinity enriched from fetal liver, or immunoaffinity enriched syncytiotrophoblasts, depending on the nature of the probe. RNA from the cells is reverse-transcribed and used as a template in PCR amplification. Two primers, based on segments of the probe that are about 100-200 base pairs apart, are used in the reaction. PCR amplification is then conducted, and the rate of amplification is determined (measured as the amount of PCR product formed of the correct size after a certain number of cycles). Compared with cells enriched from adult bone

marrow or peripheral blood, the rate of amplification is typically at least about 10, 30, or 100 times higher.

8. *DIAGNOSTIC TEST RUNS.* Model diagnostic analysis is conducted using spiked adult blood samples. Fresh peripheral blood is combined with either blood cells obtained from fetal liver, erythroid cells purified from fetal liver, cultured or uncultured cord blood cells (preferably from about 12 weeks gestation), or cytotrophoblast cells. The fetal cells are from a male fetus and added to an adult female blood sample, so that the Y chromosome can be used to follow specificity. The combined blood sample is then processed by all the intended steps leading up to hybridization with the probe, including density gradient separation and immunoaffinity enrichment. The cells are then processed with the probe according to the intended method to obtain validation. Where the probe is intended for *in situ* hybridization, the cells are processed accordingly, probed with the probe, and then counterstained for X and Y chromosomal markers. The count of X and Y chromosomes in each cell can also be determined by MGG/benzidine staining. Validation is obtained if the probe correctly distinguishes the fetal cells in the field from the adult blood cells. Similar experiments are then conducted using actual maternal blood samples taken from women carrying a single male fetus. In order of increasing preference, the probe sequence identifies at least about 25%, 50%, 75%, 90%, or 95% of fetal cells in the field. In order of increasing preference, the relative rate of true positive to false positive identification of individual cells (based on Y chromosome counterstaining) is 3, 10, 30, or 100.

5.2. NUCLEIC ACIDS OF THE INVENTION

The present invention provides nucleic acids that relate to the nucleic acids of the invention, *i.e.*, the nucleic acids of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42. Fragments, partially identical homologs, and longer nucleic acids including such sequences are included in the invention. Nucleic acids encoding the polypeptide translation products of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42, and their fragments and derivatives, are also provided.

The nucleic acids of the invention encompass adaptations of the fetal cell associated sequences, particularly those adaptations that facilitate their use in the separation and identification methods described in this disclosure. Additions, deletions, and substitutions of residues can be made for any worthwhile purpose, such as enhancing stability of hybrids formed with the target sequence, adapting towards a consensus of sequence variants, and decreasing cross-reactivity with sequences present in maternal cells.

Nucleic acid analogs of this invention include backbone chemistry not found in naturally occurring nucleic acids that improves stability or shelf-life. Labels or moieties for subsequently attaching labels can be attached or inserted into the sequence at any point that does not disturb the desired specificity.

The nucleic acids of the invention, especially those of about 50 nucleotides in length or less, can be conveniently prepared from the sequence data provided in this disclosure by chemical synthesis. Several methods of synthesis are known in the art, including the triester method and the phosphite method. In a preferred method, nucleic acids are prepared by solid-phase synthesis using mononucleoside phosphoramidite coupling units. See, for example, Beaucage *et al.*, 1981, Tetra. Lett. 22:1859; Kumar *et al.*, J. Org. Chem. 49:4905, and U.S. Patent No. 4,415,732.

Longer nucleic acids can also be prepared by chemical synthesis, but are more typically prepared by amplification or replication techniques. For example, nucleic acids can be amplified by PCR from RNA obtained from fetal tissue or cord blood cells, or from a cDNA library prepared from such tissue. Alternatively, nucleic acids can be amplified by PCR from human genomic DNA libraries. Nucleic acids prepared by any of these methods can be further replicated to provide a larger supply by any standard technique, such as by PCR amplification or molecular cloning.

One aspect of the invention pertains to isolated nucleic acid molecules that encode a polypeptide of the invention or a biologically active portion thereof, as well as nucleic acid molecules sufficient for use as hybridization probes to identify nucleic acid molecules encoding a polypeptide of the invention and fragments of such nucleic acid molecules suitable for use as PCR primers for the amplification or mutation of nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded.

An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid molecule. Preferably, an "isolated" nucleic acid molecule is free of sequences (preferably protein encoding sequences) which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kB, 4 kB, 3 kB, 2 kB, 1 kB, 0.5 kB or 0.1 kB of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free, *e.g.*, at least 60%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, free of other cellular material, or culture medium

when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. As used herein, the term "isolated"when referring to a nucleic acid molecule does not include an isolated chromosome.

In instances wherein the nucleic acid molecule is a cDNA or RNA, *e.g.*, mRNA, molecule, such molecules can include a poly A "tail", or, alternatively, can lack such a 3' tail. Although cDNA or RNA nucleotide sequences may be depicted herein with such tail sequences, it is to be understood that cDNA nucleic acid molecules of the invention are also intended to include such sequences lacking the depicted poly A tails. Where a nucleic acid molecule of the invention is used as a probe, it is preferred the that the probe lacks the polyA tails.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having all or a portion of the nucleotide sequence of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42, or a complement thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequences of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42 as a hybridization probe, nucleic acid molecules of the invention can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, eds., *Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

A nucleic acid molecule of the invention can be amplified using cDNA, mRNA or genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to all or a portion of a nucleic acid molecule of the invention can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42, or a portion thereof.

Moreover, a nucleic acid molecule of the invention can comprise only a portion of a nucleic acid sequence or SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42, a fragment which can be used as a probe or primer (*e.g.*, as described in Section 5.3) or a fragment encoding a biologically active portion of a polypeptide of the invention (*e.g.*, as described in Section 5.4).

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42 due to degeneracy of the genetic code and thus encode the same protein as that encoded by the nucleotide sequence of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42.

In addition to the nucleotide sequences of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 40, 41 or 42, it will be appreciated by those skilled in the art that DNA sequence polymorphism, including silent polymorphisms and those that lead to changes in the amino acid sequence may exist within a population (*e.g.*, the human population). Such genetic polymorphisms may exist among individuals within a population due to natural allelic variation. An allele is one of a group of genes which occur alternatively at a given genetic locus. As used herein, the phrase "allelic variant" refers to a nucleotide sequence which occurs at a given locus. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of a given gene. Alternative alleles can be identified by sequencing the gene or its corresponding mRNA of interest in a number of different individuals. This can be readily carried out by using hybridization probes to identify the same genetic locus in a variety of individuals. Any and all such nucleotide variations and resulting amino acid polymorphisms or variations that are the result of natural allelic variation and that do not alter the properties of the nucleic acids (*e.g.*, ability to hybridize to a fetal cell associated RNA) are intended to be within the scope of the invention.

In various embodiment of the present invention, an isolated nucleic acid molecule of the invention is at least 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900 or 2000 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42, or a complement thereof.

Accordingly, in other embodiments, an isolated nucleic acid molecule of the invention is at least 50, 100, 200, 300, 400, 500, 600, 700, 800 or 900 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42, or a complement thereof.

As used herein, the term "hybridizes under stringent conditions" means conditions for hybridization and washing under which nucleotide sequences at least 70% identical to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular*

Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6, which is incorporated by reference here in its entirety. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45° C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65° C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42, or a complement thereof, corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature.

In addition to naturally-occurring allelic variants of a nucleic acid molecule of the invention sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation thereby leading to changes in the amino acid sequence of the encoded protein, without altering the biological activity of the protein. For example, one can make nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are not conserved or only semi-conserved among homologues of various species may be non-essential for activity and thus would be likely targets for alteration. Alternatively, amino acid residues that are conserved among the homologues of various species may be essential for activity and thus would not be likely targets for alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding a polypeptide of the invention that contain changes in amino acid residues that are not essential for activity. Such polypeptides differ in amino acid sequence from any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78 yet retain biological activity. In one embodiment, the isolated nucleic acid molecule includes a nucleotide sequence encoding a protein that includes an amino acid sequence that is at least about 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 75%, 85%, 95%, or 98% identical to the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, and 78.

An isolated nucleic acid molecule encoding a variant protein can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 or 42 such that

one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Alternatively, mutations can be introduced randomly along all or part of the coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for biological activity to identify mutants that retain activity. A most preferred biological activity for the purposes of the present invention is antigenicity or immunogenicity. Following mutagenesis, the encoded protein can be expressed recombinantly and the activity of the protein, *i.e.*, the ability of be bound by an antibody against the non-mutant protein, can be determined.

The present invention encompasses antisense nucleic acid molecules, *i.e.*, molecules which are complementary to a sense nucleic acid encoding a polypeptide of the invention, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire coding strand, or to only a portion thereof, *e.g.*, all or part of the protein coding region (or open reading frame). An antisense nucleic acid molecule can be antisense to all or part of a non-coding region of the coding strand of a nucleotide sequence encoding a polypeptide of the invention. The non-coding regions ("5' and 3' untranslated regions") are the 5' and 3' sequences which flank the coding region and are not translated into amino acids.

An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides or more in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be

used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a selected polypeptide of the invention to thereby inhibit expression, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

An antisense nucleic acid molecule of the invention can be an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier et al. (1987) *Nucleic Acids Res.*

15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. (1987) *FEBS Lett.* 215:327-330).

The invention also encompasses ribozymes. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of the protein encoded by the mRNA. A ribozyme having specificity for a nucleic acid molecule encoding a polypeptide of the invention can be designed based upon the nucleotide sequence of a cDNA disclosed herein. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742. Alternatively, an mRNA encoding a polypeptide of the invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. *See, e.g.*, Bartel and Szostak (1993) *Science* 261:1411-1418.

The invention also encompasses nucleic acid molecules which form triple helical structures. For example, expression of a polypeptide of the invention can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the gene encoding the polypeptide (*e.g.*, the promoter and/or enhancer) to form triple helical structures that prevent transcription of the gene in target cells. *See generally* Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14(12):807-15.

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

PNAs can be used in diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression

by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. In a more preferred embodiment, PNAs are used for fetal cell detection and diagnosis, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup (1996), *supra*; or as probes or primers for DNA sequence and hybridization (Hyrup (1996), *supra*; Perry-O'Keefe et al. (1996) *Proc. Natl. Acad. Sci. USA* 93: 14670-675).

In another embodiment, PNAs can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNAse H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996), *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996), *supra,* and Finn et al. (1996) *Nucleic Acids Res.* 24(17):3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs. Compounds such as 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite can be used as a link between the PNA and the 5' end of DNA (Mag et al. (1989) *Nucleic Acids Res.* 17:5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) *Nucleic Acids Res.* 24(17):3357-63). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser et al. (1975) *Bioorganic Med. Chem. Lett.* 5:1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (*see, e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W0 88/09810) or the blood-brain barrier (*see, e.g.*, PCT Publication No. W0 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (*see, e.g.,* Krol et al. (1988) *Bio/Techniques* 6:958-976) or intercalating agents (*see, e.g.,* Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

5.3. PROBES OF THE INVENTION

Probe sequences of this invention include those that hybridize with an encoding region or a non-encoding region of the transcript, or span both. Non-encoding regions in some instances are preferred, since they are generally less functionally constrained and less likely to cross-hybridize with other targets. In addition, they may be part of a splice variant which is tissue specific.

The probes of the invention reliably distinguish human fetal cells in a majority of random maternal blood samples. In a single maternal blood sample, optionally enriched using an antibody specific to a fetal cell specific or selective antigen), the probe sequences generally identify at least 25%, and in order of increasing preference, identify about 50%, 75%, 90%, or 95% of fetal cells of a particular phenotype (such as erythroid cells or trophoblast cells) in the population. In a panel of maternal blood samples of mixed ethnic heritage, the relative rate of true positive to false positive identification of individual cells (fetal versus maternal cells identified) is generally at least 3, and is more typically 10, 30, or 100 in order of increased preference.

Preferred probe sequences are those that have minimal cross-reactivity with cells that are abnormally present in certain maternal blood samples due to a disease condition. Relevant diseases include cancer (particularly leukemias, lymphomas, and other myeloid or lymphoid malignancies, and certain endothelial cell and other malignancies that result in sluffing of malignant cells into the circulation), and hemoglobin abnormalities.

The target sequence is described as being "prominent" or "preferentially detected" in fetal cells compared with other cells in the mixed population, if the level of detection (according to the method used) is typically at least 5 times higher, more preferably at least about 25 times higher, and even more preferably at least about 100 times higher than other cells in the population, such as maternal cells of a similar phenotype. Low levels of expression of the sequence are acceptable in maternal cells, as long as the quantitative difference is sufficiently large and consistent to provide a reliable test according to the detection method used. In addition, certain preferred probe sequences have one or more of the properties described in the validation tests of Section 5.1.2.

Of special interest are probes that contain a sequence of consecutive nucleotides that is at least partly identical to a sequence in one of fetal cell associated RNAs of the invention. The length of consecutive nucleotides is generally at least 10 nucleotides, and may be 15, 25, 30, 40, 50, 70, 100, or 200 nucleotides in length. The degree of identity between the region of the probe that corresponds to a nucleic acid of the invention is typically at least 50%, and may be about 70%, 80%, 90%, 95% or 100%. The degree of identity between the region of the probe that corresponds to the fetal cell associated RNA and the corresponding region of the fetal cell associated RNA is typically at least 50%, and may be about 70%, 80%, 90%, 95% or 100%.

One of skill in the art will appreciate that nucleic acids with a longer matching sequence are preferred as more likely to distinguish the target sequence. Longer sequences can be incorporated with more labeling moieties per strand, and need not be as closely identical to the target in order to uniquely identify it. However, shorter sequences generally provide more tissue penetration and more rapid hybridization kinetics. Preferred hybridization probes are 10 to 200 nucleotides in length, more preferably 25 to 100 nucleotides in length. To combine the advantages of a long probe sequence with multiple labeling moieties and the efficiency of shorter-length probes, the probe sequence can be subdivided into nucleic acids of about 25 to 100 residues in length, provided as a reagent mixture. Thus, in certain embodiments of the invention, for a given fetal cell detection assay, a biological sample such as a maternal blood sample is contacted with multiple probes, *e.g.*, of 25-100 nucleotides in length. In one embodiment, the multiple probes comprise nucleic acid sequences that correspond to RNAs transcribed from one gene. The multiple probes can be designed to hybridize to one or more alternative splice forms of the same transcript. In another embodiment, the multiple probes comprise nucleic acids sequences that correspond to RNAs transcribed from more than one gene.

Preferred oligonucleotide probes for use as PCR primers are preferably 10 to 100 nucleotides in length and more typically 15 to 50 nucleotides in length. Individual primers may not necessarily hybridize with unique nucleic acid sequences on the target, and yet still be capable of specifically amplifying a unique sequence when used with a second primer, or when used in a nested amplification reaction with still other primers.
The probe/primer typically comprises substantially purified oligonucleotide. In one embodiment, the oligonucleotide comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 30, 40, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 or 400 consecutive nucleotides of the sense or anti-sense sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 and 42 or of a naturally occurring mutant of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 and 42.
In another embodiment, the oligonucleotide comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least 400, preferably 450, 500, 530, 550, 600, 700, 800, 900, 1000 or 1150 consecutive oligonucleotides of the sense or antisense sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 and 42 or of a naturally occurring mutant of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 and 42.

The fetal cell probes of the invention may additionally comprise features of the antisense nucleic acid molecules and PNAs described in Section 5.2, *supra*, including

but not limited the inclusion of modified nucleotide residues that impart greater stability on the probe-fetal cell RNA hybrids formed when performing fetal cell detection and/or diagnosis.

Probes for detection of a fetal cell preferably comprise a label group incorporated or attached thereto, *e.g.*, a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor.

5.4. POLYPEPTIDES OF THE INVENTION

In addition, to the foregoing nucleic acids, the present invention provide polypeptides encoded by the nucleic acids of the invention. The nucleic acid sequences provide a gateway for analyzing polypeptides encoded by the fetal cell associated nucleic acids. Since the target transcript is preferentially expressed in fetal cells, the polypeptide product is expected to have a similar expression pattern, and may also serve as a marker for fetal cells. Of particular interest are polypeptide products predicted to contain a membrane spanning region, since they are more likely to be expressed at the cell surface. Also of interest are polypeptide products with enzymatic activity, especially for production of a cell-surface marker, or for conversion of a chromogenic substrate. Epitope containing amino acid sequences from the encoding region that are preferably 10, 15, 25, 50 or greater residues in length can also be used to elicit and select specific antibody according to the general methods provided elsewhere in this disclosure. In turn, these antibodies can be used for fetal cell detection or immunoenrichment from a mixed cell population by contacting with the cells under conditions that permit the antibody to bind to the expressed antigen. Although not all nucleic acid molecules of the invention encode a full open reading frame, including a start and stop codon, one skill in the art would recognize that the encoded polypeptides can be recombinantly expressed by inserting the nucleic acid into the appropriate vector with start, stop, and/or translation initiation signals; additionally, such sequences can be encoded by a fusion partner if a polypeptide of the invention is to be expressed in the form of a fusion protein. The following table indicates which polypeptide SEQ ID NOs. correspond to which nucleic SEQ ID NOs. of the invention; where no SEQ ID NO. is given for a corresponding polypeptide is indicative that the nucleic acid in question comprises largely or solely noncoding sequences:

Fetal Cell Specific Transcript Name	Corresponding Nucleic Acid SEQ ID NO.	Corresponding Polypeptide SEQ ID NO:
1503-7E (*tag*)	10	43
J42-4d (*FL*)	11	44

	J2r(3) (*ASF*)	12	-
	J2r(12) (*ASF*)	13	45
	J2r(13) (*ASF*)	14	46
	305-4G (*tag*)	15	47
5	K1-1a (*FL*)	16	48
	K2r/1f(50) (*ASF*)	17	49, 50, 51
	K2r/1f(59) (*ASF*)	18	52
	K(1)157-2A (*ASF*)	19	53
	K3r(HIGH)76 (*ASF*)	20	-
10			
	597-10C (*tag*)	21	54
	NT7-T3 (*FL*)	22	55
	N9r/Mf (*ASF*)	23	56
15	334-2C (*tag*)	24	57, 58
	O19r-T3 (*FL*)	25	59, 60, 61
	O1-1a (*ASF*)	26	62, 63
	332-9E (*tag*)	27	-
20	P60-1a (*FL*)	28	64, 65, 66
	P1-1a (*ASF*)	29	67
	P3r(9) (*ASF*)	30	68
	305-9E (*tag*)	31	69
25	R5'-T3 (*FL*)	32	70
	R6r/1-6H (*ASF*)	33	71
	369-8G (*tag*)	34	
	U2f-T3 (*FL*)	35	72
30			

	305-6G (*tag*)	36	73
	L15-1a (*FL*)	37	74
	L21-1a	38	75
	252	39	-
	120r	40	76
	Clone-1	41	77
	D19-2g	42	78

One aspect of the invention pertains to isolated polypeptides, and biologically active portions thereof, including but not limited to polypeptide fragments suitable for use as immunogens to raise antibodies directed against a polypeptide of the invention. In one embodiment, the native polypeptide can be isolated from cells or tissue sources by an appropriate purification scheme using standard polypeptide purification techniques. In another embodiment, polypeptides of the invention are produced by recombinant DNA techniques. Alternative to recombinant expression, a polypeptide of the invention can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" polypeptide or biologically active portion thereof is substantially free of cellular material or other contaminating polypeptides from the cell or tissue source from which the polypeptide is derived, or substantially free of chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of polypeptide in which the polypeptide is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, polypeptide that is substantially free of cellular material includes preparations of polypeptide having less than about 30%, 25%, 20%, 15%, 10%, 5%, 2% or 1% (by dry weight) of heterologous polypeptide (also referred to herein as a "contaminating polypeptide"). When the polypeptide or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 25%, 20%, 15%, 10%, 5%, 2% or 1% of the volume of the polypeptide preparation. When the polypeptide is produced by chemical synthesis, it is preferably substantially free of chemical precursors or other chemicals, *i.e.*, it is separated from chemical precursors or other chemicals which are involved in the synthesis of the

polypeptide. Accordingly such preparations of the polypeptide have less than about 30%, 25%, 20%, 15%, 10%, 5%, 2% or 1% (by dry weight) of chemical precursors or compounds other than the polypeptide of interest.

Biologically active portions of a polypeptide of the invention include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the polypeptide (*e.g.*, the amino acid sequence shown in any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, and 78), which include fewer amino acids than the full length polypeptide, and exhibit at least one activity of the corresponding full-length polypeptide. Typically, biologically active portions comprise a domain or motif with at least one activity of the corresponding polypeptide. A biologically active portion of a polypeptide of the invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover, other biologically active portions, in which other regions of the polypeptide are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of the native form of a polypeptide of the invention.

Preferred polypeptides have the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, and 78. Other useful polypeptides are substantially identical (*e.g.*, at least about 45%, preferably 55%, 65%, 75%, 85%, 95%, or 99%) to any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, and 78, and retain the functional activity of the polypeptide of the corresponding naturally-occurring polypeptide yet differ in amino acid sequence due to natural allelic variation or mutagenesis.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (, % identity = # of identical positions/total # of positions (*e.g.*, overlapping positions) x 100). In one embodiment, the two sequences are the same length.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a

mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the invention. BLAST polypeptide searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a polypeptide molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules (*Id.*). When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. *See* http://www.ncbi.nlm.nih.gov.

Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, CABIOS (1989). Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the CGC sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Additional algorithms for sequence analysis are known in the art and include ADVANCE and ADAM as described in Torellis and Robotti (1994) *Comput. Appl. Biosci., 10*:3-5; and FASTA described in Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. 85*:2444-8. Within FASTA, ktup is a control option that sets the sensitivity and speed of the search. If ktup=2, similar regions in the two sequences being compared are found by looking at pairs of aligned residues; if ktup=1, single aligned amino acids are examined. ktup can be set to 2 or 1 for polypeptide sequences, or from 1 to 6 for DNA sequences. The default if ktup is not specified is 2 for polypeptides and 6 for DNA. For a further description of FASTA parameters, see http://bioweb.pasteur.fr/docs/man/man/fasta.1.html#sect2, the contents of which are incorporated herein by reference.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, typically exact matches are counted.

The invention also provides chimeric or fusion polypeptides. As used herein, a "chimeric polypeptide" or "fusion polypeptide" comprises all or part (preferably biologically active) of a polypeptide of the invention operably linked to a heterologous polypeptide (*i.e.*, a polypeptide other than the same polypeptide of the invention). Within the fusion polypeptide, the term "operably linked" is intended to indicate that the

polypeptide of the invention and the heterologous polypeptide are fused in-frame to each other. The heterologous polypeptide can be fused to the N-terminus or C-terminus of the polypeptide of the invention.

One useful fusion polypeptide is a GST fusion polypeptide in which the polypeptide of the invention is fused to the C-terminus of GST sequences. Such fusion polypeptides can facilitate the purification of a recombinant polypeptide of the invention.

In another embodiment, the fusion polypeptide contains a heterologous signal sequence at its N-terminus. For example, if a polypeptide of the invention comprises a signal sequence, the native signal sequence can be removed and replaced with a signal sequence from another polypeptide. For example, the gp67 secretory sequence of the baculovirus envelope protein can be used as a heterologous signal sequence (*Current Protocols in Molecular Biology*, Ausubel et al., eds., John Wiley & Sons, 1992). Other examples of eukaryotic heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase (Stratagene; La Jolla, California). In yet another example, useful prokaryotic heterologous signal sequences include the phoA secretory signal (Sambrook et al., *supra*) and the protein A secretory signal (Pharmacia Biotech; Piscataway, New Jersey), which may be useful for recombinant expression and/or purification of the polypeptide of the invention.

In yet another embodiment, the fusion polypeptide is an immunoglobulin fusion polypeptide in which all or part of a polypeptide of the invention is fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion polypeptides of the invention can be used as immunogens to produce antibodies directed against a polypeptide of the invention.

Chimeric and fusion polypeptides of the invention can be produced by standard recombinant DNA techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (*see, e.g.,* Ausubel et al., *supra*). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the polypeptide of the invention.

The present invention also pertains to variants of the polypeptides of the invention. Variants can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation. An agonist can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of the polypeptide. An antagonist of a polypeptide can inhibit one or more of the activities of the naturally occurring form of the polypeptide

by, for example, competitively binding to an antibody the binds to the native polypeptide of interest.

The polypeptides of the invention can be modified to exhibit reduced or increased post-translational modifications, including, but not limited to glycosylations, (*e.g.*, N-linked or O-linked glycosylations), myristylations, palmitylations, acetylations and phosphorylations (*e.g.*, serine/threonine or tyrosine).

5.5. ANTIBODIES OF THE INVENTION

The present invention further encompasses the use of antibodies that bind to the polypeptides of the invention for fetal cell detection and diagnostics. The antibodies are preferably monoclonal, and may be multispecific, human, humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab') fragments, fragments produced by a Fab expression library, and binding fragments of any of the above. The term "antibody," as used herein, refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site that immunospecifically binds a polypeptide of the invention. The immunoglobulin molecules of the invention can be of any type (*e.g.*, IgG, IgE, IgM, IgD, IgA and IgY), class (*e.g.*, IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2) or subclass of immunoglobulin molecule.

An isolated polypeptide of the invention, or a fragment thereof, can be used as an immunogen to generate antibodies using standard techniques for polyclonal and monoclonal antibody preparation. The full-length polypeptide can be used or, alternatively, the invention provides antigenic peptide fragments for use as immunogens. The antigenic peptide of a polypeptide of the invention comprises at least 8 (preferably 10, 15, 20, or 30) amino acid residues of the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, and 78, and encompasses an epitope of the polypeptide such that an antibody raised against the peptide forms a specific immune complex with the polypeptide.

Preferred epitopes encompassed by the antigenic peptide are regions that are located on the surface of the polypeptide, *e.g.*, hydrophilic regions. Hydrophobic regions can be identified by hydropathy plots.

An immunogen typically is used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse or other mammal). An appropriate immunogenic preparation can contain, for example, recombinantly expressed or chemically synthesized polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent.

In certain embodiments of the invention, the antibodies are human antigen-binding antibody fragments of the present invention and include, but are not limited to, Fab,

Fab' and $F(ab')_2$, Fd, single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a V_L or V_H domain. Antigen-binding antibody fragments, including single-chain antibodies, may comprise the variable region(s) alone or in combination with the entirety or a portion of the following: hinge region, CH1, CH2, CH3 and CL domains. Also included in the invention are antigen-binding fragments also comprising any combination of variable region(s) with a hinge region, CH1, CH2, CH3 and CL domains. Preferably, the antibodies are human, murine (*e.g.*, mouse and rat), donkey, sheep, rabbit, goat, guinea pig, camelid, horse, or chicken. As used herein, "human" antibodies include antibodies having the amino acid sequence of a human immunoglobulin and include antibodies isolated from human immunoglobulin libraries, from human B cells, or from animals transgenic for one or more human immunoglobulin, as described *infra* and, for example in U.S. Patent No. 5,939,598 by Kucherlapati *et al.*

Antibodies that bind to the polypeptides of the invention may be monospecific, bispecific, trispecific or of greater multispecificity. Multispecific antibodies may be specific for different epitopes of a polypeptide of the invention or may be specific for both a polypeptide of the invention as well as for a heterologous polypeptide. *See, e.g.*, PCT publications WO 93/17715; WO 92/08802; WO 91/00360; WO 92/05793; Tutt, *et al.*, 1991, J. Immunol. 147:60-69; U.S. Patent Nos. 4,474,893; 4,714,681; 4,925,648; 5,573,920; 5,601,819; Kostelny *et al.*, 1992, J. Immunol. 148:1547-1553.

The present invention encompasses the use of derivatives of the antibodies of the invention that are modified, *i.e,* by the covalent attachment of any type of molecule to the antibody such that covalent attachment does not prevent the antibody from binding to a polypeptide of the invention. For example, but not by way of limitation, the antibody derivatives include antibodies that have been modified, *e.g.*, by glycosylation, acetylation, pegylation, phosphylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other polypeptide, *etc.* Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, metabolic synthesis of tunicamycin, *etc.* Additionally, the derivative may contain one or more non-classical amino acids.

The antibodies of the present invention may be generated by any suitable method known in the art. Polyclonal antibodies of the invention can be produced by various procedures well known in the art. For example, polypeptides of the invention can be administered to various host animals including, but not limited to, rabbits, mice, rats, *etc.* to induce the production of sera containing polyclonal antibodies specific for the polypeptide. Various adjuvants may be used to increase the immunological response, depending on the host species, and include but are not limited to, Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as

lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and corynebacterium parvum. Such adjuvants are also well known in the art.

Monoclonal antibodies can be prepared using a wide variety of techniques known in the art including the use of hybridoma, recombinant, and phage display technologies, or a combination thereof. For example, monoclonal antibodies can be produced using hybridoma techniques including those known in the art and taught, for example, in Harlow *et al.*, Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratory Press, 2nd ed., 1988); Hammerling, *et al.*, in: Monoclonal Antibodies and T-Cell Hybridomas 563-681 (Elsevier, N.Y., 1981) (said references incorporated by reference in their entireties). The term "monoclonal antibody" as used herein is not limited to antibodies produced through hybridoma technology. The term "monoclonal antibody" refers to an antibody that is derived from a single clone, including any eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced.

Methods for producing and screening for specific antibodies using hybridoma technology are routine and well known in the art. In a non-limiting example, mice can be immunized with a polypeptide of the invention or a cell expressing a polypeptide of the invention. Once an immune response is detected, *e.g.*, antibodies specific for the polypeptide of the invention are detected in the mouse serum, the mouse spleen is harvested and splenocytes isolated. The splenocytes are then fused by well known techniques to any suitable myeloma cells, for example cells from cell line SP20 available from the ATCC. Hybridomas are selected and cloned by limited dilution. The hybridoma clones are then assayed by methods known in the art for cells that secrete antibodies capable of binding the polypeptide of the invention. Ascites fluid, which generally contains high levels of antibodies, can be generated by injecting mice with positive hybridoma clones.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, Fab and $F(ab')_2$ fragments of the invention may be produced by proteolytic cleavage of immunoglobulin molecules, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce $F(ab')_2$ fragments). $F(ab')_2$ fragments contain the variable region, the light chain constant region and the CH1 domain of the heavy chain.

For example, the antibodies of the present invention can also be generated using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the nucleic acid sequences encoding them. In a particular embodiment, such phage can be utilized to display antigen binding domains expressed from a repertoire or combinatorial antibody library (*e.g.*, human or murine). In phage display methods, functional antibody

domains are displayed on the surface of phage particles which carry the nucleic acid sequences encoding them. In particular, DNA sequences encoding V_H and V_L domains are amplified from animal cDNA libraries (*e.g.*, human or murine cDNA libraries of lymphoid tissues). The DNA encoding the V_H and V_L domains are recombined together with an scFv linker by PCR and cloned into a phagemid vector (*e.g.*, p CANTAB 6 or pComb 3 HSS). The vector is electroporated in *E. coli* and the *E. coli* is infected with helper phage. Phage used in these methods are typically filamentous phage including fd and M13 binding domains expressed from phage with Fab, Fv or disulfide stabilized Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Phage expressing an antigen binding domain that binds to polypeptide of the invention or a binding portion thereof can be selected or identified with antigen *e.g.*, using labeled antigen or antigen bound or captured to a solid surface or bead. Examples of phage display methods that can be used to make the antibodies of the present invention include those disclosed in Brinkman *et al.*, 1995, J. Immunol. Methods 182:41-50; Ames *et al.*, 1995, J. Immunol. Methods 184:177-186; Kettleborough *et al.*, 1994, Eur. J. Immunol. 24:952-958; Persic *et al.*, 1997, Gene 187:9-18; Burton *et al.*, 1994, Advances in Immunology, 191-280; PCT Application No. PCT/GB91/O1 134; PCT Publications WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/1 1236; WO 95/15982; WO 95/20401; and U.S. Patent Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, *e.g.*, as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and $F(ab')_2$ fragments can also be employed using methods known in the art such as those disclosed in PCT publication WO 92/22324; Mullinax *et al.*, BioTechniques 1992, 12(6):864-869; and Sawai *et al.*, 1995, AJRI 34:26-34; and Better *et al.*, 1988, Science 240:1041-1043 (said references incorporated by reference in their entireties).

Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Patents 4,946,778 and 5,258,498; Huston *et al.*, 1991, Methods in Enzymology 203:46-88 ; Shu *et al.*, 1993, PNAS 90:7995-7999; and Skerra *et al.*, 1988, Science 240:1038-1040 . For some uses, including *in vivo* use of antibodies in humans and *in vitro* proliferation or cytotoxicity assays, it is preferable to use chimeric, humanized, or human antibodies. A chimeric antibody is a molecule in which different portions of the antibody are derived from different animal species, such as antibodies having a variable region derived from a murine monoclonal antibody and a

human immunoglobulin constant region. Methods for producing chimeric antibodies are known in the art. *See e.g.*, Morrison, Science, 1985, 229:1202; Oi *et al.*, 1986, BioTechniques 4:214; Gillies *et al.*, 1989, J. Immunol. Methods 125:191-202; U.S. Patent Nos. 5,807,715; 4,816,567; and 4,816,397, which are incorporated herein by reference in their entirety. Humanized antibodies are antibody molecules from non-human species antibody that binds the desired antigen having one or more CDRs from the non-human species and framework and constant regions from a human immunoglobulin molecule. Often, framework residues in the human framework regions will be substituted with the corresponding residue from the CDR donor antibody to alter, preferably improve, antigen binding. These framework substitutions are identified by methods well known in the art, *e.g.*, by modeling of the interactions of the CDR and framework residues to identify framework residues important for antigen binding and sequence comparison to identify unusual framework residues at particular positions. (*See, e.g.*, Queen *et al.*, U.S. Patent No. 5,585,089; Riechmann *et al.*, 1988, Nature 332:323 , which are incorporated herein by reference in their entireties.) Antibodies can be humanized using a variety of techniques known in the art including, for example, CDR-grafting (EP 239,400; PCT publication WO 91/09967; U.S. Patent Nos. 5,225,539; 5,530,101; and 5,585,089), veneering or resurfacing (EP 592,106; EP 519,596; Padlan, Molecular Immunology, 1991, 28(4/5):489-498; Studnicka *et al.*, 1994, Protein Engineering 7(6):805-814; Roguska. *et al.*, 1994, PNAS 91:969-973), and chain shuffling (U.S. Patent No. 5,565,332).

Completely human antibodies can be used in the present methods. Human antibodies can be made by a variety of methods known in the art including phage display methods described above using antibody libraries derived from human immunoglobulin sequences. See also, U.S. Patent Nos. 4,444,887 and 4,716,111; and PCT publications WO 98/46645, WO 98/50433, WO 98/24893, WO 98/16654, WO 96/34096, WO 96/33735, and WO 91/10741; each of which is incorporated herein by reference in its entirety.

Human antibodies can also be produced using transgenic mice which express human immunoglobulin genes. For example, the human heavy and light chain immunoglobulin gene complexes may be introduced randomly or by homologous recombination into mouse embryonic stem cells. The mouse heavy and light chain immunoglobulin genes may be rendered non-functional separately or simultaneously with the introduction of human immunoglobulin loci by homologous recombination. In particular, homozygous deletion of the JH region prevents endogenous antibody production. The modified embryonic stem cells are expanded and microinjected into blastocysts to produce chimeric mice. The chimeric mice are then bred to produce homozygous offspring which express human antibodies. The transgenic mice are immunized in the normal fashion with a selected antigen, *e.g.*, all or a portion of a polypeptide of the invention. Monoclonal antibodies directed against the antigen can be obtained from the immunized, transgenic

mice using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce useful IgG, IgA, IgM and IgE antibodies. For an overview of this technology for producing human antibodies, *see*, Lonberg and Huszar, 1995, Int. Rev. Immunol. 13:65-93. For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, *see, e.g.*, PCT publications WO 98/24893; WO 92/01047; WO 96/34096; WO 96/33735; European Patent No. 0 598 877; U.S. Patent Nos. 5,413,923; 5,625,126; 5,633,425; 5,569,825; 5;661,016; 5,545,806; 5,814,318; 5,885,793; 5,916,771; and 5,939,598, which are incorporated by reference herein in their entirety. In addition, companies such as Abgenix, Inc. (Freemont, CA) and Genpharm (San Jose, CA) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, *e.g.*, a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers *et al.*, 1994, Bio/technology 12:899-903).

Further, antibodies to the polypeptides of the invention can, in turn, be utilized to generate anti-idiotype antibodies that "mimic" the polypeptides of the invention using techniques well known to those skilled in the art. (*See, e.g.*, Greenspan & Bona, 1989, FASEB J. 7(5):437-444; and Nissinoff, 1991, J. Immunol. 147(8):2429-2438).

5.6. Recombinant Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a polypeptide of the invention (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors, expression vectors, are capable of directing the expression of genes to which they are operably linked. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids (vectors). However, the

invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell. This means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operably linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce polypeptides or peptides, including fusion polypeptides or peptides, encoded by nucleic acids as described herein.

The recombinant expression vectors of the invention can be designed for expression of a polypeptide of the invention in prokaryotic (*e.g.*, *E. coli*) or eukaryotic cells (*e.g.*, insect cells (using baculovirus expression vectors), yeast cells or mammalian cells). Suitable host cells are discussed further in Goeddel, *supra*. Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of polypeptides in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion polypeptides. Fusion vectors add a number of amino acids to a polypeptide encoded therein, usually to the amino terminus of the recombinant polypeptide. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant polypeptide; 2) to increase the solubility of the recombinant polypeptide; and 3) to aid in the purification of the recombinant polypeptide by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant polypeptide to enable separation of the

recombinant polypeptide from the fusion moiety subsequent to purification of the fusion polypeptide. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant polypeptide.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann et al., (1988) *Gene* 69:301-315) and pET 11d (Studier et al., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

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

In another embodiment, the expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari et al. (1987) *EMBO J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al. (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and pPicZ (Invitrogen Corp, San Diego, CA).

Alternatively, the expression vector is a baculovirus expression vector. Baculovirus vectors available for expression of polypeptides in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith et al. (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and pMT2PC (Kaufman et al. (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's

control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook et al., *supra*.

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

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operably linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to the mRNA encoding a polypeptide of the invention. Regulatory sequences operably linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub et al. (*Reviews - Trends in Genetics*, Vol. 1(1) 1986).

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such

terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic (*e.g.*, *E. coli*) or eukaryotic cell (*e.g.*, insect cells, yeast or mammalian cells).

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*supra*), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, for resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

In another embodiment, the expression characteristics of an endogenous (*e.g.*, J42-4d or D19-2g gene) within a cell, cell line or microorganism may be modified by inserting a DNA regulatory element heterologous to the endogenous gene of interest into the genome of a cell, stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous gene (*e.g.*, J42-4d or D19-2g gene) and controls, modulates or activates. For example, endogenous J42-4d or D19-2g genes which are normally "transcriptionally silent", *i.e.*, a J42-4d or D19-2g gene which is normally not expressed in maternal erythroid cell, or are expressed only at very low levels in a cell line or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, transcriptionally silent, endogenous J42-4d or D19-2g genes may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such that it is operatively linked with and activates expression of endogenous genes, using techniques, such as targeted homologous recombination, which

are well known to those of skill in the art, and described *e.g.*, in Chappel, U.S. Patent No. 5,272,071; PCT publication No. WO 91/06667, published May 16, 1991.

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce a polypeptide of the invention. Accordingly, the invention further provides methods for producing a polypeptide of the invention using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding a polypeptide of the invention has been introduced) in a suitable medium such that the polypeptide is produced. In another embodiment, the method further comprises isolating the polypeptide from the medium or the host cell.

5.7. METHODS FOR FETAL CELL ENRICHMENT

The invention provides methods of identifying and diagnosing fetal cells in a sample. Most commonly, the fetal cells will be present in a maternal blood sample, where the fetal cells comprise an extremely low percentage of the total cells present. As mentioned above, prior to carrying out the fetal cell detection and diagnosis methods of the present invention, it is preferable that the biological sample, *e.g.*, maternal blood sample, that will be subject to detection or diagnosis is enriched for rare fetal cells.

A mixed blood sample may be enriched for fetal cells by fractionation. Fractionation methods include density gradient centrifugation and differential lysis. For example, density gradients can be used to remove maternal red blood cells and lymphocytes (see, *e.g.*, Durrant *et al.*, 1996, Early Hum Dev 47 Suppl:S79-83). Similarly, maternal blood cells can be removed from a maternal blood sample by differential lysis of the maternal cells, as described by Furbetta *et al.*, 1980, Br J Haematol 44(3):441-50.

The foregoing fractionation techniques can be done as an alternative or, more preferably, in addition to the immunoenrichment methods described in Section 5.7.1 below.

5.7.1. IMMUNOENRICHMENT

In a preferred embodiment, the fetal cell enrichment step utilizes a fetal cell-associated or maternal cell-associated antibody (an "immunoenrichment" step).

Immunoenrichment can be positive immunoenrichment, whereby the mixed cell population of interest is contacted with a fetal cell-associated antibody and cells bound to the antibody are selected for. Preferably, the antibodies are directed to fetal blood cell associated antigens or trophoblast associated antigens. The antibodies are preferably specific or selective for antigens which are not found on maternal blood cells in a maternal blood sample. More preferably, the antibodies are specific or selective for a fetal blood cell-specific or trophoblast-specific antigen. One exemplary antibody for use in immunoenrichment is Clone 1 which is specific for a fetal erythroid cell antigen. An

antibody comprising the heavy chain or light chain of Clone 1, or one or more heavy or light chain CDRs of Clone 1, can also be used.

Immunoenrichment can also be negative immunoenrichment, whereby the mixed cell population of interest is contacted with a maternal cell-associated antibody and cells bound to the antibody are selected against. Preferably, the antibodies used in negative immunoenrichment have little to no binding to fetal erythroid and trophoblast cells.

Immunoenrichment may be carried out using any appropriate methodology known in the art. Preferred methods include fluorescence-activated cell sorting (FACS), immunomagnetic technologies, immunoprecipitation methods, and solid-phase separation methods (*e.g.*, panning). Generally, the antibody used for immunoenrichment is modified in a way that allows separation of cells with bound antibody from cells without bound antibody. In the case of FACS, the immunoenrichment antibody is labeled with a fluorescent compound (or secondarily labeled with a fluorescent compound), incubated with the maternal sample, then the cells of the maternal sample are separated into labeled and unlabeled fractions using an automated sorter. Immunoenrichment using immunomagnetic technology generally involves binding cells in the maternal sample with an immunoenrichment antibody primarily or secondarily labeled with paramagnetic or ferromagnetic particles, followed by separation of labeled cells with a magnetic field. Other useful techniques involve binding, adsorbing, or otherwise linking an immunoenrichment antibody to a solid phase, such as a bead or a plastic substrate, binding the antibody to cells in a maternal sample, and retaining bound cells on the basis of the properties of the solid phase (*e.g.*, by collection of beads). Preferred forms of immunoenrichment antibody include immunoenrichment antibodies which are haptenized, directly labeled with a fluorescent dye, adsorbed to magnetic particles, linked to a suspendable solid phase such as beads, or adsorbed to a solid phase such as a plastic dish.

Normally, the immunoenrichment step comprises incubating the sample containing the mixed cell population with an immunoenrichment antibody for a sufficient period of time to allow binding of the immunoenrichment antibody to target cells present in the sample. The incubation period is typically from about 10 or 15 minutes up to several hours. The incubation may be carried out at elevated temperatures (*e.g.*, about 30° to 37° C), at room temperature (RT, approximately 19° to 22° C) or at reduced temperatures (*e.g.*, from about 4° to about 15° C). The incubation is normally carried out in a physiologically-acceptable solution containing a pH buffer, salts, and optionally containing dextrose and/or blocking agents such as serum albumin, gelatin, and the like.

The steps following incubation of the maternal sample with the immunoenrichment antibody will depend on the immunoenrichment antibody and any modifications thereto, as will be apparent to one of skill in the art. Generally, the sample

will be processed to separate cells bound to the immunoenrichment antibody from cells not bound to the immunoenrichment antibody.

When magnetic particles are used, the sample is subjected to a magnetic field, which is generally oriented to segregate the antibody-bound cells to the wall of the incubation vessel, and the incubation solution, including any bound, is removed. Where positive immunoenrichment is used, *i.e.*, the antibody bound cells are the fetal cells of interest, the incubation solution is discarded. Where negative immunoenrichment is used, the incubation solution will contain the fetal cells of interest and antibody-bound cells are discarded. Similarly, for suspendible solid phase-based systems, the suspendible solid phase is allowed to settle. The supernatant, including unlabeled cells, is removed where a positive immunoenrichment antibody is used, and collected for further processing where a negative immunoenrichment antibody is used. Where the immunoenrichment antibody is adsorbed to a non-suspendible solid phase (*e.g.*, the bottom of a plastic dish), unbound cells and incubation solution are simply removed or collected, as desired. For positive immunoenrichment, the cells thus separated may be washed by simply resuspending the sample (where magnetic or suspendible solid phase technology is used), or simply adding additional buffer (where non-suspendible solid phase technology is used) and repeating the separation procedure. Preferably the separated cells are washed at least once.

When FACS is used for separating antibody-bound and non-bound cells, the cells are normally processed to render bound cells detectable, if the immunoenrichment antibody is not primarily labeled. For positive immunoenrichment, such processing generally involves washing away any unbound immunoenrichment antibody, then incubating with a detection reagent (*e.g.*, rhodamine-derivatized avidin for a biotinylated immunoenrichment antibody, or a labeled secondary antibody of appropriate specificity for an unmodified immunoenrichment antibody), washing again, then FACS processing.

Washing is typically carried out using the solution which was used for the antibody incubation, minus the antibody, although simple buffered solutions such as phosphate buffered saline (PBS) and tris buffered saline (TBS) may also be used. The cells of the maternal sample are typically washed several times, generally about 2-3, although a larger or smaller number of washes may be used as long as sufficient excess immunoenrichment antibody is removed and the cells are not unduly damaged by the washing procedure.

Optionally, additional rounds of immunoenrichment are utilized to further enrich the mixed cell population for fetal cells. The increase in enrichment of fetal cells in the maternal sample with additional rounds of immunoenrichment must be balanced against loss of cells due to processing during immunoenrichment. Preferably, immunoenrichment is performed in 1 to 4 rounds, 1 to 3 rounds, or 1 to 2 rounds. Where more than one round

of immunoenrichment is carried out, it is preferred that different immunoenrichment antibodies are used for each round of immunoenrichment.

Depending on the technology used for positive immunoenrichment, it may be desirable to "release" antibodies bound to the fetal cells enriched from the maternal sample after immunoenrichment is completed and/or between rounds of immunoenrichment. Release of bound antibodies is generally accomplished by altering pH or ionic conditions in the fluid medium.

5.7.2. METHODS FOR IDENTIFYING FETAL CELL ASSOCIATED ANTIBODIES

The selection step of the instant methods for enriching fetal cells in a sample prior to fetal cell detection and/or diagnosis is preferably performed using one or more antibodies specific for a specific or selective marker on the target cell. The instant invention provides antibodies specific for fetal erythroid specific markers and selective markers, which can be used in the selection step, as well as methods for isolating new target cell specific and selective antibodies.

The inventors have discovered that fetal liver is an excellent source for erythroblast cells expressing suitable markers for antibody development. Human fetal liver samples are preferably collected between about 10-18 weeks of gestation, taking care to chill the tissue very shortly after harvesting. Preferably, the fetal livers experience no more than 15 minutes of warm hypoxia. The fetal livers are dissociated by gentle mechanical dispersion (*e.g.*, trituration or pressing between sterile glass plates), and erythroid cells are separated from hepatic parenchymal cells by, for example, low-speed or gradient centrifugation.

The cells may then be used as the "target preparation" for immunization or antibody selection. It is recommended that the cells be used immediately and without further manipulation, so as not to affect antigen display. However, cultured cells or preserved cells with or without mild fixation may also be used as the target preparation, and it is also possible to use cellular extracts, purified membranes, or antigen fractions as the target preparation.

Antibodies can be raised against antigens in the target preparation by immunizing animals with the target preparation. Preferably, the animals have been previously tolerized to adult human blood cells, preferably nucleated red blood cells isolated from adult human peripheral blood. Serum may be collected from such immunized animals, and polyclonal antibodies may be purified from the serum. For methods of antibody production, see generally the Handbook of Experimental Immunology (D.M. Weir & C.C. Blackwell, eds.); and Current Protocols in Immunology (J.E. Coligan *et al.*, eds., 1991).

When animals are immunized with the target preparation, it is preferable to prepare monoclonal antibodies. Monoclonal antibody production is well known in the art, and generally involves isolation of immunoglobin-producing cells (or immunoglobin producing cell precursors) from immunized animals. The isolated cells are immortalized by, for example, fusion with a myeloma cell line which does not produce immunoglobin or by transformation with Epstein-Barr virus (EBV). Clones of immortalized cells which produce antibodies of interest are isolated by screening the clones (or supernatant from cultures of the clones) against an antigen of interest, typically the target preparation. Methods of monoclonal antibody production can be found in, for example, U.S. Patent Nos. 4,491,632, 4,472,500, and 4,444,887, and Galfre et al. (1981, Meth. Enzymol., 73B:3-46).

In a particularly preferred method, the target preparation is used to select antibody-producing clones from an established library of immunocompetent cells or particles. Preferably, the library is a "naïve" library, which means that it is not biased by previous immunization events. The preferred naïve library will either be a germ-line library, or a library prepared from a young, immunologically naïve animal neither tolerized nor sensitized against any foreign antigen. Especially preferred is a "germ-line" library, in which an array of variable regions (usually V_H and V_L) are obtained in germ-line form and assembled in the library in random heterodimeric combinations. The variable regions in a germ line library will not have gone through the somatic mutation events that normally occur in cells of the B lymphocyte lineage during affinity maturation. The number of theoretical combinations of germ-line V_H and V_L regions (the product of the numbers of encoded V_H and V_L variants) can exceed 10^9, 10^{11}, or even 10^{13}, especially when encoding sequences from a large plurality of out-bred individuals of the same species are used in preparing the library. Higher numbers of V_H-V_L combinations are preferred, since this increases the probability of obtaining a specific antibody with a higher affinity. A key advantage of a germ line library is that it will not have immunological blind spots due to tolerization for self antigens, as would be present in a library obtained, say, from the rearranged immunoglobulin genes of a mature B lymphocyte population. Thus, antibodies against rare self-antigens are obtainable. For preparation of germ line antibody libraries, see generally Marks et al. (1996, N. Engl. J. Med. 335(10):730-733), and McGuinness *et al.* (1996 Nat. Biotechnol. 14(9):1149-1154). Particularly preferred is the Griffiths library, described in Griffiths *et al.* (1993, EMBO J. 12(2):725-734), in which single-chain variable regions (scFv) are displayed on phage.

To perform the selection, the cells or viral particles are contacted with the target preparation under conditions that permit the antibody to bind the cells if they display an antigen binding site specific for a cell antigen, as will be undersood by one of skill in the art. The bound cells or viral particles are separated from unbound cells or viral particles, then the bound cells or viral particles are released from the target preparation and the

process is preferably repeated several times. Negative selection can optionally be conducted by contacting with mature erythrocytes or other non-erythroblasts and collecting the unbound cells or viral particles. The cells or viral particles can be replicated at various points during selection if necessary to replenish the supply.

Selected antibodies are preferably further validated by clonally replicating the particles or cells expressing the selected antibody, then testing the clones (or the antibody produced by them) in positive and negative screens. Positive screening may be accomplished by testing the antibodies with cells or antigen preparations which the antibody should bind to, such as erythroblasts, preferably fetal erythroblasts. Negative screening may be accomplished by testing the antibodies against cells or antigen preparations to which the antibody should not react, such as mature erythrocytes, monocytes, granulocytes, or lymphoid cells from periperal blood. Optionally, the antibodies may be additionally negatively screened against erythroblasts and bone marrow from adults. Antibodies that react in the positive screen and do not react in the negative screen may be used in the fetal cell enrichment methods of the invention, but are preferably further selected and/or characterized.

Preferably, the antibodies which pass validation testing are also tested in an immunoaffinity purification assay, if such an assay has not been part of the validation testing. As will be understood by one of skill in the art, any given antibody may have varying effectiveness across different assays. For example, an antibody which is highly efficacious in immunostaining may perform poorly in quantitative immunoassays such as an ELISA. Accordingly, it is recommended that antibodies which pass validation testing be further tested for the ability to enrich erythroblasts from amniotic cord blood samples. Optionally, antibodies which perform well in enriching erythroblasts from amniotic cord blood samples are further tested for the ability to enrich fetal erythroblasts from a maternal blood sample.

At any time during or following the selection or validation process, further adaptations of the antibody molecule both within and outside the variable region can be conducted. It has been found that a proportion of scFv antibodies, when not expressed on the surface of a phage, undergo denaturation upon incubation for several hours or days at 37°C. This is attributed to a weak affinity between the V_H and V_L chains along the interface. Antibodies with this property can be tested while attached to the phage, or converted to another construct such as an antibody consisting of or containing an Fab fragment. The V_H and V_L interface is stabilized in the Fab due to interaction of the CL and CH1 immunoglobulin domains. Conversion of genetic constructs encoding scFv to those that encode Fab is a matter of standard genetic manipulation, and is illustrated herein.

The antibodies of the invention include antibody molecules having the V_H or V_L sequence of the exemplary antibodies, with or without modifications in the amino acid

sequence. Acceptable modifications to the V_H or V_L sequence of the exemplary antibodies include amino acid insertions, deletions, and substitions, so long as the modified antibodies retain the specificity of the 'parent' antibody (*e.g.*, the antibody upon which the modified antibody is based). A wide range of alterations of the variable region framework are typically available that do not compromise specificity. Alterations in buried residues, interface residues, and antigen-binding residues are less frequent, as are non-conservative substitutions and excisions that affect folding, but all such alterations are permissible as long as the specificity of the parent antibody is maintained. Methods used for 'humanization' of non-human antibody variable regions, such as those disclosed in International Patent Applications Nos. WO 94/11509 and WO 96/08565 may be applied to the antibodies of the invention, or may simply be used as guides for selecting residues to be altered. Alterations are also permitted that improve specificity, including mutations in the CDR and substitution of either the V_H or V_L with a variable region chain from another antibody.

Certain embodiments of the invention are antibodies having the complementarity determining regions (CDRs) of either the V_H or the V_L (preferably both) that are homologous to those of one of the exemplary antibodies described herein. Preferably, the homologous CDRs contain no more than about 5 alterations per V_H or V_L chain in comparison with the prototype.

Antibodies having any of the alterations indicated above can be identified as having desirable specificity without undue experimentation, by simply conducting binding or purification assays similar to that used to validate the specificity of the parent molecule, as illustrated herein.

Certain embodiments of the invention comprise antibodies that compete with one of the exemplary antibodies for binding to an antigen preferentially expressed on human erythroblasts. Such antibodies can be identified, for example, by adapting any binding or validation assay for the parent molecule to a competition format. In a preferred example, the exemplary antibody is used for immunofluorescent labeling or immunoaffinity purification of erythroblasts in a mixed cell population as already described. However, the cells are preincubated or the separation step is carried out in the presence of the antibody being tested in an unlabeled form. Ability to compete with the exemplary antibody is indicated by decreased effectiveness of the exemplary antibody in labeling or purification. Competition can also be assayed by antibody binding in a blot or antigen immunoassay format.

Certain embodiments of the invention comprise antibodies that bind the same antigen as one of the exemplary antibodies. Such antibodies can be identified, for example, by competition assays using one of the exemplary antibodies and purified antigen. Included are antibodies that are specific for an erythroblast antigen with an apparent

molecular weight of 78 kDa or 90 kDa as determined by polyacrylamide gel electrophoresis in sodium dodecyl sulfate (SDS-PAGE) under disulfide reducing conditions.

Particular antibodies of this invention can be prepared based on the amino acid sequence data provided in this disclosure, incorporating any desired amino acid deletions, additions, or substitutions. Peptide synthesis and assembly is one possible approach, but it is usually more convenient to prepare proteins of the length of variable region chains by expressing a nucleic acid encoding it in a suitable prokaryotic or eukaryotic host cell. One example is the phagemid vector VODOX1, which can be used as a backbone for expressing V_H and V_L polypeptide sequences as an Fab fragment. The construct can encode recognition sites such as polyhistidine or a c-myc tag that permit later purification by affinity methods. Alternatively, antibody can be purified from cell supernatants, lysates, or ascites fluid by a combination of traditional biochemical separation techniques, such as amonium sulfate precipitation, ion exchange chromatography on a weak anion exchange resin such as DEAE, hydroxyapatite chromatography, and gel filtration chromatography.

The antibodies used in the invention have a variety of utilities, including enriching cells in mixed samples, purification of antigens, and imaging, detection, identification, and quantitation of cells.

The erythroid cell antibodies of the invention may be used for direct or indirect immunostaining. Accordingly, the antibodies may be used for imaging, detecting and/or identifying erythroid cells in biological samples, by contacting cells in a sample with an antibody of the invention, permitting formation of a stable complex, and then visualizing cells bearing the stable antigen-antibody complex by any method known in the art. The antibodies may also be used to quantitate erythroid cells in a sample, using the antibodies in a quantitative immunoassay. To the extent that the target antigen is also expressed on cells that are dedifferentiating during oncogenesis, the antibody may also be used to image, detect, identify and/or quantitate such cells.

Antibodies of this invention can be used to raise anti-idiotypes for erythroid antigens, according to any method known in the art. Generally, anti-idiotype antibodies are prepared by using an anti-erythroid cell antibody of the invention as an immunogen or to select antibody producing particles, for example from a phage library. Selection of the anti-idiotype clones is done using the anti-erythroid cell antibody as a positive selector, and using antibodies of unrelated specificity, but generally of the same isotype, as negative selectors. Validation of initially selected clones is performed by inhibition experiments, in which desired clones block binding between anti-erythroid cell antibody and either erythroid cells or the target antigen. Anti-idiotype clones may be further selected for their ability to elicit a specific anti-erythroid cell antibody in a naïve mammal, or selecting a

specific anti-erythroid cell antibody from an antibody library. Anti-idiotypes can then be used to obtain additional clones of anti-erythroid cell antibodies.

The antibodies of the instant invention are particularly advantageous for enriching erythroid cells in biological samples. A mixed cell population containing erythroid cells is contacted with an antibody of the invention under conditions that permit the antibody to bind to an erythroid cell antigen and form a stable complex, then the cells bearing the stable antigen-antibody complex are separated from cells not bearing the stable complex. The mixed cell population may be any population containing erythroid cells, including bone marrow cells and other blood cell progenitor and precursor populations. Of particular interest are obtaining fetal erythroid cells from maternal blood for purposes of prenatal genetic diagnosis, as described herein.

Antibodies of this invention can also be used to identify, purify, or characterize their target antigen. The Examples provide an illustration of the immunoaffinity purification of a 78-90 kDa antigen from a lysate of erythroblasts, using the antibody produced by the antibody referred to as Clone 1, whose heavy and light chain coding sequences are SEQ ID NOs:8 and 9, respectively. The expression pattern of this antigen along the erythrogenic pathway is compared with that of other cell markers in Table 1.

TABLE 1

	Relative Expression (- to ✓✓)					
Cell Phenotype	**Hemo-globin**	**CD71 (TfR)**	**CD45**	**CD36**	**Glyco-phorin A**	**Clone 1 Antigen**
Proerythroblast	-	-	-	✓	✓	✓✓
Basophilic Erythroblast	-	?	-	✓	✓	✓✓
Polychromatophilic Erythroblast	✓✓	✓	-	✓	✓	✓✓
Orthochromatic Erythroblast	✓✓	✓	-	-	✓✓	✓✓
Reticulocyte	✓✓	✓	-	-	✓✓	✓
Mature Erythrocyte	✓✓	✓	-	-	✓✓	-
Other Blood Cells	-	✓	✓✓	✓	-	-

Once the amino acid sequence of the target antigen is obtained, the full length antigen or a fragment of the antigen can be prepared synthetically for further use.

Generally, polypeptides can be prepared either by chemical synthesis, or by expression of a nucleic acid encoding it in a cell-free translation system or in a host cell. Short polypeptides of about 30 or fewer amino acids in length are conveniently prepared

from sequence data by chemical synthesis. A preferred method is solid phase synthesis, in which the C-terminal amino acid is attached to a solid phase and the peptide is grown towards the N-terminal, as is well known in the art, using iterative cycles of deprotection of the growing protein on the solid phase and coupling the next amino acid, followed by cleavage of the completed peptide from the solid phase and deprotection of the amino acid side chains. Recombinant expression is the preferred method for production of longer polypeptides. A large variety of recombinant expression systems are known in the art, utilizing a variety of constructs and host cells. Generally, a nucleic acid encoding the desired protein is operatively linked to a suitable promoter in an expression vector, and transfected into a suitable host cell. The host cell is then cultured under conditions that allow transcription and translation of the protein, which is subsequently recovered and purified.

The epitope to which a particular antibody binds can be mapped by preparing fragments and testing the ability of the antibody to bind. For example, sequential peptides of 12 amino acids are prepared covering the entire sequence, and overlapping by 8 residues. The peptides can be prepared on a nylon membrane support by F-Moc chemistry, using a SPOTSÔ kit from Genosys according to manufacturer's directions. Prepared membranes are then overlaid with the antibody, washed, and overlaid with ß-galactose conjugated anti-human IgG. The test is developed by adding the substrate X-gal. Positive staining indicates an antigen fragment recognized by the antibody.

Purified erythroblast antigens and antigen fragments may in turn be used to prepare additional erythroblast-specific antibodies according to the general techniques already described.

Antibodies against a fetal cell associated or specific antigen may be derivatized for use in the methods of the present invention. Details of production of antibody fragments and derivatives are well known in the art and may be found, for example in , "Antibody Engineering," 2nd edition (C. Borrebaeck, ed., Oxford University Press, 1995) and "Immunoassay" (E. P. Diamandis & T.K. Christopoulos, eds., Academic Press, Inc., 1996). The term "antibody" also refers to fusion polypeptides comprising an antibody of the invention and another polypeptide or a portion of a polypeptide (a "fusion partner"), such as an affinity tag, an enzyme or other fusion partner.

For use in certain aspects of the instant invention, antibodies may be "primarily" or "secondarily" labeled. A primarily labeled antibody is an antibody which is directly conjugated to a composition which permits detection of the antibody. A secondarily labeled antibody is an antibody which is bound to a detection composition through at least one intermediate composition. For example, an antibody may be primarily labeled by covalent linkage to an enzyme or fluorescent molecule or by adsorption to a magnetic particle. A secondarily labeled antibody may be unmodified and labeled by

binding a labeled antibody-binding protein (such as Protein A, Protein G, or an anti-immunoglobin antibody which may be primarily or secondarily labeled itself), or modified, and labeled by a compound which specifically binds the modification (*e.g.*, covalent modification with a hapten such as biotin followed by labeling with labeled hapten binding protein such as avidin or streptavidin).

5.8. METHODS OF FETAL CELL DETECTION

The present invention provides methods of detecting rare fetal cells in a mixed cell population. Such methods utilize the nucleic acids identified herein as being selectively or specifically expressed in fetal cells relative to other cell types in the mixed cell populations of interest.

Identification of fetal cells with nucleic acid probes is normally carried out using an *in situ* approach, generally fluorescent *in situ* hybridization (FISH), although *in situ* amplification methods are also contemplated. For FISH, a nucleic acid probe is modified (or synthesized with modified nucleotides) so that it can be detected by fluorescence. The nucleic acid probe may incorporate or be covalently bound to a fluorescent dye, it may be modified with a hapten to allow a fluorescent reagent to bind to the nucleic acid probe, or it may be primarily or secondarily labeled with an enzyme which is detected by the use of a fluorogenic substrate. Hapten/hapten binding polypeptide pairs, useful for detection of nucleic acid probe hybridization, include (but are not limited to) biotin/avidin or streptavidin, digoxigenin/ a-digoxigenin antibodies, and dinitrophenol (DNP)/ a-DNP antibodies.

At least one nucleic acid probe is used to identify fetal cells in a maternal sample, although the use of at least 2, 3, 4, 5 or more nucleic acid probes is contemplated. When more than one nucleic acid probe is used, each different nucleic acid probe can be detected using a different fluorescent dye, so that cells expressing multiple nucleic acid probes can be identified. In addition to the nucleic acids of the invention, probes corresponding to genes that are preferentially expressed in fetal cells in a maternal blood sample include, but are not limited to, fetal hemoglobin probes, paternal HLA determinant probes, Y chromosome specific probes, With respect to fetal hemoglobin probes, probes to transcripts of the γ-, ϵ-, or ζ-globin probes can be used, although an ϵ-globin probe is preferred. Because γ-globin transcripts are expressed in RBCs of adults with hereditary persistence of fetal hemoglobin or $\sigma\beta$ thalassemia, and ζ-globin probes transcripts are expressed in RBCs of adults with α thalassemia.

When multiple dyes are used in conjunction with multiple fetal cell probes, it is preferable that the various dyes have non-overlapping fluorescent spectra, or at least that the emissions spectra be distinguishable through the use of narrow pass filters. A large

number of fluorescent dyes are known in the art and commercially available. Commonly used fluorescent dyes include fluorescein, rhodamine, texas red, phycoerythrin, Hoechst 33258, Cascade Blue, Cy3, and derivatives thereof. Alternatively, multiple nucleic acids probes can comprise the same type of label for the purpose of improving the signal to noise ratio of a single probe.

Methods for *in situ* hybridization (ISH) are well known in the art. Because the cells analyzed by the methods of the invention are generally in a suspension, ISH is normally carried out on fixed, permeabilized cells which have been fixed to an insoluble substrate, such as a poly-L-lysine-coated glass slide or polystyrene plate or dish, although ISH may also be carried out on fixed cells in suspension. Where the cells are adhered to a substrate, the substrate is preferably transparent to visible and ultraviolet light (*e.g.*, glass), to allow for use of fluorescent dyes as labels. As will be appreciated by one of skill in the art, materials and solutions used in preparation of cells for ISH and for ISH itself are preferably RNase-free.

Generally, a suspension of cells, preferably at least about 106, 107 or 2 x107 cells/milliliter is made in a solution comprising little or no added protein (*e.g.*, serum free medium or a balanced salt solution) and placed on substrate which has been derivatized to allow attachment of cells by use of a crosslinking agent. Preferably, the substrate is modified by coating with poly-L-lysine or by "subbing" with gelatin. The cell suspension is placed on the substrate, generally as a small "pool" or drop on the surface of the substrate, and the cells are allowed to attach to the substrate by settling under normal gravity for a period of time, preferably at least about 10, 20 or 30 minutes, although the cells may be "spun" onto the substrate by the use of a centrifuge with an approprate rotor adapted to hold the substrate. Attachment of the cells onto the substrate is preferably accomplished under conditions of humidity approaching 100%, as will be apparent to one of skill in the art.

After the cells have attached to the substrate, the cells are crosslinked to the substrate (or to the derivative bound to the substrate) using a fixative. Any appropriate fixative may be used, including acid alcohol solutions, acid acetone solutions, aldehyde fixatives, homobifunctional crosslinking agents such as N-hydroxysuccinimide (NHS) esters (*e.g.*, disuccinimidyl suberate, disuccinimidyl glutarate, and the like) and heterobifuncational crosslinking agents known in the art. Preferably, an aldehyde fixative such as formaldehyde, paraformaldehyde or glutaraldehyde, is used to crosslink cells to poly-L-lysine or gelatin coated substrates. Preferably, the cells are fixed to the substrate by placing the substrate with attached cells into a bath of fixative solution, although fixation may be accomplished by replacing the pool or drop of liquid containing the cells with a similar volume of fixative. The attached cells and substrate are incubated in the fixative for a period of time appropriate to the particular fixative selected by the practitioner, preferably about 20 minutes in the case of 4% paraformaldehyde.

After fixation, the substrate may be rinsed, typically with a buffered saline solution such as phosphate buffered saline or tris-buffered saline, dehydrated using a series of ethanol baths (*e.g.*, by incubating the fixed cells in 50%, 70%, 95%, and 100% ethanol for 2-5 minutes each) air dried, and stored for later ISH procesing. Where the cell/substrate preparation is intended for immediate ISH processing, the cells must still be permeabilized, preferably by incubating the cell/substrate preparation in 50% ethanol, although detergent solutions, such as 0.01 to 0.1% t-octylphenoxypolyethoxyethanol or polyoxyethylenesorbitan monolaurate, may also be used.

Alternatively, the cells may be fixed in solution using an appropriate fixative, rinsed, dehydrated and embedded in paraffin, then sectioned and adhered to glass slides using conventional histologic processing techniques. Prior to processing for ISH, cells processed in this matter must be de-paraffinized, typically by use of a xylene bath, and rehydrated by processing through progressively less concentrated ethanol solutions, as is well known in the art.

The cells to be analyzed are first denatured, generally by use of extreme pH (*e.g.*, 0.2 N HCl for 10-30 minutes at room temperature) followed by high temperature (*e.g.*, 10-20 minutes at 70° C in 2 x SSC), and an additional digestion with a non-specific protease (*e.g.*, pronase) may be included as well. After denaturation, a post-fixation step is preferably performed by incubating the denatured cells in fixative (*e.g.*, five minutes in 4% paraformaldehyde at room temperature), followed by rinsing in a buffered salt solution.

Non-specific binding sites on the cell/substrate preparation are preferably blocked prior to hybridization with probes, typically by acetylation and modification of free sulfur groups. Preferably such blocking is carried out by incubating the cell/substrate preparation in a sulfur reducing agent (*e.g.*, 10 mM dithiothreitol, DTT, in buffered saline at elevated temperature, such as 10 minutes at 45° C), followed by incubation with DTT, iodoacetamide, and N-ethylmaleimide (*e.g.*, 10 mM DTT, 10 mM iodoacetamide, 10 mM N-ethylmaleimide for 30 minutes at 45° C). Additional blocking of polar and charged groups may be accomplished by incubation of the cell/substrate preparation in acetic anhydride (*e.g.*, 0.25 to .5% for 5-10 minutes at room temperature).

Probe nucleic acid probe is denatured prior to hybridization with the prepared cells. Normally, the probe is precipitated in ethanol, then redissolved in a small volume of solvent such as 2 x SSC, 1 x TEA, or formamide, heat denatured by incubating at 70° C or higher for 10-20 minutes, then added to a hybridization mixture. A non-specific, unlabeled DNA, such as sonicated salmon sperm DNA is preferably denatured along with the probe. Generally, when more than one probe is used, the probes are hybridized with the cells at the same time, although use of multiple probes does require use of divergent labeling systems to avoid signal crossover.

Hybridization is typically carried out at elevated temperature in hybridization mix containing a buffered salt solution (*e.g.*, 4 x SSC), a high molecular weight polymer to increase the effective concentration of the probe(s) (*e.g.*, 20% dextran sulfate), and a protein blocking agent (*e.g.*, 2 mg/mL high purity bovine serum albumin). Hybridization is typically carried out under a coverslip which may be anchored in place with rubber cement or any other material which serves to temporarily anchor the coverslip and reduce evaporation of the hybridization mixture. Hybridization is preferably carried out under conditions where the hybridization temperature is 12-20° C below the melting temperature (Tm) of the probe. The Tm of a long nucleic acid can be found as Tm = 81.5 - 16.6(log10[Na+]) + 0.41(%G + C) - 0.63(%formamide) - 600/N, where N = the length of the selectively hybridizable nucleic acid under study, while the Tm of oligonucleotides from about 70 to 15 nucleotides in length may be found as Tm = 81.5 - 16.6(log10[Na+]) + 0.41(%G + C) - 600/N, and the Tm of short oligonucleotides of <14 nucleotides may be found as Tm = 2(A+T) + 4(G+C), where A, T, G and C are the numbers of adenosine, thymidine, guanosine and cytosine residues, respectively. Hybridization may be accomplished in as short a period as 2-4 hours, although longer hybridization incubations are also acceptable. Alternatively, glycerol-based ISH technology, such as that disclosed in International Patent Application No. WO 96/31626 or U.S. Patent No. 5,948,617, may be used.

After the hybridization incubation is completed, the hybridization solution is removed, and the cells are washed, typically for 15 minutes each in 50% formamide/2 x SSC at 37° C, 2 x SSC at 37° C, and 1 x SSC at room temperature. After washing is completed, the cells are incubated in the detection reagent (*e.g.*, fluorescently-labeled avidin or streptavidin for a biotinylated probe). The exact conditions of the incubation with the detection reagent will vary depending on the exact identity of the detection reagent, but is typically accomplished by incubation for 30-60 minutes at 37° C in a chamber protected from ambient light (to reduce photobleaching of the fluorescent label), although signal amplification techniques generally require multiple incubations, as will be apparent to one of skill in the art. Amplification techniques such as the use of secondary antibodies which bind to a primary detection reagent or enzymatic amplification may be employed if so desired. Excess detection or amplification reagent is washed away, typically by rinsing with a buffered salt solution (*e.g.*, 4 x SSC) at room temperature. Optionally, a rinse including a detergent (*e.g.*, 0.1% t-octylphenoxypolyethoxyethanol) in the buffered salt solution may be incorporated in the wash protocol.

Genomic DNA in the cells may be counterstained by incubation with a double-stranded DNA-binding dye, such as propidium iodide or 4,6-diamidino-2-phenylindole (DAPI) and rinsing away unbound dye.

Where immunoenriched cells are processed as cells in suspension, the cells are carried through a substantially similar process, except that the cells are collected by centrifugation or filtration after each step (*e.g.*, after fixation, each wash step, etc.).

After hybridization, labeling with detection reagent and counterstaining, the cells are preferably sealed under a coverslip with an anti-fading reagent appropriate to the fluorescent dye(s) used in the detection reagent. The appropriate anti-fading reagent can be easily selected by the skilled practitioner.

Fetal cells may be detected by the use of any convenient fluorescent microscopy technique, including epifluorescence microscopy, confocal fluorescence microscopy, and other techniques known in the art. Results of microscopy may be stored on photographic negatives, photographic plates, or on magnetic or optical storage media when a CCD camera or other electronic imaging equipment is used. Alternatively, cells which are processed as cells in suspension may be analyzed using FACS technology.

Nucleated cells which are present in the sample following immunoenrichment and are labeled by at least one of the nucleic acid probes are considered identified as fetal cells.

In situ single cell PCR, for example using PCR primers corresponding to the nucleic acids of the invention, also offers a method for detection of single cells of fetal origin. With this method, each cell, fixed either in suspension or on a solid support, and either as a single cell or in the context of surrounding tissue, functions individually as a reaction chamber for the PCR. With proper fixation and permeabilization conditions, the oligonucleotide primers and other reaction components are able to diffuse into the cells, and, upon thermal cycling, are able to amplify available specific target sequences. The product DNA retained within the source cell can be readily detected by standard *in situ* hybridization (Brezinschek *et al.*,1995, J. Immunol. 155:190). For diagnostic purposes, single fetal cells can be isolated and product DNA can alternatively be extracted and subjected to gel electrophoresis or southern blotting.

Specific or selective acting fetal cell probes of the invention can be labeled with radioactive labels including radionucleides (*e.g.* ^{35}S,^{32}P or ^{3}H) and hybridized to nucleic acid that has been extracted or amplified via various PCR techniques from single cells or samples of cells. Southern and Northern blot analyses standard for practitioners in the art can then be utilized to confirm presence of fetal cells in nucleic acid extracted from the samples. If RT-PCR is used in conjunction with the fetal cell associated primers of the invention to produce cDNA's specific to fetal cells in a mixed fetal maternal sample, then detection of amplified cDNA in cells could also be accomplished with radioactive labeling of cDNA followed by autoradiography or scintillation counting.

In a preferred embodiment, non-isotopic labels (*e.g.* biotin or digoxigenin) for RNA probes could ideally be used for non-radioactive *in situ* and Northern blotting

applications to detect fetal cells. Non-isotopic labeled RNA probes offer several advantages over other types of probes: RNA/RNA hybrids are more stable than RNA/DNA hybrids; RNA probes are single stranded and don't re-anneal on themselves; RNA probes can be labeled throughout the molecule; and RNase A can be used to eliminate unhybridized single stranded probe. These factors result in RNA probes that are more sensitive and have lower background than either cDNA or oligonucleotide probes. Thus, the fetal cell probes of the invention, labeled with non-isotopic identifiers offer a superior technique for detection. In addition the fetal cell probes of the invention, with non-isotopic labels, enable simultaneous use of probes specific for genetic disorders or traits and aimed at nuclear DNA.

If the non-isotopic labeled RNA probes contain fluorescence markers, then fetal cells may be detected by the use of any convenient fluorescent microscopy technique, including epifluorescence microscopy, confocal fluorescence microscopy, and other techniques known in the art. Results of microscopy may be stored on photographic negatives, photographic plates, or on magnetic or optical storage media when a CCD camera or other electronic imaging equipment is used. Alternatively, cells which are processed as cells in suspension may be analyzed using FACS technology.

Identification of fetal cells with nucleic acid probes is normally carried out using an *in situ* approach, generally fluorescent *in situ* hybridization (FISH), although *in situ* amplification methods are also contemplated, as discussed above. For FISH, a nucleic acid probe is modified (or synthesized with modified nucleotides) so that it can be detected by fluorescence.

A fetal cell probe is a reagent for detecting a fetal cell RNA contained in a fetal cell potentially present in a sample of interest by a hybridization reaction. Usually, a probe will comprise a label or a means by which a label can be attached, either before or subsequent to the hybridization reaction. Means for attaching labels include biotin moieties that couple with avidin or streptavidin, haptens that couple with anti-hapten antibody, and particular nucleic acid sequences (optionally on a branch or fork) that hybridize with a reagent nucleic acid having a complementary sequence, any of which ultimately lead to the attachment of a label. Suitable labels include radioisotopes, fluorochromes, chemiluminescent compounds, dyes, and proteins, including enzymes. The probe may incorporate or be covalently bound to a fluorescent dye, it may be modified with a hapten to allow a fluorescent reagent to bind to the probe, or it may be primarily or secondarily labeled with an enzyme which is detected by the use of a fluorogenic substrate. Hapten/hapten binding protein pairs, useful for detection of nucleic acid specifier hybridization, include (but are not limited to) biotin/avidin or streptavidin, digoxigenin/ a-digoxigenin antibodies, and dinitrophenol (DNP)/ a-DNP antibodies. In preferred embodiment, the signal arising from the probe which indicates hybrid formation between a probe and its target is described in FIG. 18 and Section 8, *infra*. Such modifications are

also contemplated for the diagnostic probes that are used in conjunction with the fetal cell probes of the invention.

In other embodiment of the present invention, instead of, or in conjunction with, using fetal cell associated probes to identify rare fetal cells in a maternal blood sample, an antibody that immunospecifically binds to a fetal cell antigen, such as an antibody directed against a polypeptide of the invention or an antibody that is identified by the methods described in Section 5.7.2, *supra*, including but not limited to the anti-Clone-1 antibody or derivatives thereof, can be used for fetal cell detection. A maternal blood sample, which has been optionally immunoenriched for fetal cell, is contacted by an antibody against a fetal cell associated antigen, including but not limited to the antibodies described of Sections 5.5 and 5.7.2. Antibody-bound cells can be identified by routine immunostaining methods known in the art. As will be readily apparent to one of skill in the art, the signal amplification step of FIG. 18 and Section 8 can be readily adapted to methods where the agent bound to fetal cells is an antibody rather than a nucleic acid probe.

5.9. METHODS OF FETAL CELL DIAGNOSIS

The identification methods of the present invention allow for non-invasive prenatal diagnostics.. As discussed above, in a preferred embodiment, the identification of fetal cells involves contacting the biological sample containing the cell or a cell extract with the probe under conditions where the nucleic acid can selectively or specifically hybridize with the target transcript. The target transcript may be RNA or a cDNA copy. Where a cell extract is used, formation of a stable hybrid will indicate that at least one cell containing the transcript was present in the original cell population. Where permeabilized whole cells are used, detection of hybrid formation will indicate which cells in the population contain the transcript.

Optionally, the fetal cells are separated from the maternal cells prior to carrying out the diagnostic methods of the invention. Thus, the probe sequences can also be used to separate fetal cells expressing the target transcript from maternal cells in a mixed population. An example of an intracytoplasmic staining method for cell separation using nucleic acid sequences is described generally in U.S. Patent No. 5,648,220. Briefly, the cell is lightly fixed with 2-8% paraformaldehyde and permeabilized with aqueous alcohol, such that the cell remains sufficiently intact to retain the target sequence. The cells are then contacted with a probe sequence or plurality of sequences to which a detectable label (such as a fluorescence marker) is attached. After washing, the identified cells are then separated from other cells, either by micromanipulation, or by an automated method such as fluorescence-activated cell sorting.

Where the detection and diagnostic methods of the invention entail the use of PCR, the PCR reaction can be performed *in situ*. For the diagnostic methods of the

invention, the PCR reaction can be performed on a single cell that has been identified by the fetal cell probes and antibodies of the invention to be a fetal cell. Micromanipulation methods are known in the art and can be used to separate a fetal cell from the maternal blood sample and place into a suitable container for the PCR reaction.

Identified fetal cells from a maternal sample can be used in diagnostic assays, particularly assays for genetic diseases, as will be apparent to one of skill in the art. Normally, such diagnostic assays are carried out using FISH technology and a diagnostic probe. As will be apparent to one of skill in the art, diagnostics assays on the fetal cells may be carried out after the ISH procedure with the fetal cell probe or antibody, or may be carried out concurrently. The exact size and sequence of the diagnostic probe will depend on the identity of the genetic disorder which is the subject of testing. For example, when testing for a trisomy (*e.g.*, Down's Syndrome or trisomy 21), a probe specific for the chromosome of interest is utilized, while testing for genetic diseases will utilize one or more probes specific for disease-causing or associated alleles. In a preferred embodiment, the trisomy 21 probe is the AneuVysion® probe (Vysis).

When the diagnostic assay is carried out sequentially (*e.g.*, after identification of fetal cells with a probe DNA), the location of fetal cells in the sample can be recorded, then the DNA probes and detection reagents can be removed from the sample by stripping. Generally, stripping is accomplished by denaturing the sample using extreme pH or elevated temperatures. After denaturation, the sample is processed using the desired diagnostic assay. As will be apparent to one of skill in the art, the details of conducting the assay will depend on the exact identity of the assay and the form of the sample.

As will be apparent to one of skill in the art, diagnostic assays carried out concurrently with the DNA probe ISH step should be assays which do not interfere with the DNA probe and detection system utilized. Accordingly, a diagnostic assay run concurrently with the specification step will normally utilize a non-overlapping detection system (*e.g.*, where the DNA probe step utilizes a biotinylated probe, the diagnostic assay utilizes a different detection technology, such as digoxigenin-modified probes, and the fluorescent dyes utilized in the detection system will be different). However, the same detection system may be used if the subcellular localization of the fetal cell vs. diagnostic probe (*e.g.*,

A wide variety of diagnostic assay technologies and probes are available for detection of chromosomal abnormalities and/or genetic diseases. For example, U.S. Patent No. 5,447,841 discloses probes specific for chromosome 21, which may be utilized in a diagnostic assay for trisomy 21 (*i.e.*, Down's syndrome). Multiple genetic disorders may be assayed in a single test utilizing the multiplex FISH methods disclosed in U.S. Patent No. 6,007,994.

5.9.1. DIAGNOSIS OF FETAL GENETIC ABNORMALITIES

Once the fetal cell probes of the invention have been employed to identify fetal nRBC in maternal blood samples, several possibilities emerge for diagnosis and genotyping of genetic disorders. Fluorescence DNA probes specific to interphase stage nuclei have been developed to identify chromosomal disorders of aneuploidy (Down syndrome, Klinefelter syndrome, and trisomy 13) (Simpson and Elias, 1995, Human Reproductive Upate 1(4):409-418). FISH analysis (fluorescence *in situ* hybridization) using dual color X and Y specific DNA probes has also been developed to determine fetal sex. The advantage of such techniques is that these nuclear DNA probes can be used simultaneously with the RNA tag based probes of this invention, allowing for a non-laborious method of multiple diagnoses through multiprobe florescence *in situ* hybridization.

If identified fetal cells have been sufficiently isolated from maternal cells, for example by single cell micromanipulation techniques and PCR described below, then mutation detection by fetal DNA analysis can be conducted. The genes responsible for many single gene disorders have been mapped and cloned, including spina bifida, sickle-cell anemia, thalassaemias, Marfan Syndrome, and Duchenne Muscular Dystrophy. For the single gene mutation causing Cystic Fibrosis, PCR or ARMS multiplex tests are typically used to detect the known causal mutations (Ferrie *et al.*, 1992, Am. J. Hum. Genet. 51(2): 251-262). Amplification proceeds with PCR primers specific to known mutations in the gene. A diagnosis can be made which is then confirmed by DNA sequence analysis of the gene.

Another category of single gene disorders encompassed by the diagnosis methods of the present invention relates to diseases caused by expansion of blocks of repeating nucleotide triplets within a gene. For many of these disorders, PCR based primers have been developed for the responsible genes, including Fragile X syndrome (FMR1 gene)(Chong *et al.*, 1994 Am J Med Genet 51:522-526), Friedreich's ataxia (Filla *et al*, 1996, Am J Hum Genet 59:554-560), myotonic dystrophy (Brook *et al.*, 1992, Cell 21;68(4):799-808.), and Huntington's Disease (Warner *et al.*, 1993, Mol Cell Probes 7: 235-139). In Huntington disease, the DNA sequence, CAG, is part of this sequence. This sequence may be duplicated many times in individuals, up to 26 times in the general population. The duplication of this segment is called a "trinucleotide repeat" in which these three nucleotides (CAG pattern) are repeated over and over again. Individuals with Huntington disease may have from 40 to over 100 repeated CAG segments. The normal number of CAG repeats is from 11-24. Since the sizing of alleles is essential to diagnosis, DNA sequencing is performed. Southern blotting is also used to back up the PCR test, especially for large amplifications or individuals with a single normal allele. The sequences labeled and used as probes for lengths of repeats could serve as a basis for developing probes which could be utilized *in situ* with fetal nRBC's. Such nuclear DNA probes could

be used simultaneously with the RNA cytoplasmic probes based on the tag sequences of this invention. Thus allowing for detection, by florescent microscopy screening, of individual fetal cells with genetic disorder markers. The DNA specific probes and associated assays would not interfere with the RNA-based fetal cell detection system, providing and added benefit. The present invention might also allow for diagnosis of fetal infections, including but not limited to retroviral infections (*e.g.*, HIV). The fetal cell genome can be diagnosed by contacting the fetal cell identified by the methods of the invention (before or after identification) with a probe that will hybridize to genomes of infectious agents of interest.

Certain techniques for acquiring genetic information, especially pertaining to human genetic disorders can be used following or before detection of target cells using fetal cell probes of the invention, or simultaneously with the fetal cell probes or fetal cell antibodies of the invention. Used in combination with available genetic diagnostic procedures, the fetal cell probes and antibodies of the invention aid in detection and confirmation of genetic disorders of a developing fetus. Once the fetal cell probes and antibodies of the invention have been employed to identify fetal nRBC's in maternal blood samples, several possibilities emerge as techniques which can be utilized for diagnosis and genotyping of genetic disorders and traits in the fetus.

5.9.2. DIAGNOSIS OF OTHER FETAL CHARACTERISTICS

Fluorescence probes specific for certain fetal characteristics exist and can be simultaneously or successively utilized with the instant invention. For example, the sex of a fetus is commonly desired knowledge. FISH analysis (fluorescence in situ hybridization) using dual color X and Y specific DNA probes has been developed to determine fetal sex. The advantage of such techniques is that these nuclear DNA probes can be used simultaneously with the fetal cell probes and antibodies of the invention, allowing for a non-laborious method of multiple diagnoses through multiprobe florescence in situ hybridization. Fluorescence probes specific to Y chromosome (*i.e.* the Vysis® LSI SRY DNA FISH probes or the Vysis® WCP Y DNA DNA FISH probe) are targeted at nuclear genetic material not cytoplasmic and thus do not interfere with the fetal cell probes of the invention. Fetal cell probes designed for specific or selective markers in the fetal cell could be utilized, since the maternal genome does not contain Y specific genes. Probes specific to the X chromosome exist (i.e. the Vysis® CEP X probes) as well and could be used in a similar manner to determine sex of the fetus if such probes were used in conjunction with fetal cell probes of the invention which in this case must be targeted at specific markers of the fetal cells.

PCR offers an alternative method to hybridization for determination of sex. PCR primers specific to genes exclusive to the Y chromosome can also be utilized in

conjunction with the fetal cell probes of the invention to determine fetal sex. The PCR reactions may precede, succeed, or occur simultaneous with the fetal probe hybridization technique or use of fetal antibodies of the invention.

Allele-specific PCR primers for the alleles of genes encoding for the proteins responsible for blood types exist. Primers specific to the RhD gene responsible for Rh factor (Gassner *et al.*, 1997, Transfusion 37:1020) are a good example. Such primers can provide genotype data that can be used to determine blood type of the fetus. The fetal cell probes of the invention can provide some degree of confirmation of diagnosis based on PCR results and in conjunction with the PCR can provide exact data necessary to determine the genotype of a fetus with respect to the RhD gene alleles provided the fetal probes of the invention are developed based on specific markers. The PCR reactions may precede, succeed, or occur simultaneous with the fetal cell probe hybridization technique or use of the fetal cell antibodies of the invention.

In cases of multiple pregnancy, the fetal cell probes or antibodies of the invention could be combined with single cell isolation and standard DNA fingerprinting techniques to detected the presence of multiple fetuses at early stages of pregnancy, provided the fetuses are not genetically identical.

With detection of target cells using the fetal cell probes of the invention, any human trait for which the gene(s) the trait controls have been identified can be examined, provided probes or PCR primers specific to alleles responsible for the trait of interest have been developed. The fetal cell probes and the fetal cell antibodies of the invention when utilized in conjunction with existing and future molecular diagnostic techniques will result in an increase of potentially valuable fetal genetic information available to physicians during gestation and after birth.

5.9.3. *IN SITU* DETECTION OF GENETIC ABNORMALITIES

In situ fetal diagnoses of genetic abnormalities can be achieved by combining the fetal cell probes or antibodies of the invention with existing probes aimed at nuclear genetic material that enable one to determine the number of chromosome copies present in fetal cells. Flow cytometry followed by the use of one or more of the various Vysis® DNA FISH probes specific to human chromosomes or portions thereof or other such probes specific to human chromosomes enables detection of fetal aneuploidy (Down syndrome, Klinefelter syndrome, and trisomy 13). Fluorescence DNA probes specific to interphase stage nuclei have been developed to identify chromosomal disorders of aneuploidy (Simpson and Elias, 1995, Human Reproductive Upate 1(4):409-418). If a sample has been enriched for fetal cells, *e.g.*, using an antibody of the invention, it may not necessary to identify fetal cells in the situation as described above, because maternal cells lack such abnormalities. However, the cytoplasmic fetal cell probes of the invention, aimed

at either specific or selective fetal cell markers, whether used before, after, or at the same time as other diagnostic techniques provide and additional confirmation that the diagnosis is limited to the fetal genome. The fetal cell probes and antibodies of the invention, used in conjunction with probes such as those mentioned above, would reduce the number of cells screened in a sample of mixed fetal/maternal blood, reducing the time necessary for diagnosis procedures. In addition the RNA-based cytoplasm specific fetal cell probes do not interfere with the procedures for the nuclear-based probes mentioned above.

Fluorescence probes for microdeletion syndromes also have directed to nuclear DNA that can with easily be utilized in conjunction with the fetal cell probes and antibodies of the invention. Probes for microdeletion syndromes such as DiGeorge, Velocardiofacial, Cri Du Chat, Miller-Dieker, LSI Prader-Willi/Angelman, Smith-Magenis, Wolf-Hirschhorn, and LSI Steroid Sulfatase can be synthesized or purchased. Again, the hybridization probes specific to genetic material responsible for these disorders may be used preceding, succeeding, or simultaneously with the fetal probe hybridization technique or use of fetal antibodies of the invention.

Similarly, the fetal cell probes and antibodies of the invention can be used in conjunction with the other fluorescence probes that enable detection of abnormalities of chromosome telomeric regions, chromosome rearrangements, chromosome deletions & additions, and chromosome translocations. Such probes are commercially available, for example from Vysis. As the availability and number of probes for genetic traits and disorders increases so will the utility of the fetal cell probes and antibodies of the invention. Combined, the probes provide a means to increase the specificity and speed with which a diagnosis can be made.

5.9.4. DETECTION OF GENETIC ABNORMALITIES BY PCR

For single gene genetic disorders where the mother has been genotyped as a carrier, the fetus may have the same genotype with respect to gene responsible for disorders. Techniques such as those described in the previous section are insufficient in this situation and the fetal cell probes and antibodies of the invention then have added value and necessity, since diagnosis cannot generally be made from maternal or enriched fetal cell blood samples in the absence of methods for identifying the fetal target cells with specificity. PCR techniques utilized in connection with the fetal cell probes and antibodies of the invention provide a means to diagnose genotypes in situations where both the mother and fetus are carriers.

The use of PCR techniques in detection of genetic abnormalities can be done *in situ* or following DNA extraction from single identified fetal cells. The efficiency of methods for DNA extraction from single cells is continually being improved http://www.aps.org/meet/MAR01/baps/abs/S2730006.html (Findlay *et al.*, 1997, Nature

389:555-556; Ray and Handyside, 1996, Mol.Hum. Reprod, 2:213-218). These techniques can utilized with isolated fetal cells. Another option is to conduct PCR without any extraction procedure (Küppers *et al.*, 1997, Handbook of Exp. Immunol. 5th ed., Eds. D.M.Weir *et al.*, Blackwell Scientific). With this method, each cell, fixed either in suspension or on a solid support, and either as a single cell or in the context of surrounding tissue, functions individually as a reaction chamber for the PCR. With proper fixation and permeabilization conditions, the oligonucleotide primers and other reaction components are able to diffuse into the cells, and, upon thermal cycling, are able to amplify available specific target sequences. Product DNA is retained within the source cell and is readily detectable by standard in situ hybridization. (Brezinschek *et al.*,1995, J. Immunol. 155, 190). Nucleic acid sequences can now be amplified within the environment of the cell (Komminoth *et al.*, 1992, Diagn. Mol. Pathol. 1:85; Nuovo *et al.*, 1991, Am. J. Pathol. 139:1239). *In situ* PCR can be performed on a single fetal blood cell samples allowing for detection of genetic disorders for which specific primers have been designed. By incorporating molecular beacons into the PCR reaction, a fluorescence can be observed in cells possessing mutated gene copies responsible for the genetic disorder being tested. (Pierce *et al.*, 2000, Molec Human Reproduction, 6(12):1155-1164; Giesendorf *et al.*, 1998, Clin Chem 44:482-486; Bonnet *et al.*,1999, Proc Natl Acad Sci USA 96:6171-6176; Kostrikis *et al.*, 1998, Science, 279:1228-1229). In addition, single cell PCR has been successfully used to identify heterozygous loci http://www.promega.com/geneticidproc/eusymp2proc/21.pdf. Such techniques could be used to determine if a fetus is a carrier for the genetic disorder being tested.

For fetal cells identified the cytoplasm specific fetal cell probes, further isolation from maternal cells in a sample can be achieved by several methods known to those of skill in the art. Techniques for single cell micromanipulation have been developed for embryo manipulation in preimplantation genetic diagnosis and fertility treatments and have successfully been applied to other cell types (Leary, 1994. *In*: Methods in Cell Biology. Flow Cytometry, Darzynkiewicz *et al.*, eds. vol. 42:pp. 331-358; Iritani, 1991, Mol. Reprod. Dev., 28:199-207). The essential equipment consists of an inverted phase fluorescence microscope suitable for observation of single cells that has been fitted with a Narishige micromanipulation/microinjection system for single cell manipulation. The manipulation is dependent on drawn glass capillaries. An alternative method, using a similar microscope, employs an optical trapping system (lazer tweezers). In this technique a laser beam is capable of catching and holding both static and motile cells (Moravcik Z. et al. 1998. The Journal of Eukaryotic Microbiology conference procedings). Grover has had success in using an optical trapping system with erythrocytes (Grover *et al.*, 2000, journal of Optical Society of America 7(13):533). Thus to isolate a single tag-labeled fetal cell in a

sample of maternal blood cells can be accomplished either by using micromanipulation or an optical trapping system.

In situations where the maternal genotype may have copies of genes responsible for the genetic disorders being tested for in the fetus, the fetal cell probes and antibodies of the invention designed to specific markers are required and isolation of target cells may be necessary if DNA extraction is required. The isolation may be mechanical, followed by single cell PCR, or visual utilizing fluorescence microscopy. For example, if PCR primers specific to copies of genes responsible for the trait or disorder being tested for are used in connection with fluorescence markers such as molecular beacons, then the PCR reaction could be conducted before the use of fetal cell probes and cells with both cytoplsamic fluorescence and nuclear (preferably of differing colors) could be identified. If specific nuclear PCR primers have been developed for both normal and mutated copies of disease genes or for mutant copies of genes with intermediate expression, then multi-colored probes could be employed to identify single fetal cells and determine the carrier status of the fetus or the likely severity of the disorder based on genetic compliment the fetus has inherited. The possibility that the fetal cell probes and antibodies of the invention could be used before, after, or simultaneously with such PCR gene specific primers makes the combined use of technologies a strong one.

If identified fetal cells have been sufficiently isolated form maternal cells my methods described in the present application, then mutation detection by fetal DNA analysis can be conducted. Allele-specific PCR primers for alleles of the RhD gene (Gassner *et al.*, 1997, Transfusion 37:1020) which determines human Rh factor have potential in diagnosing the possibility of erythroblastosis fetalis when utilized in connection with the fetal cell probes and antibodies of the invention. Additional single gene disorders or fetal genetic characteristics, *e.g.*, gender or infection, that can be diagnosed by PCR-based techniques are described in Sections 5.9.1 and 5.9.2, *supra*. Amplification proceeds with PCR primers specific to known mutations in the gene. A diagnosis can be made which is then confirmed by DNA sequence analysis of the gene.

5.9.5. METHODS FOR SIMULTANEOUS RNA AND GENOMIC DNA HYBRIDIZATION

As discussed above, in a preferred embodiment of the invention, the fetal cell probes of the invention are used simultaneously with one or multiple fluorescence probes designed to detect specific chromosomes or genomic markers, for example for diagnosis of genetic disorders. The present inventors have developed techniques that allow for the first time simultaneous hybridization with a probe directed to a cytoplasmic RNA and a probe directed to a nuclear DNA. These technique maintain the best possible cell/sample conservation, a strong distinguishable signal in the highest number of cells, eliminate of background autoflourescence, and allow detection of the highest possible fetal cell number

in the maternal cell sample. The simultaneous hybridization techniques of the invention entail the use of one, preferably more than one, preferably at least three or four of the following features. These techniques have been used successfully to detect X and Y chromosome sequences (using probes purchased from Vysis) in the nuclei and epsilon and gamma globulin RNAs in fetal cord blood cells (see FIG. 20)

One feature entails, prior to hybridization, coating slides with anti-cell-surface antibodies, for example anti- IgG antibodies (GPA), followed by the addition of cell suspension and centrifugation. This enables more cells to remain intact, minimizing cell loss.

Another feature entails using probes for both RNA and genomic DNA that comprise at least 45%, most preferably at least 50% GC content. The probes are preferably 25-300, most preferably 30-200, *e.g.*, approximately 30, 50, 70, 100, 150 or 200 nucleotides in length.

Probe accessability can be improved by addition of the detergent Tween-20/PSB (0.2%) incubation at RT for 20 min. and/or addition of the protease Proteinase-K at 0.1ug/ml incubation for 10 min at RT.

Probe hybridization is preferably carried out for a period of 1.5-4 hours at 40-50°C, most preferably for approximately 2.5 hours at 45°C.

The two most preferable features for inclusion in the simultaneous hybridization techniques relate to fixation. A 4% neutral Formalin/PBS, at a neutral pH of 6 to8, more preferably at a pH of 6.5-7.5, most preferably at a pH approximately 7, allows fixation of the freshly prepared cell samples. The cell samples are incubated with the fixative for approximately 20 min at room temperature. Following fixation and addition of probes, preferably after the cells are washed to remove probe, a post-fixation step is preferably performed. Post-fixation entails using a 4% Formalin/PBS for approximately 5 minutes at room temperature. The post-fixative is preferably at an acidic pH, for example at a pH of approximately 2, 2.5, 3, 3.5, 4, 4.5, 5 or 5.5.

Following the FISH procedure, samples can be dehydrated and stored at -20°C to aid in retaining long term cell and probe integrity.

5.10. KITS

The present invention yet further provides kits comprising in one or more containers a first probe which is a fetal cell of the invention, as described in Section 5.3, *supra*, or an antibody against a fetal cell associated polypeptide, as described in Sections 5.5 and 5.7.2, *supra*. The kits of the invention further instructions for diagnostic use and/or a label indicating regulatory approval for diagnostic use. The kits can further comprise one or more antibodies for immunoenriching for fetal cells in a maternal blood sample, for example an antibody that selectively or specifically binds to fetal cells in a maternal blood

sample. The kits can also optional comprise a second fetal cell probe, including but not limited to a fetal cell probe of the invention or a fetal globulin probe. The second probe can correspond to the same or a different fetal cell mRNA as the first probe. The kits of the invention can further include diagnostic reagents for determining the gender of the fetal cells or for identifying abnormalities associated with the fetal cells.

The probes and antibodies contained in the kits of the invention are preferably labeled, for example by a radioactive or fluorescent label, a colorimetric reagent, or an enzyme. Optionally, a kit of the invention further comprises reagents for colorimetric detection of the labeled probes and antibodies.

6. EXAMPLE: Identification of Fetal Cell Associated Antibodies

6.1. EXAMPLE 1: CELL PREPARATIONS

Human fetal livers were harvested from terminated pregnancies for use in antibody selection and validation. For the first round of antibody selection, the liver cells were obtained at between 8-26 (optimally at about 10-18) weeks of gestation. Wherever possible, the individual liver was placed immediately on ice after dissection. Optimally, the warm ischemia time is less than 15 min. The liver was gently divided into small pieces, and then the pieces were disaggregated into individual cells between microscope slides. The preparation was then centrifuged gently in phosphate buffered saline (PBS) at ~3000 rpm to remove liver parenchymal cells. On some occasions, the fetal erythroblasts were characterized by flow cytometry, using mouse anti-CD36 and anti-Glycophorin A (Edelman et al.). The antibodies were labeled with fluorescein and phycoerythrin respectively, and used according to the directions of the flow cytometer manufacturer (Coulter). Typically, 80-90% of the cells were positively stained.

In some of the phage-display antibody selection experiments, fetal erythroblasts were used that had been treated with papain to improve exposure of antigen. Papain digestion was performed by adding several milligrams of the enzyme to the resuspended erythroblast fraction of a single fetal liver preparation and incubating for several hours at 37°C. The treated cells were then washed and used for antibody selection. Some of the antibody validation experiments were performed using cultured erythroblasts. Umbilical cord blood was obtained after delivery of human newborns. The blood was diluted 1:1 with alpha medium (Alpha MEM, Sigma) containing 2% fetal bovine serum (FBS, PAA Laboratories). The cells were layered onto an equal volume of Histopaque™ 1083 (Sigma) and centrifuged at 390 g for 30 min at 18°C to obtain mononuclear cells. The culture method used was a modification of Weinberg et al. (Blood 81:2591, 1993). Briefly, cells were cultured in alpha medium containing 30% FBS, 1% BSA (Boehringer-Mannheim fraction V), 100 μM ß-mercaptoethanol (Sigma), 50 μg/mL Gentamicin (R and D systems), 10 ng/mL IL-6 (R and D systems), 1.3 U/mL Erythropoietin

(Boehringer-Mannheim), and 1 mM L-glutamine (Sigma). Cultures were inoculated with 1-2 x10^7 cells into 10 mL of culture medium, and incubated for 10-21 days at 37°C, 5% CO_2. At day 15, the majority of cells in culture are erythroid, and express high levels of CD36 with a range of Glycoprotein A expression.

For the generation or validation of assays using trophoblast markers, syncytiotrophoblasts are obtained as follows: First trimester placentas are obtained from apparently healthy pregnancies electively terminated by aspiration at 6-10 weeks gestation. Clotted blood and any adherent decidua are carefully dissected from the placentas. Syncytiotrophoblasts are isolated by gently teasing the placentas through a 250-mesh sieve. The sheets of syncytiotrophoblast, being significantly larger than contaminating cells, readily sediment at unit gravity in isotonic medium. After sedimentation for approximately 2 min, the supernatant is decanted and the cells resuspended in fresh solution. This washing procedure is performed three times. The success of trophoblast isolation is confirmed by measuring the synthesis of human chorionic gonadotrophin in culture after three days by immunoassay (*e.g.*, Hybritech). Trophoblast cells can be cultured as described in US. Patent No. 5,503,981. The choriocarcinoma line, JEG-3, can be obtained from the American Type Culture Collection and cultured in RPMI-1640 medium supplemented with 10% fetal calf serum.

6.2. EXAMPLE 2: PREPARATION OF ERYTHROBLAST SPECIFIC ANTIBODIES

A large naive human phage-display library was constructed by recloning the heavy and light chain variable regions from the lox library vectors into the phagemid vector pHEN2 (Griffiths et al., Vaughan et al.). The library displays antibody as a single-chain variable region (scFv) molecule, comprising random combinations of germ-line V_H and V_L regions linked together as part of a single polypeptide chain.

Briefly, the kappa and lambda light chain variable regions were PCR amplified from the fdDOG-2lox Vκ and Vλ phage constructs using the following primers: 5'-GAG TCA TTC TCG ACT TGC GGC CGC ACG TTT GAT TTC CAS CTT GGT CCC-3' (SEQ. ID NO:1) or 5'-GAG TCA TTC TCG ACT TGC GGC CGC ACC TAG GAC GGT CAG CTT GGT CCC-3' (SEQ. ID NO:2) and "FdPCRback": 5'-GCG ATG GTT GTT GTC ATT GTC GGC-3' (SEQ. ID NO:3). The PCR fragments were purified and digested with ApaL1 and Not1. The gel purified fragments were then ligated into the vector pHEN2 in several aliquots. DNA was then purified from the ligation mixtures, resuspended in water, and electroporated into *E.coli* TG1. Vk-pHEN2 or VL-pHEN2 library pools of 3.5 x 107 and 1.67 x 107 respectively, were obtained. V_H regions were PCR amplified from the pUC19-2lox V_H vector using the primers "LMB3" 5'-CAG GAA ACA GCT ATG AC-3' (SEQ. ID NO:4) and "CH1.LIBSEQ" 5'-GGT GCT CTT GGA GGA GGG TGC-3' (SEQ. ID NO:5).

The PCR fragments were purified and digested with Sfi 1 or Nco 1 and Xho 1. The gel purified fragments were then ligated into the vectors Vκ-pHEN2 or Vλ -pHEN2. DNA was purified from the ligation mixtures, resuspended in water, and used for several hundred electroporations into *E. coli* TG1 to obtain a total library size of 2×10^9. The scFv fragments contain a small c-myc peptide fused at the C-terminus as a tag to facilitate detection of the soluble scFv fragment using an anti-c-myc monoclonal antibody conjugated to horseradish peroxidase (HRP).

Human fetal erythroblast cells were prepared as described in Example 1, with or without papain treatment. The cell population from one liver was re-suspended in filtered PBS/2% marvel, which acted as a blocking agent against non-specific binding of phage to the cells.

Approximately 10^{13} phage from the phage display library (~$6x\ 10^9$ human scFv clones) were incubated for 16 h with $3x\ 10^6$ cells in filtered PBS/2% marvel to a final volume of 250 µL at 4°C. Cells were washed 5-6 times within 1 h, and then lysed in distilled water. The debris containing the phage was collected and used to infect an exponentially growing culture of *E. coli* TG1. Infected cells were grown overnight on plates containing 100 µg/mL of ampicillin and 2% glucose. The plates were scraped the next day, and phage were rescued from the selected population using M13 K07 as described in the art (Marks et al., 1991). Rescued phage were used to perform another selection, and the process was repeated until 3 rounds of selection had been carried out.

Individual colonies were grown in 96 well plates, and production of scFv was induced using 1 mM IPTG for 16 h. Clones were selected in an ELISA-format assay, using density-purified erythroblasts from fetal livers (positive selection), and adult red cells (negative selection). Cells were spun at 600 rpm for 5 min at 4°C onto poly-L-lysine coated 96 well plates (NUNC, Immunosorb) 5×10^4/well in 50 µL. Fixation was carried out by adding either 50 µL of 0.1% glutaraldehyde in PBS, or 2.5% paraformaldehyde in PBS to each well, and leaving for 15 min at room temperature (Forster *et al.*). Plates were washed 3 times in PBS and blocked for 1 h with PBS containing 2% skimmed milk protein (PBS-M) at 37°C. Culture supernatants were adjusted to 2% skimmed milk protein in PBS, and then incubated with the different cell populations. Bound scFv was detected with monoclonal mouse antibody 9E10 (Sigma), which recognizes the myc tag on the scFv, followed by alkaline phosphatase conjugated goat anti-mouse immunoglobulin (Sigma) (Griffiths et al.). An antibody recognizing carcinoembryonic antigen (CEA-6, Vaughan et al.) was used as a negative control. The majority of positive clones were specific for erythroblasts. PCR fingerprinting using restriction enzyme BstN 1 was performed in the manner of Clackson et al. to identify clones with unique sequences.

Nucleic acid sequencing was performed by PCR amplification of the scFv insert using primers specific for flanking phagemid sequences. Inserts were amplified using

primers 5'-CAG GAA ACA AGC TAT GAC-3' (SEQ. ID NO:6), which sits upstream from the pelB leader sequence;, and "fdSeq1" 5'-GAA TTT TCT GTA TGA GG-3' (SEQ. ID NO:7) which sits in the 5' end of gene 3 (Marks et al. 1991). Sequencing templates were prepared using a Qiagen plasmid Midi Kit and sequenced using the ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kit with Amplitaq FS (ABI/Perkin Elmer).

Partial sequence has been obtained for Clones 22, 23, and 28. Complete sequence has been obtained for Clones 1 and 27. Surprisingly, Clone 1 comprised murine variable region sequences rather than human sequences. The nucleic acid sequence at the ends of the V_H and V_L region indicated that it had been constructed using murine-specific PCR primers. The clone had probably entered the human library or a subpopulation during replication or selection, either from contaminated glassware or from contaminated helper phage. The other selected clones all had human scFv sequences.

Clones expressing selected antibody were grown in 50-500 mL cultures, induced with 1 mM IPTG for 3-4 h, and a periplasmic extract was prepared. Immobilized metal affinity chromatography (IMAC) was used to purify the scFv using a hexahistidine tag at the carboxy terminus on NTA-Agarose (Qiagen).

6.3. EXAMPLE 3: CHARACTERIZATION OF ERYTHROBLAST SPECIFIC ANTIBODIES

Validation of binding specificity was performed by determining the ability of each antibody to identify or enrich erythroblasts from mixed cell populations.

The purified scFv was tested for its ability to label cord blood mononuclear cells. The cord blood cells were separated on Ficoll® as described in Example 1, and washed with PBE (PBS containing 0.5% BSA and 5 mM EDTA). The cells were incubated with the primary antibody at ~25 μg/mL in 100 μL PBE for 1 h. After washing, the cells were incubated with a 1/50 dilution of phycoerythrin-conjugated anti-mouse antibody (Jackson Laboratories) in 100 μL PBE, incubated, and rewashed. Fluorescence was measured on a Coulter EPICS™ XL-MCL flow cytometer.

Exemplary clones were analyzed using flow cytometry. Some clones (*e.g.*, Clones 17 and 18) have a positive shoulder beside the bulk of negative cells. Other clones (*e.g.*, Clone 23) have two obvious peaks. Clones showing no staining in the direct labeling experiment (*e.g.*, Clone 14) were judged as negative. Ficoll® purified adult and cord blood represent very heterogeneous cell populations. In this preparation, 15% of the cells were erythroid as judged by staining for glycophorin A and CD36. The labeling of rare cell populations or low-density antigen by the scFv could easily be obscured.

As an alternative to direct labeling, the antibodies were characterized by their ability to enrich for erythroblast cells by magnetic activated cell sorting (MACS). Ficoll® purified cord cells were labeled with scFv and the 9E10 secondary antibody, and enriched using paramagnetic beads coated with goat anti-mouse immunoglobulin. The studies were

conducted using either cord blood mononuclear cells, or adult whole blood or buffy coat preparations doped with cultured cord blood cells. The cells were incubated at 4°C in 200 ml of PBE containing 1/10 dilution of purified scFv. After 1 h, the cells were washed by spinning in 5 mL PBE at 390 g for 5 min. The cells were resuspended in 200 mL of PBE containing 500 ng of 9E10, incubated for 1 h at 4°C, and then washed. This was followed by incubating with microbeads coated with rat anti-mouse IgG1 (Miltenyi Biotec Ltd.) for 15 min at 4°C. The cells were washed once more, resuspended in 20 µL PBE, and loaded onto a pre-equilibrated MACS MS+/RS+ column clamped in a MiniMACS magnet. The column was washed with 2 ′ 1 mL of PBE, removed from the magnet, and the cells were eluted with 1 mL PBE pushed through with the supplied plunger. Eluted cells were either analyzed by flow cytometery, or were spun onto poly-L lysine coated slides and stained with benzidine/Wrights Giemsa stain.

In some experiments, enriched cells were analyzed using antibody specific for hemoglobin (Parsons et al.). Cell samples were spun onto poly-L-lysine coated slides (Shandon) at 800 rpm for 10 min in a Cytospin™ 3 (Shandon). Slides were left for at least 30 min to dry, and then fixed for 2 min in acetone:methanol:ethanol 3:1:1 at room temp (Thorpe et al.). Slides were washed for 5 min in PBS, and then blocked for 10 min with 10% goat serum in a humid box. They were then incubated with purified Hb-1 scFv diluted in PBS containing 10% goat serum for 1 h at room temp. The slides were washed in PBS and incubated with 9E10 antibody at 5 µg/mL for 1 h, rewashed, and incubated for 30 min with FITC-conjugated anti-mouse immunoglobulin (Sigma). After a final wash, the slides were mounted using Vectashield™ mounting medium (Vector Laboratories Inc.) and viewed using a fluorescence microscope (Olympus BX 40).

The various clones were analyzed using flow cytometry for their ability to enhance MACS enrichment. Clones 18 and 28 appear to bind erythroid cells covering a wider range of CD36 expression levels than the rest. Glycophorin A positive cells with lower levels of CD 36 expression are enriched, as shown by the high trailing edge of glycophorin-A staining cells with low CD36 expression. These are probably reticulocytes recovered on the Ficoll® gradient. Benzidine Wrights Giemsa staining showed these cells to be non-nucleated. It was concluded that the antigen is present on both immature and mature erythroblasts, and at least some reticulocytes (probably immature reticulocytes high in CD36, which are more common in cord blood), but not on mononuclear cells of adult blood. The antigen is believed to be distinct from CD36, since CD36 is expressed on monocytes which are not recognized by these clones.

Clones 17 and 20 give a lower signal by direct labeling of Ficoll® purified cord cells, but clearly enrich a population of erythroid cells. The trailing edge (high Glycophorin A and low CD36) present in the cells enriched using Clones 18 and 28 was not apparent for these clones, suggesting that antigen expression is turned off earlier. No cells

were enriched from adult blood using Clone 17. It was concluded that these clones recognize an antigen present on all the erythroid stages recognized by CD36, but distinct from that recognized by Clones 18 and 28.

Clones 4, 11, and 23 enrich a similar population of erythroid cells from cord blood as Clones 17 and 20. However, the antigens recognized by these clones are different, since they also pull out a population of CD36 positive, Glycophorin A negative cells from adult and cord blood. This additional cell population had the size and granular morphology of monocytes.

Clones 9 and 14 are similar in V_H and V_L amino acid sequences, except for the CDR3 region of V_H. They also have a similar binding profile on adult and cord blood cells. By direct labeling, Clone 9 was barely above background and Clone 14 appeared to be negative. However, both enriched erythroid cells from cord blood. Two populations of Glycophorin A positive cells can be distinguished, particularly in the case of Clone 14. As well as the main $CD36^{med}$ population observed in other clones, there is an additional $CD36^{high}$ population obtained using Clone 14.

The majority of erythroid cells in cord blood are non-nucleated reticulocytes or red cells. In cord blood, there are approximately 137 x 10^9 L^{-1} reticulocytes and 0.89 x 10^9 L^{-1} erythroblasts. Even after Ficoll® purification, there is still a significantly higher proportion of non-nucleated erythroid cells compared with nucleated erythroid cells in cord blood. The binding profile of the antibody clones across the erythroid lineage was performed using cells from erythroid culture as a more even representation of cell types arising during erythropoiesis.

6.4. EXAMPLE 4: CHARACTERIZATION OF UNIQUE ERYTHROBLAST ANTIGENS

The phagemid vectors containing the cloned scFv antibodies allow for expression either as pIII fusion protein on the surface of filamentous phage, or as soluble single-chain molecules. Staining of fetal erythroblasts was tested using both the whole phage (developed with anti-M13 antibody) and purified scFv (developed with 9E10 anti-myc). Whole phage of each of the first seven clones (Clones 1, 17, 18, 22, 23, 27, and 28) effectively stained fetal erythroblasts. The purified scFv from Clone1 showed consistent high levels of signal, but scFv from the other clones demonstrated irregular or less intense staining. This difference may be due to the amplification of the signal that occurs when using the phage but not the soluble scFv.

To increase stability and facilitate detection, the Clone 1 scFv was converted into an Fab antibody by fusing the sequences for V_H (SEQ. ID NO.8) and V_L (SEQ. ID NO.9) (a κ-chain variable region) to CH1 and C κ. Fab clone VODOX1 was used as a backbone vector. The V_H and V_L in the vector were substituted with the V_H and V_L of Clone 1.

Figure 1 shows the strategy for the substitution. An Nco1/Bste2 fragment containing V_H from clone1 was inserted into Nco1/Bste2 digested VODOX1. DNA was prepared from the cloning intermediate, and digested with ApaL1 and Xho1. Since these restriction sites are not present in Clone 1, the V_L was amplified as a PCR fragment with oligos containing these sites. The amplicon was digested and cloned as an ApaL1/Xho1 insert. The resulting clone, designated Clone 1 Fab or e-Fab, was verified by sequencing through the regions indicated by the horizontal arrows.

Clone 1 Fab was expressed and successfully used to stain erythroblasts, with detection by either anti-myc or anti-human kappa. Clone 1 Fab was prepared using immobilized metal affinity chromatography (IMAC), in which the (His)6 tag of the antibody is captured on a nickel-loaded NTA column, purified by gel filtration, and then biotinylated. This reagent was used for both cell staining and cell separations.

To identify the antigen recognized by the erythroblast antibodies, each antibody was used for affinity isolation from cell extracts. Antibody was rescued from the extract along with bound antigen by IMAC. Both the scFv and the Fab constructs contain the (His)6 tag, so either can be used. Fab was generally chosen when it was available, in part because it is expressed at much higher levels than the scFv.

Erythroblast cells were surface labeled with biotin and lysed using a Cellular Labeling and Immunoprecipitation Kit (Boehringer Mannheim) according to the manufacturer's instructions. Lysates from 10^7 cells were incubated with Clone 1, either in the form of scFv or Fab, for 2 hrs. at 4°C. Antibody-antigen complexes were then recovered with nickel-NTA resin. The resin was then eluted with 250 mM imidizole, and the eluted protein was analysed by polyacrylamide gel electrophoresis in sodium dodecyl sulfate (SDS-PAGE). Gels were stained directly, or electroblotted to PVDF membrane for Western analysis. Since proteins from the cell surface had been biotinylated, they were detected using a conjugate of streptavidin-horseradish peroxidase (HRP), followed by a chemiluminescent substrate for HRP. The advantage of this system is that antigen molecular weight information can be obtained from crude cellular extracts and crude antibody preparations.

Figure 2 shows the results of IMAC immunopurification using Clone 1 Fab. The top panel shows the silver stained gel (total protein); the two lower panels show the chemiluminescent patterns from the Western blots at two different exposures. A band corresponding to an apparent size of 90 kDa (arrow) was seen in the blot in the lane corresponding to the Fab absorbed biotinylated fetal cell extraction (Bio FC ext.), but not in the extract-only or Fab-only control lanes. A number of minor bands appear in the other lanes after a long exposure, but not with the same intensity or molecular weight as the antigen band. The 90 kDa antigen detected on the Western blot did not correspond to any

of the prominent bands on the silver stained gel. Most of the protein represented by the silver stain corresponds to material present in the Fab antibody preparation.

In an alternative purification strategy, unlabeled extract was prepared from $5x10^8$ erythroblasts (a 50-fold increase from the previous purification). The extract was combined with biotinylated Fab under conditions that permitted binding to the solubilized erythroblast antigen. The antigen-antibody complexes were then captured using streptavidin-coated Dynabeads™. Antigen was eluted and analyzed.

Figure 3 shows quantitation and molecular weight analysis of antigen obtained from preparative-scale isolations by Ni-NTA purification (upper panel) or Dynabead purification (lower panel). Apparent molecular weights were calculated from the relative mobility on a semi-log plot using six molecular weight standards between about 150 and 30 kDa. In addition to the ~90 kDa band seen previously, another specific but less intense band was seen at ~78 kDa. It is not known whether the two bands represent separate cross-reacting antigens, or whether the 78 kDa species is an alternative form of the 90 kDa species.

The Coomassie stained transfer blot shown in the lower panel was used to obtain purified material for amino acid sequencing. The band at 90 kDa was cut out from each of the four bands, and pooled, yielding approximately 3μg of purified material. No sequence was obtainable, and apparently the amino terminus of the protein is blocked.

Figure 4 shows the results of using the other cloned antibodies in the first group to purify biotinylated erythroblast membranes. Upper Panel: Silver stain; Lower Panel: Western blot from two separate experiments. The arrows indicate the position of the 90 kDa and 78 kDa bands identified by Clone 1. Clone 18 and Clone 28 appear to recognize the same bands, although the 90 kDa and 78 kDa species appear to be recognized in different proportions by the different clones. It is not clear from this experiment what antigen is recognized by Clone 17, 22, or 23. The antigens may be less abundant, or they may label with biotin less efficiently. Further characterization is performed using gel-purified scFv or Fab in a scaled-up procedure.

A summary of the Clones with established anti-erythroblast activity and their known antigen characteristics is shown in Table 2.

TABLE 2

Designation	Cell Specificity	Antigen Characteristics
Clone 1 & 27		90 kDa, 78 kDa
Clone 4	erythroblasts & monocytes	
Clone 9	early erythroblasts	

Clone 11	erythroblasts & monocytes	
Clone 13		
Clone 14	early erythroblasts	
Clone 17	early erythroblasts	
Clone 18	erythroblasts & early reticulocytes	78 kDa, (90 kDa)
Clone 20	early erythroblasts	
Clone 22		
Clone 23	erythroblasts & monocytes	
Clone 28	erythroblasts & early reticulocytes	90 kDa, 78 kDa

6.5. EXAMPLE 5: ERYTHROCYTE-SPECIFIC ANTIBODIES OBTAINED BY BOTH POSITIVE AND NEGATIVE SELECTION

Additional erythrocyte-specific antibodies were obtained using a modified scFv library. A naïve library of Fab expressing phagemids (about 10^{10} species) was converted so as to express the variable regions as scFv. The heavy chain CDR3 regions were scrambled to provide additional diversity.

Specific anti-erythroblast antibodies were obtained by a combination of positive and negative selection. Erythroblasts from fetal liver that had been cultured for 1-2 weeks were used for positive selection. Adult peripheral blood leukocytes (PBL) (pooled Ficol™ separated white cells) were used for negative selection. Briefly, the phagemid library was mixed with the erythroblasts. The bound phagemids were then recovered from the cells by adding 0.1 M glycine buffer pH 2.2, inclubating for 5 min, and centrifuging out the cells. The supernatant was neutralized by adding concentrated Tris buffer. The recovered particles were then replicated. Positive selection using the erythroblasts was repeated twice for a total of three rounds. The selected phage were then negatively selected by incubating with peripheral blood leukocytes, the supernatant was recovered. The phagemids in the supernatant were then positively selected with erythroblasts, and replicated as before. The recovered phagemids were then subjected to another round of negative and positive selection.

An aliquot of phagemids was saved from each of the selection steps for subsequent analysis. Specificity was determined by conjugating with biotin, incubating with erythroblasts, and developing with streptavidin coupled with Texas Red™.

The results of this analysis demonstrated weak staining using the phage mixture obtained after three rounds of positive selection. Much stronger selection was obtained after one or two subsequent rounds of PBL subtraction followed by erythroblast enrichment.

7. EXAMPLE: Identification of Fetal Cell Associated Transcripts

7.1. EXAMPLE 1: CONSTRUCTION OF SUBTRACTED cDNA LIBRARIES

This example is directed at identifying nucleic acid sequences that are expressed at the mRNA level in fetal cells appearing in the maternal circulation, but not in any type of circulating maternal cells that might be present in a test sample after antibody enrichment. Where the target fetal cell is a blood cell precursor such as an erythroblast, the sequence should be able to distinguish fetal cells from maternal cells at the same stage of differentiation.

To accomplish this, a number of cDNA subtraction libraries were prepared in which sequences specifically expressed in fetal cell precursors are enriched. The libraries were prepared by Suppression Subtraction Hybridization (SSH), a PCR-based method that combines normalization (the matching of mRNA levels) and subtraction (obtaining differentially expressed mRNA) in a single procedure, and requires less mRNA than other subtraction methods.

Different tissues were obtained as both the source of the differentially expressed mRNA (referred to as the "tester") and the source of baseline mRNA that would be subtracted (referred to as the "driver").

RNA Isolation: Total RNA was isolated using the TRIzol™ Reagent (Cat # 15596-026) from Life Technologies (Gaithersburg, MD) according to manufacturer's instructions. For mRNA isolation either the Straight A's™ mRNA Isolation system (Cat # 69962-1) from Novagen (Madison, WI) or the mRNA Purification System (Cat # 27-9258-02) from Pharmacia (Piscataway, NJ) was used according to manufacturer's instructions. The RNA preparations were used immediately or stored at -70°C.

cDNA synthesis: The RNA was reverse transcribed using either conventional methods as described by Klickstein, L. B., Neve, R. L., Golemis, E. A., and Gyuris, J., 1995, in Current Protocols in Molecular Biology, Ausubel, F. M., et. al., Eds, John Wiley & Sons, Inc., New York, NY, pp. 5.5.1-5.5.10 for at least 2 μg mRNA or CapFinder™ kit (Cat # K1052-1) from Clontech Laboratories, Inc. Palo Alto, CA for less than 1 μg total RNA. The CapFinder™ synthesis was performed essentially as described in the product insert; the only change was that the PCR amplification conditions were conducted with 27 cycles of 95°C for 12 seconds and 68°C for 4 minutes.

SSH: Subtraction suppression hybridization (SSH) was conducted using a PCR-Select™ kit (Cat # K1084-1) from Clontech Laboratories, Inc. Palo Alto, CA

according to manufacturers instructions except for the following modifications. The first and second hybridizations were performed for 14 and 22 hours, respectively. After adaptor extension, the PCR amplification conditions were 28 cycles of 95°C for 15 seconds, 65°C for 25 seconds and 72°C for 2 minutes. Then, the PCR product was re-amplified for 15 cycles at 94°C for 10 seconds, 68°C for 25 seconds and 72°C for 2 minutes.

The following table shows the different combinations of tester cDNA and driver cDNA that have been used for preparing suitable subtraction libraries:

TABLE 3

Summary of Subtracted cDNA Libraries

Subtracted cDNA Library Series	Tester			Driver			Biobloc k
FL10/22	10w FL	mRNA	CM	22w FL	mRNA	CM	N/A
FBP10-12/BM	10-12w FBP	mRNA	CF		mRNA	CF	N/A
HFt10-11/BM	10-11w HF	total RNA	CF	BM	total RNA	CF	20-99
HF13/BMPB	13w HF	mRNA	CF	BM & PB (1:1)	mRNA	CF	100-
HFt 12-14/24	12-14w HF	total RNA	CF	24w HF	total RNA	CF	200-
HFt 12-14/BMPB	12-14w HF	total RNA	CF	BM & PB (1:1)	total RNA	CF	400-
FBLt12-14/BMPB	12-14w FBL	total RNA	CF	BM & PB (1:1)	total RNA	CF	300-
FBLt12-14/24BMPB	12-14w FBL	total RNA	CF	24w FBL & BM & PB (1:1:1)	total RNA	CF	500-
FBt12-14/22-24	12-14w FB	total RNA	CF	22-24wFB	total RNA	CF	1000-
FBt12-14/BMG	12-14w FB	total RNA	CF	BM & g-GLOBIN (5:1)	total RNA	CF	1500-

Abbreviations:

FL: Human fetal liver
CM: cDNA from conventional methods
FBP: Porcine fetal blood
CF: cDNA from CapFinder™ synthesis
HF: Human fetal cord or circulating blood
BM: Human adult bone marrow
PB: Human adult peripheral blood

FBL: Human fetal blood from liver

The amplified cDNA was digested with Rsa I to remove the adaptor sequences, size selected on a 2% agarose gel, and subcloned into the PCR-Script™ (SK+) vector (Cat # 211189, Stratagene, La Jolla, CA). This represents a selected cDNA library by SSH.

7.2. EXAMPLE 2: IDENTIFICATION AND CHARACTERIZATION OF SHORT FRAGMENT cDNAs (TAGS) FROM SUBTRACTED cDNA LIBRARIES

Random clones from the subtracted libraries were picked and grown to provide sufficient material for characterization. The cDNA insert was PCR amplified for further testing using primers (T3 and T7) corresponding to the flanking sequences of the pCR-Script™ (SK+) vector (Cat # 211189, Stratagene, La Jolla, CA) as designated in the package insert. The PCR products were purified according to manufacturers instructions using the PCR purification kit (Cat# 28106) from Qiagen (Valencia, CA). The tags were single pass sequenced from one end using Dye-Terminator chemistry on a 377 ABI fluorescent DNA sequencer (PE Biosystems, Inc. Foster City, CA). Using either the BLAST (Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410.) or the BLAST2 algorithm (Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," Nucleic Acids Res. 25:3389-3402), the tag sequences were compared to the Genbank public database (NCBI, The National Center for Biotechnology Information) to see if the tag represented a fragment of a known gene. If not, the tag was subjected to further analysis. These clones were further analyzed first by Southern blotting of amplified cDNA fragments from a panel of 6 - 8 tissues representing mostly adult and fetal blood and blood forming organs. Briefly, 0.5 μg of amplified cDNA from relevant tissues were electrophoresed on a 1.2% agarose gel, transferred to a nylon filter and hybridized with a 32P randomly labeled (Tabor, S., Struhl, K., Scharf, S. J., and Gelfand, D. H.,1997, in Current Protocols in Molecular Biology, Ausubel, F. M., et. al., Eds, John Wiley & Sons, Inc., New York, NY, pp. 3.5.9-3.5.10) candidate tag. The hybridization conditions used were as shown by George M. Church and Walter Gilbert, 1984, Genome Sequencing, Proc. Natl. Acad. Sci. 81:1991-1995.

Tags showing evidence of specificity for fetal cells compared to adult cells were further analyzed by Southern blotting with an extended cDNA panel representing tissues from more individuals, tissue types and gestational time points. This analysis was then expanded to testing the differential tags directly on total RNA and mRNA isolated from adult and fetal blood and blood forming organs on a Northern blot. For preparation of

the Northern blot, 20 μg of total RNA or 2 μg of mRNA isolated from relevant tissues was loaded onto a denaturing agarose gel. After electrophoresis the separated RNA was then transferred to a nylon membrane and hybridized with the tag under analysis. Tags were labeled with 32P either by random priming (Tabor, S., Struhl, K., Scharf, S. J., and Gelfand, D. H.,1997, in Current Protocols in Molecular Biology, Ausubel, F. M., et. al., Eds, John Wiley & Sons, Inc., New York, NY, pp. 3.5.9-3.5.10) or by asymmetric PCR (Peter C. McCabe, Production of Single Stranded DNA by Asymmetric PCR, in PCR Protocols, A Guide to Methods and Applications Michael A. Innis et. al., Eds. 1990 Academic Press, pp 76-83. The hybridization conditions used were as described by Brown T., and .Mackey, K., 1997, Current Protocols in Molecular Biology, Ausubel, F. M., et. al., Eds, John Wiley & Sons, Inc., New York, NY, pp. 4.9.1 - 4.9.8. For each candidate tag, direct RNA analysis (by Northern blot) was performed to assess the following: validate the expression patterns seen in the cDNA blots, determine the number of mRNAs that hybridized with each promising tag, assess the size of the mRNA and to determine which strand of the tag represented the coding strand of the messenger RNA. Results of blotting experiments are shown in Figures 9-15.

Tags that passed to this point were then evaluated by mRNA Fluorescence *in situ* hybridization for their expression in the cytoplasm of individual fetal and adult erythroblast cells, and all adult end-stage nucleated peripheral blood cells.

To begin with, riboprobes representing each tag to be evaluated were synthesized and titered on fetal liver blood cells and adult peripheral blood. After validation of each of the probes' reactivity in the cytoplasm of fetal liver erythroblasts and not in adult nucleated peripheral blood cells they were tested on many other relevant tissues representing numerous individuals. These included circulating fetal erythroblasts (fetal cord blood pools from 8 - 12 week gestation human fetuses), adult erythroblasts (adult human bone marrow from iliac crest enriched for erythroblasts using an anti-transferrin receptor antibody) and adult human nucleated peripheral blood cells (white cell fraction from whole adult blood). To allow more adult erythroblasts for analysis, adult bone marrow mononuclear cells were erythroid enriched using an anti-transferrin receptor antibody attached to solid phase; this resulted in bone marrow preparations that were ~90% erythroid compared to 20-35% without enrichment.

Experiments were set up to test all these populations with each probe beginning with hybridization conditions of lower stringency and moving to higher stringency. This was done by varying experimental parameters such as: increasing the temperature of hybridization (from 55°C to 60°C), increasing the temperature of the washes (from 55°C to 60°C), decreasing the salt concentration in the washes (0.2 X SSPE to 0.05X SSPE), varying the number of washes of each type (2-3), and finally, the addition of 100 mM TMAC to the hybridization, wash and moist chamber buffers. Through all these

experimental alterations cellular morphology was preserved and by accessing the hybridization signal in multiple tissues across many stringency conditions the likelihood of a spurious positive was minimized. Hybridization signal should rise and fall according to stringency condition in a predictable manner if it is based on specific binding interactions of a nucleic acid probe with its target.

Single cell preparations were made; all according to standard cell biology methods (Cell Biology, A Laboratory Handbook 2nd Edition , J. E. Celis, Ed., Academic Press, 1998), from pooled and washed 8-12 week gestation fetal human cord blood (high in nucleated red blood cells), adult bone marrow mononuclear cells (Cat # 1M-125A, Biowhittaker, Gaithersburg, MD) enriched for the nucleated red cell precursors and progenitors with an anti-transferrin receptor antibody attached to a solid phase (A. A. Neurauter, et. al., Immunomagnetic Separation of Animal Cells, pp 197-204, in Cell Biology, 2nd Edition , J. E. Celis, Ed., Academic Press, 1998), adult peripheral blood mononuclear cells either by density gradient fractionation on Histopaque 1077 (Sigma Chemical, St. Louis, MO) according to manufacturer's instructions; or by lysis of red blood cell fraction, as described by McCoy Jr., J. P., 1998, in Current Protocols in Cytometry, Robinson J. P., et. al., Eds, John Wiley & Sons, Inc., New York, NY, pp. 5.1.2 - 5.1.3, and developing blood cells from first trimester human fetal liver. Briefly, blood cells were released from first trimester human fetal liver by floating the liver in a small petri dish in minimal essential alpha medium (Gibco BRL / Life Technologies, Grand Island, NY; Cat# 32561-037) and gently scoring the surface with a scalpel and swirling the dish to release the cells. Medium containing the fetal blood cells in the filtered through a 74 μm mesh screen (Costar / Corning, Corning, NY; Cat# 3479) into a 50 ml centrifuge tube and cells were pelleted by centrifugation in a Megafuge® 1.0R / 2.0R (Kendro Laboratory Products, Newtown, CT) at 300xg for 10 minutes at 4°C. Cells were counted and attached to coated glass microscope slides (Shandon®) using the Shandon® Cytospin 3 system and centrifugation conditions of 600 rpm, medium acceleration for 3 minutes (Shandon Lipshaw, Inc., Pittsburg, PA). Slides were handled with gloves and always in an RNAase free manner. Slides were fixed with 4% Paraformaldehyde/5% Acetic acid for 20 minutes, washed 2 times in PBS and once in Molecular Grade water (Genotech Cat # 78672), dried on a 37°C slide warmer and either used immediately or stored at -20°C in airtight containers containing dessicant. The *in situ* hybridization method used was based on Rosen B. and Beddington R., Detection of mRNA in whole mounts of mouse embryos using digoxigenin riboprobes (1994) in Methods in Molecular Biology Vol 28, Issac P. G., Ed. Humana Press Inc., Totowa, NJ. Since we were using single cells some modifications were made. Slides containing fixed cells were permeabilized with Proteinase K (0.1 μg/ml PBS) for 10 minutes at room temperature. Cells were then postfixed with 4% paraformaldehyde / 5% acetic acid for 5 minutes, rinsed in Molecular grade water twice for 5 minutes each,

incubated in 0.05% Saponin/PBS twice for 10 minutes each, and rinsed in PBS twice for 1 minute each. Hydrogen peroxidase activity was blocked using 3% Hydrogen peroxide / 1% Sodium azide / PBS for 40 minutes. Slides are rinsed in Molecular grade water twice for 1 minute, subjected to an ethanol gradient (70%, 95%, and 95% in Molecular grade water for 1 minute each) and allowed to air dry for 10 minutes.

For each tag, digoxigenin labeled riboprobes that were complementary to the sense mRNA strand were synthesized using a method based on (Signer, S. N., Digoxigenin Labeling of RNA Transcripts from Multi- and Single-Locus DNA Minisatellite Probes pp. 77- 81, in Methods in Molecular Biology Vol. 28, Ed: Issac, P. G., 1994 Humana Press, Inc., Totowa, NJ). The following modifications were made: Molecular Grade water was used throughout, the final reaction volume was increased to 23 µl, and the template was destroyed with RQ1 DNase (Promega, Madison, WI, Cat # M610A). To control the final probe size to approximately 150 -200 base pairs the precipitated riboprobes were subjected to controlled alkali hydrolysis at 65°C (Anderson, M. L. M., 1999, Nucleic Acid Hybridization, Springer-Verlag New York Inc., pp. 125). Hydrolyzed probes were then re-precipitated and 10% of each was analyzed on a denaturing agarose gel to determine the extent of the hydrolysis.

Slides were hybridized with denatured probe (mass empirically determined) in 20 µl of hybridization buffer composed of 50% deionized formamide, 5 X SSPE, 1 X Denhardt's solution, 50 µg/ml Yeast tRNA, 50 µg/ml denatured Salmon sperm DNA, 10% Dextran sulphate, 0.2% CHAPS and Molecular Grade water. Slides were placed in a sealed moist chamber (50% deionized formamide / 5X SSPE) and incubated overnight at 55°C, 58°C or 60°C depending on the experimental set up. The following day the slides were washed with two to three washes of 0.2X SSPE / 0.05% Saponin pre-warmed to either at 55°C, 58°C or 60°C (10 minutes each), then with two to three washes of 0.1X SSPE / 0.05% Saponin pre-warmed to either at 55°C, 58°C or 60°C (10 minutes each), then incubated with 0.2X SSPE (10 minutes at room temperature), and then rinsed in 1X blocking buffer (Fluorescent Antibody Enhancer Set for DIG Detection, Roche Molecular Biochemicals, Cat# 1768506) for 30 minutes at room temperature. In some cases, a third wash set of of 0.05X SSPE / 0.05% Saponin pre-warmed to either at 55°C, 58°C or 60°C (10 minutes each) was done. The temperature of the washes was performed at either the temperature of the hybridization or higher; up to 60°C and with some agitation. Temperatures greater than 60°C resulted in poor morphology and impaired data analysis. The Digoxigenin label on the bound riboprobe was detected using two rounds of the first two reagents from Fluorescent Antibody Enhancer Set for DIG Detection, Roche Molecular Biochemicals, Cat# 1768506 according to manufacturer's instructions, then followed by a sheep anti-DIG Fab labeled with Horseradish peroxidase (Roche Molecular Biochemicals, Cat# 1207733) diluted 1:500 in PBS. Finally, TSA™-Fluorescein (NEN™ Life Science

Products, Inc., Boston, MA) was added as a substrate for HRP resulting in activation and covalent deposition of Fluorescein -tyramide according to manufacturer's instructions. After washing the slides in PBS, the cells were counter stained with 4',6-Diamidino-2-phenylindole (DAPI, Molecular Probes, Inc., Eugene, OR; Cat # D-1306). Slides were washed twice in PBS and once in Molecular Grade water, dried in 38°C oven and mounted with ProLong™ mounting medium (Molecular Probes, Inc., Eugene, OR; Cat # P7481). Slides were stored in the dark at 4°C until fluorescent microscopic analysis. Cellular epifluorescent signal was visualized with a stationary multi-band beamsplitter and emitter mounted in the body of a Zeiss Axioskop microscope (Zeiss; Thornwood, NY) with single and multi-band excitation filters fitted in a ludl filter wheel (Ludl; Hawthorne, NY) using Chroma's 83000 filter set (Chroma; Brattleboro, VT), an AttoArc power source (Zeiss; Thornwood, NY), a charged coupled device (Photometrics, Tucson, AZ, Model SenSys™) and QUIPS™ SmartCapture™ image capture software, Version 3.1.2 (Vysis, Inc., Downer's Grove, IL) resident on a Macintosh Power PC G3/400 (Apple Computer; Cupertino, CA).

The *in situ* data illustrated in Figures 9-16 show the results of each tag (probe) with multiple preparations (6 - 7 each) of fetal erythroblasts and adult erythroblasts. For all the tags, green signal (fluorescein) representing specific hybridization of the probe is seen in the cytoplasm of the fetal erythroblast cells to a higher degree than in the adult erythroblasts. For orientation, the nucleus is counter stained blue with the nuclear dye, DAPI. For direct comparison of signals in fetal erythroblasts vs. adult erythroblasts and adult peripheral white cells, multiple sets of these tissue types were performed in the same experiment. They were all treated identically. To create these figures, the comparable images were pulled from the server and a screen shot was captured for the images shown for each tag (probe). The adult nucleated white blood cells done in the same experiments were all negative for signal in the cytoplasm and the data is not shown.

7.3. EXAMPLE 3: FULL LENGTH cDNA LIBRARY SCREENING

Since the nine tags detailed above showed fetal erythroid specificity and that the tags identified from our subtracted cDNA libraries represent only fragments (Rsa I digested short fragments) of the entire mRNA, full length libraries were screened to pull the full length mRNA for each of the nine tags.

Full length cDNA library construction: The construction of the full-length cDNA libraries were performed according to manufacturer's instructions. cDNA synthesis from mRNA greater than 5 μg was made by using ZAP-cDNA ® Synthesis Kit (Cat # 200400, Stratagene, La Jolla, CA). From total or mRNA less than 1ug, the cDNA synthesis was performed with CapFinder™ PCR cDNA library construction kit (Cat# K1051-1) from

Clontech Laboratories, Inc. Palo Alto, CA according to manufacturer's instructions except for the following modifications. After second strand cDNA synthesis, cDNA was size selected by low melting agarose gel -electrophoresis, followed by phenol/chloroform extraction. The cDNA was then ethanol precipitated, washed and resuspended in water and ligated to EcoRI digested and CIAP-treated lambda ZAP® II vector (Cat # 236211, Stratagene, La Jolla, CA). The ligated cDNA was packaged and used to infect *E. coli* XL-1 Blue MRF'. The titer of each library was more than $5x10^6$ plaque forming units (pfu). cDNA libraries were then amplified by either PCR or phage infection into bacteria.

Full length screening by plaque hybridization: The short-fragment tags from the subtracted cDNA libraries were used as probes to isolated full-length sequences of the tags. The procedure was according to Quertermous, T., 1996, in Current Protocols in Molecular Biology, Ausubel, F. M., et. al., Eds, John Wiley & Sons, Inc., New York, NY, pp. 6.1.1-6.1.4. Phagemid containing cDNA inserts were excised from lambda ZAP® II vector (Cat # 236211, Stratagene, La Jolla, CA) by *in vivo* excision. The clones containing the longest cDNA inserts were sequenced and compared using the BLAST2 algorithm (Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," Nucleic Acids Res. 25:3389-3402), to GenBank (release 117; NCBI,Bethesda, MD) / EMBL (release 62; EBI, Cambridge, UK), human EST (release 117, NCBI, Bethesda, MD), LifeSeq® Gold full length and component sequences Version 5.1, May 2000 release (Incyte Genomics, Palo Alto, CA) for DNA identity.

The following table shows the relationship between the tags, their full-length genes and their respective lengths in base pairs (bp).

TABLE 4

Original Tag	SIZE of Tag (bp)	SEQ. ID NO.	Name of FL	SIZE of FL (bp)	SEQ. ID NO.
1503-7E	711	10	J42-4d	3194	11
305-4G	378	15	K1-1a	2256	16
597-10C	1068	21	NT7-T3	2186	22
334-2C	1159	24	O19r-T3	3561	25
332-9E	1126	27	P60-1a	3215	28
305-9E	1014	31	R5'-T3	3230	32
369-8G	454	34	U2f-T3	3661	35
305-6G	1095	36	L15-1a	2103	37

7.4. EXAMPLE 4: IDENTIFICATION OF ALTERNATIVELY SPLICED FORMS OF TAG-RELATED FULL-LENGTH GENES

The initial purpose of this work was to confirm that the full-length genes identified by plaque hybridization contained their entire 5' and 3' ends. As a function of this work, alternatively spliced forms were found for almost all of the eight tag-related full length gene sequences. Some were only alternative polyadenylation sites so are not mentioned here. However, others that reflect sequence variation are included and are referred to as tag-related splice variants. Since the variant regions may be even more tissue specific and useful for designing cellular identification probes, we decided to extensively study the alternatively spliced forms for each of the tag-related full-length genes. 5' and 3' RACE (rapid amplification of cDNA ends) was conducted using a SMART™ RACE cDNA Amplification Kit (Cat# K1811-1) from Clontech Laboratories, Inc. Palo Alto, CA according to manufacturer's instructions. PCR products were cloned into PCR-Script™ (SK+) vector (Cat # 211189, Stratagene, La Jolla, CA) and sequenced with T3 and T7 primers using Dye-Terminator chemistry on a 377 ABI fluorescent DNA sequencer (PE Biosystems, Inc. Foster City, CA). Sequences were then compared with tags and tag-related full length genes by using Sequencher 3.1 and 4.0 softwares (Gene Codes Corp, Ann Arbor, MI). For tag 305-4G (SEQ. ID NO.15), 334-2C (SEQ. ID NO.24), 332-9E (SEQ. ID NO. 27) and 305-6G (SEQ. ID NO. 36) we also used plaque hybridization and PCR to increase the likelihood for obtaining variants.

The tags, and the respective tag-related full length genes and tag-related splice variants are listed in the table below.

TABLE 5

Name Tag	Name of FL	SEQ. ID NO.	Name of alternative spliced form	SEQ ID NO.	Method used for obtaining alternative spliced form
1503-7E	J42-4d	11	J2r(3)	12	5' RACE
			J2r(12)	13	5' RACE
			J2r(13)	14	5' RACE

305-4G	K1-1a	16	K2r/1f(50)	17	PCR
			K2r/1f(59)	18	PCR
			K(1)157-2A	19	plaque hybridization
			K3r(HIGH)76	20	3' RACE
597-10C	NT7-T3	22	N9r/Mf	23	5' RACE
334-2C	O19r-T3	25	O1-1a	26	plaque hybridization
332-9E	P60-1a	28	P1-1a	29	plaque hybridization
			P3r(9)	30	5' RACE
305-9E	R5'-T3	32	R6r/1-6H	33	5' RACE
369-8G	U2f-T3	35	-		-
305-6G	L15-1a	37	-		-

All tags, tag-related full length genes and tag-related splice variants can be used for the purposes of this invention. Specifically, all of these nucleic acid sequences are useful to distinguish fetal cells from maternal cells.

8. Example: Detection of Tag Sequences by *In Situ* Hybridization

Dioxygenin (DIG)-labeled riboprobes corresponding to tags identified in the screening methods of Section 7, *supra* (SEQ ID NOs:10, 15, 21, 24, 27, 31, 34, 36 and 41) were synthesized and titered on fetal liver blood and adult peripheral blood. After validation of each of the probes' reactivity in the cytoplasm of fetal liver erythroblast and not in adult nucleated peripheral blood cells, the probes were tested in other relevant tissues, such as circulating fetal erythroblasts (fetal cord blood pools from 8-12 week gestation human fetuses), adult erythroblast (adult bone marrow from iliac crest enriched for erythroblast using an anti-transferrin receptor antibody attached to a solid phase, which resulted in bone marrow preparations that were ~90% erythroid compared to 20-35% prior to enrichment.) and adult nucleated peripheral blood cells (white cell fraction from whole adult blood). Experiments were set up to test all these population with each probe, beginning with hybridization conditions of lower stringency and moving to higher stringency conditions. This was accomplished by varying experimental parameters such as: the temperature of hybridization (from 55°C to 60°C), the temperature of the washes (from 55°C to 60°C), the salt concentration in the washes (0.2 X SSPE to 0.05X SSPE), the number of washes of each type (two - three times), and finally, the addition of 100 mM TMAC to the hybridization, wash and moist chamber buffers.

8.1. MATERIALS AND METHODS

Mononuclear fractions of adult bone marrow, peripheral blood and washed 8-12 week fetal cord blood were prepared according to standard cell biology methods (Celis, 1998, " Cell Biology, A Laboratory Handbook", 2nd Ed., Academic Press, San Diego, CA). The *in situ* hybridization method used was based on Rosen B. and Beddington R. (Rosen and Beddington, 1994, Detection of mRNA in whole mounts of mouse embryos using digoxigenin riboprobes. *In* "Methods in Molecular Biology" (Isaac, P. G., Ed), Vol 28, pp. 201-208, Humana Press Inc., Totowa, NJ) with some modifications. Paraformaldehyde (4%) / acetic acid (5%) fixed cells on slides were permeabilized with Proteinase K (0.1 mg/ml in PBS) for 10 minutes at room temperature. Cells were then post fixed with 4% paraformaldehyde/5% acetic acid for 5 minutes and incubated in 0.05% Saponin/PBS twice for 10 minutes each. Hydrogen peroxidase activity was blocked using 3% Hydrogen peroxide / 1% Sodium azide / PBS for 40 minutes. Digoxigenin labeled riboprobes were synthesized using a method based on both Signer's protocol (Signer, 1994, Digoxigenin labeling of RNA transcripts from multi- and single - locus DNA minisatellite probes. *In* "Methods in Molecular Biology" (Isaac, P. G., Ed), Vol 28, pp. 77-81, Humana Press Inc., Totowa, NJ) and controlled alkali hydrolysis at 65°C (Anderson, 1999 "Nucleic Acid Hybridization"(Rickwood, D., Ed.) p.125, Springer-Verlag New York Inc., New York). In the analyses the average hydrolyzed non-radioactive riboprobe was targeted to be between 150 and 200 nucleotides in length. In addition, riboprobes synthesized using Digoxigenin-11-UTP (Roche Molecular Biochemicals) on average contain one Digoxigenin-11-UTP every twenty nucleotides resulting in an average of 7 digoxigenin labels per 150 nucleotides to 10 digoxigenin labels per 200 nucleotides of hydrolyzed non-radioactive riboprobe fragment.

Slides were incubated overnight at either 55°C, 58°C or 60°C with a mixture of 20 ml of each riboprobe in hybridization buffer (50% deionized formamide, 5 x SSPE, 1 x Denhardt's solution, 50 mg /ml Yeast tRNA, 50 mg/ml denatured salmon sperm DNA, 10% Dextran sulphate, and 0.2% CHAPS). Slides were washed minimally with 0.2x SSPE/0.05% Saponin three times and 0.1x SSPE/0.05% Saponin twice at the same temperature as the hybridization. Temperature and ionic conditions of the hybridizations and wash steps did not go higher than 60°C or above 3 times with 0.1x SSPE to preserve cellular morphology.

The digoxigenin labeled riboprobes were detected using two rounds of the first two reagents from the Fluorescent Antibody Enhancer Set for DIG Detection (Roche Molecular Biochemicals), namely a mouse IgG_1 monoclonal antibody and a digoxigenin labeled anti-mouse IgG $F(ab')_2$ fragment, followed by incubation with a sheep anti-DIG horseradish peroxidase labeled Fab (Roche Molecular Biochemicals) and the HRP substrate TSA™-Fluorescein (NEN™ Life Science Products, Inc., Boston, MA). After washing the

slides with PBS the cells were counter stained with 4',6-Diamidino-2-phenylindole (DAPI), washed again with PBS, followed by water, then dried and mounted with ProLong™ mounting medium (Molecular Probes, Inc., Eugene). Cellular epifluorescent signal was visualized with a stationary multi-band beamsplitter and emitter mounted in the body of a Zeiss Axioskop epifluorescence microscope (Zeiss; Thornwood, NY) with single and multi-band excitation filters fitted in a Ludl filter wheel (Ludl; Hawthorne, NY) using Chroma's 83000 filter set (Brattleboro, VT) and equipped with a CCD camera (Photometrics, Tucson, AZ, ModelSenSys™). Images were captured using QUIPS™ SmartCapture™ image capture software, Version 3.1.2 (Vysis, Inc., Downer's Grove, IL).

8.2. RESULTS

Using the above visualization and detection systems cells were determined to be positive for riboprobe *in situ* hybridization analysis when the FITC signal (green) at least two fold greater than background and the cell contained a nucleus (blue). Negative cells were cells that either lacked FITC signal (green) within the cellular membrane borders (red) and/or the cell lacked a nucleus (blue). Thus captured images of cells positive for the non-radioactive riboprobe would have within its cellular membranes a nucleus that was labeled blue and peroxidase-antibody cascades that were labeled green.

By assessing the hybridization tissues across many conditions, the tags of SEQ ID NOs:10, 15, 21, 24, 27, 31, 34, 36 and 41 were shown to selectively or specifically hybridize to erythroblasts, as the hybridization signal varied in a predictable manner with the stringency of hybridization. While tag 252 (SEQ ID NO:39) was not found in adult peripheral blood cells, its expression level in adult and fetal erythroblasts was indistinguishable.

9. SPECIFIC EMBODIMENTS, CITATION OF REFERENCES

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

Various references, including patent applications, patents, and scientific publications, are cited herein; the disclosure of each such reference is hereby incorporated herein by reference in its entirety.

WHAT IS CLAIMED IS:

1. A method for detecting a fetal cell in a maternal blood sample, comprising the steps of:

(a) contacting said maternal blood sample with a first probe, said first probe comprising:

(i) a nucleotide sequence corresponding to SEQ ID NO: 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42,

(ii) a nucleotide sequence having at least 90% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or

(iii) a nucleotide sequence having at least 80% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42,

which first probe selectively or specifically hybridizes to mRNA in fetal cells if present in the maternal blood sample; and

(b) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell.

2. The method of claim 1, wherein the first probe specifically hybridizes to fetal cells.

3. The method of claim 1, wherein the first probe selectively hybridizes to fetal cells.

4. The method of claim 1 in which the first probe comprises a label.

5. The method of claim 4 in which the label is a radioactive label, a fluorescent label, a colorimetric reagent, or an enzyme.

6. The method of claim 1, wherein the nucleotide sequence is 20-30 nucleotides in length.

7. The method of claim 1, wherein the nucleotide sequence is 30-40 nucleotides in length.

8. The method of claim 1, wherein the nucleotide sequence is 40-60 nucleotides in length.

9. The method of claim 1, wherein the nucleotide sequence is 60-80 nucleotides in length.

10. The method of claim 1, wherein the nucleotide sequence is 80-100 nucleotides in length.

11. The method of claim 1, wherein the nucleotide sequence is 100-150 nucleotides in length.

12. The method of claim 1, wherein the nucleotide sequence is 150-200 nucleotides in length.

13. The method of claim 1, wherein the nucleotide sequence is greater than 200 nucleotides in length.

14. The method of claim 1, wherein the probe is less than 40 nucleotides in length.

15. The method of claim 1, wherein the probe is less than 50 nucleotides in length.

16. The method of claim 1, wherein the probe is less than 100 nucleotides in length.

17. The method of claim 1, wherein the probe is less than 200 nucleotides in length.

18. The method of claim 1, wherein the probe is less than 300 nucleotides in length.

19. The method of claim 1, wherein the probe is less than 400 nucleotides in length.

20. The method of claim 1, wherein the probe is less than 500 nucleotides in length.

21. The method of claim 1, wherein the probe is less than 1,000 nucleotides in length.

22. The method of claim 1, wherein the probe is less than 2,000 nucleotides in length.

23. The method of claim 1, further comprising, prior to step (b), contacting said maternal blood sample with a second probe which selectively or specifically hybridizes to fetal cells if present in the maternal blood sample.

24. The method of claim 23, further comprising detecting a cell in said maternal blood sample which comprises mRNA that hybridizes to the second probe.

25. The method of claim 24, wherein the first and second probe are labeled with the same type of label.

26. The method of claim 23, wherein the first and second probes correspond to the same mRNA.

27. The method of claim 23, wherein the first and second probes correspond to different mRNAs.

28. The method of claim 1, further comprising, prior to step (a), immunoenriching the maternal blood sample for fetal cells.

29. The method of claim 28, wherein immunoenriching the maternal blood sample comprises:

(a) contacting the maternal blood sample with an antibody that selectively or specifically binds to fetal cells in the maternal blood sample; and

(b) separating cells in the maternal blood sample that bind to the antibody from cells that do not bind to the antibody, thereby immunoenriching the maternal blood sample for fetal cells.

30. The method of claim 28, wherein immunoenriching the maternal blood sample comprises:

(a) contacting the maternal blood sample with an antibody that selectively or specifically binds to maternal cells in the maternal blood sample; and

(b) separating cells in the maternal blood sample that do not bind to the antibody from cells that bind to the antibody, thereby immunoenriching the maternal blood sample for fetal cells.

31. The method of claim 1, wherein the probe is an RNA probe.

32. The method of claim 1, wherein the probe is a DNA probe.

33. The method of claim 1, wherein the fetal cell is an erythroblast.

34. The method of claim 1, wherein the fetal cell is a trophoblast.

35. The method of claim 1, wherein nucleotide sequence has at least 95% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

36. The method of claim 35, wherein nucleotide sequence has 100% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

37. The method of claim 1, wherein nucleotide sequence has at least 85% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

38. The method of claim 37, wherein the nucleotide sequence has at least 90% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

39. The method of claim 38, wherein the nucleotide sequence has at least 95% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14,

15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

40. The method of claim 39, wherein the nucleotide sequence has 100% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

41. A method for diagnosing an abnormality in a fetal cell, comprising the steps of:

(a) contacting a maternal blood sample with a first probe, said first probe comprising:

(i) a nucleotide sequence corresponding to SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42,

(ii) a nucleotide sequence having at least 90% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or

(iii) a nucleotide sequence having at least 80% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42,

which first probe selectively or specifically hybridizes to mRNA in a fetal cell if present in the maternal blood sample;

(b) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell; and

(c) if a fetal cell is detected, determining whether the abnormality exists in said fetal cell, thereby diagnosing the abnormality.

42. The method of claim 41, wherein the first probe specifically hybridizes to the fetal cell.

43. The method of claim 41, wherein the first probe selectively hybridizes to the fetal cell.

44. The method of claim 41 in which the first probe comprises a label.

45. The method of claim 44 in which the label is a radioactive label, a fluorescent label, a colorimetric reagent or an enzyme.

46. The method of claim 41, wherein the nucleotide sequence is 20-30 nucleotides in length.

47. The method of claim 41, wherein the nucleotide sequence is 30-40 nucleotides in length.

48. The method of claim 41, wherein the nucleotide sequence is 40-60 nucleotides in length.

49. The method of claim 41, wherein the nucleotide sequence is 60-80 nucleotides in length.

50. The method of claim 41, wherein the nucleotide sequence is 80-100 nucleotides in length.

51. The method of claim 41, wherein the nucleotide sequence is 100-150 nucleotides in length.

52. The method of claim 41, wherein the nucleotide sequence is 150-200 nucleotides in length.

53. The method of claim 41, wherein the nucleotide sequence is greater than 200 nucleotides in length.

54. The method of claim 41, wherein the probe is less than 40 nucleotides in length.

55. The method of claim 41, wherein the probe is less than 50 nucleotides in length.

56. The method of claim 41, wherein the probe is less than 100 nucleotides in length.

57. The method of claim 41, wherein the probe is less than 200 nucleotides in length.

58. The method of claim 41, wherein the probe is less than 300 nucleotides in length.

59. The method of claim 41, wherein the probe is less than 400 nucleotides in length.

60. The method of claim 41, wherein the probe is less than 500 nucleotides in length.

61. The method of claim 41, wherein the probe is less than 1,000 nucleotides in length.

62. The method of claim 41, wherein the probe is less than 2,000 nucleotides in length.

63. The method of claim 41, further comprising, prior to step (b), contacting said maternal blood sample with a second probe which selectively or specifically hybridizes to fetal cells if present in the maternal blood sample.

64. The method of claim 63, further comprising detecting a cell in said maternal blood sample which comprises mRNA that hybridizes to the second probe.

65. The method of claim 64, wherein the first and second probe are labeled with the same type of label.

66. The method of claim 63, wherein the first and second probes correspond to the same mRNA.

67. The method of claim 63, wherein the first and second probes correspond to different mRNAs.

68. The method of claim 41, further comprising, prior to step (a), enriching the maternal blood sample for fetal cells prior.

69. The method of claim 68, wherein immunoenriching the maternal blood sample comprises:

(a) contacting the maternal blood sample with an antibody that selectively or specifically binds to fetal cells in the maternal blood sample; and

(b) separating cells in the maternal blood sample that bind to the antibody from cells that do not bind to the antibody, thereby immunoenriching the maternal blood sample for fetal cells.

70. The method of claim 68, wherein immunoenriching the maternal blood sample comprises:

(a) contacting the maternal blood sample with an antibody that selectively or specifically binds to maternal cells in the maternal blood sample; and

(b) separating cells in the maternal blood sample that do not bind to the antibody from cells that bind to the antibody, thereby immunoenriching the maternal blood sample for fetal cells.

71. The method of claim 41, wherein the probe is an RNA probe.

72. The method of claim 41, wherein the probe is a DNA probe.

73. The method of claim 41, wherein nucleotide sequence has at least 95% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

74. The method of claim 73, wherein nucleotide sequence has 100% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

75. The method of claim 41, wherein nucleotide sequence has at least 85% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

76. The method of claim 75, wherein the nucleotide sequence has at least 90% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14,

15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

77. The method of claim 76, wherein the nucleotide sequence has at least 95% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

78. The method of claim 77, wherein the nucleotide sequence has 100% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

79. The method of claim 41, wherein the abnormality is a chromosomal abnormality.

80. The method of claim 79, wherein the chromosomal abnormality is an aneuploidy.

81. The method of claim 80, wherein the aneuploidy is trisomy 13, trisomy 21, or Klinefelter syndrome.

82. The method of claim 41, wherein the abnormality is a single gene disorder.

83. The method of claim 82, wherein the single gene disorder is a deletion, insertion or substitution disorder.

84. The method of 83, wherein the single gene disorder is spina bifida, sickle-cell anemia, a thalassemia, Marfan Syndrome, Duchenne Muscular Dystrophy, or cystic fibrosis.

85. The method of claim 41, wherein the abnormality is a nucleotide triplet expansion in the gene.

86. The method of claim 85, wherein the gene is the Fragile X Syndrome gene, the Friedreich's ataxia gene, the myotonic dystrophy gene, or the Huntington's disease gene.

87. The method of claim 41, wherein the fetal cell is an erythroblast.

88. The method of claim 41, wherein the fetal cell is a trophoblast.

89. The method of claim 41, wherein determining whether the abnormality exists in said fetal cell comprises the steps of:

(d) contacting the maternal blood sample with a diagnostic probe under conditions that allow hybridization of the diagnostic probe to a diagnostic target sequence in the fetal cell, wherein the manner of hybridization of the diagnostic probe to the diagnostic target sequence is indicative of whether the abnormality exists in the fetal cell; and

(e) determining the manner in which the diagnostic probe hybridizes to the target sequence, thereby determining whether the abnormality exists in the fetal cell.

90. The method of claim 89, wherein steps (a) and (d) are performed simultaneously.

91. The method of claim 90, further comprising, prior to steps (a) and (d), contacting the maternal blood sample with a first fixative, wherein the first fixative comprises 4% formalin and has a pH of 6-8.

92. The method of claim 91, further comprising, following steps (a) and (d), contacting the maternal blood sample with a second fixative following, wherein the second fixative comprises 4% formalin and has a pH of less than 6.

93. A method for identifying a nucleic acid useful as a probe for fetal cells in the maternal circulation, comprising the steps of:

(a) performing differential expression analysis on RNA or cDNA obtained from fetal liver myeloid cells of less than 22 weeks of gestation relative to RNA or cDNA obtained from more mature liver or non-liver myeloid cells; and

(b) identifying an RNA or cDNA species that is selectively or specifically expressed in the fetal liver myeloid cells,

wherein the RNA or cDNA species that is selectively or specifically expressed in the fetal liver myeloid cells is useful as a probe for fetal cells in the maternal circulation.

94. The method of claim 93, wherein the fetal liver is human fetal liver.

95. The method of claim 93, wherein the fetal liver myeloid cells are obtained before 20 weeks of gestation.

96. The method of claim 95, wherein the fetal liver myeloid cells are obtained between 10 and 15 weeks of gestation.

97. The method of claim 95, wherein the more mature myeloid cells are fetal cord blood cells obtained after 22 weeks of gestation, fetal peripheral blood cells obtained after 22 weeks of gestation, or fetal liver myeloid cells obtained after about 22 weeks of gestation.

98. The method of claim 93, wherein the mature myeloid cells are adult bone marrow cells or adult peripheral blood cells.

99. The method of claim 93, wherein the differential expression analysis comprises subtraction suppression hybridization (SSH).

100. A kit comprising in one or more containers

(a) a first probe, said first probe comprising:

(i) a nucleotide sequence corresponding to SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42;

(ii) a nucleotide sequence having at least 90% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42; or

(iii) a nucleotide sequence having at least 80% sequence identity to at least 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42,

which first probe selectively or specifically hybridizes to mRNA in fetal cells if present in a maternal blood sample; and

(b) instructions for diagnostic use or a label indicating regulatory approval for diagnostic use.

101. The kit of claim 108, further comprising an antibody that selectively or specifically binds to fetal cells in a maternal blood sample.

102. The kit of claim 108, further comprising a second probe that selectively or specifically hybridizes to mRNA in fetal cells if present in a maternal blood sample.

103. The kit of claim 102, wherein the second probe comprises:

(a) a nucleotide sequence corresponding to SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42;

(b) a nucleotide sequence having at least 90% sequence identity to at least 20 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42; or

(c) a nucleotide sequence having at least 80% sequence identity to 40 consecutive nucleotides of SEQ ID NO:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42.

104. The kit of claim 102, wherein the first probe and the second probe correspond to the same mRNA.

105. The kit of claim 102, wherein the first probe and the second probe correspond to different mRNAs.

106. The kit of claim 102, wherein the first probe and the second probe are labeled with the same type of label.

107. A method for detecting a fetal cell in a maternal blood sample, comprising the steps of:

(a) performing differential expression analysis on RNA or cDNA obtained from fetal liver myeloid cells relative to RNA or cDNA obtained from mature myeloid cells;

(b) identifying an RNA or cDNA species that is selectively or specifically expressed in the fetal liver myeloid cells, thereby identifying an RNA or cDNA species that is useful as a probe for fetal cells in the maternal circulation;

(c) contacting the maternal blood sample with a probe comprising a nucleotide sequence corresponding to all or a portion of the RNA or cDNA of step (b); and

(d) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell.

108. A method for diagnosing an abnormality in a fetal cell, comprising the steps of:

(a) performing differential expression analysis on RNA or cDNA obtained from fetal liver myeloid cells relative to RNA or cDNA obtained from mature myeloid cells;

(b) identifying an RNA or cDNA species that is selectively or specifically expressed in the fetal liver myeloid cells, thereby identifying an RNA or cDNA species that is useful as a probe for fetal cells in the maternal circulation;

(c) contacting the maternal blood sample with a probe comprising a nucleotide sequence corresponding to all or a portion of the RNA or cDNA of step (b);

(d) identifying whether or not said maternal blood sample comprises a cell that comprises mRNA that detectably hybridizes to the first probe, thereby detecting whether or not said maternal blood sample contains a fetal cell; and

(e) if the maternal blood sample contains a fetal cell, determining whether the abnormality exists in said fetal cell, thereby diagnosing the abnormality.

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

(a) a nucleic acid molecule having a nucleotide sequence which is at least 90% identical to the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or a complement thereof;

(b) a nucleic acid molecule comprising at least 15 nucleotide residues and having a nucleotide sequence identical to at least 15 consecutive nucleotide residues of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16,

17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or a complement thereof,

(c) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78;

(d) a nucleic acid molecule which encodes a fragment at least 10 consecutive amino acid residues of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78;

(e) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of any SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78; wherein the fragment comprises consecutive amino acid residues corresponding to at least half of the full length of any of said SEQ ID NOs; and

(f) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78, wherein the nucleic acid molecule hybridizes with a nucleic acid molecule consisting of the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42 under stringent conditions, or a complement thereof.

110. The isolated nucleic acid molecule of claim 109, which is selected from the group consisting of:

(a) a nucleic acid having the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or a complement thereof; and

(b) a nucleic acid molecule which encodes a polypeptide having the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66,

67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78, or a complement thereof.

111. The nucleic acid molecule of claim 109, further comprising a vector nucleic acid sequence.

112. The nucleic acid molecule of claim 109, further comprising a nucleic acid sequence encoding a heterologous polypeptide.

113. A host cell which contains the nucleic acid molecule of claim 109.

114. The host cell of claim 113 which is a mammalian host cell.

115. A non-human mammalian host cell containing the nucleic acid molecule of claim 109.

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

(a) a fragment of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78;

(b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes with a nucleic acid molecule consisting of the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42 under stringent conditions, or a complement thereof; and

(c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 90% identical to a nucleic acid consisting of the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or a complement thereof.

117. The isolated polypeptide of claim 116 having the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78.

118. The polypeptide of claim 116, wherein the amino acid sequence of the polypeptide further comprises heterologous amino acid residues.

119. An antibody which selectively binds with the polypeptide of claim 116.

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

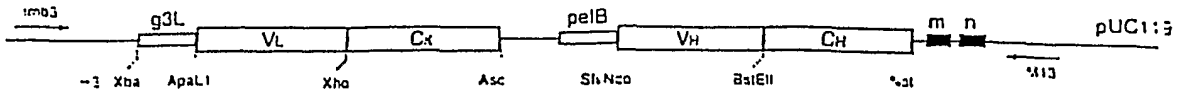
(a) a polypeptide comprising the amino acid sequence of any of SEQ ID NOs: 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78;

(b) a polypeptide comprising a fragment of at least 10 contiguous amino acids of the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78; and

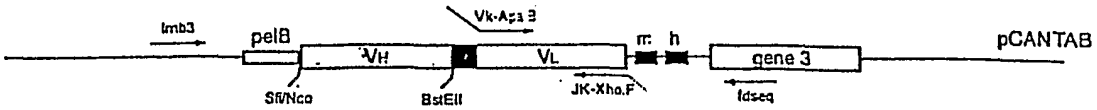
(c) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, or 78, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes with a nucleic acid molecule consisting of the nucleotide sequence of any of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41 or 42, or a complement thereof under stringent conditions;

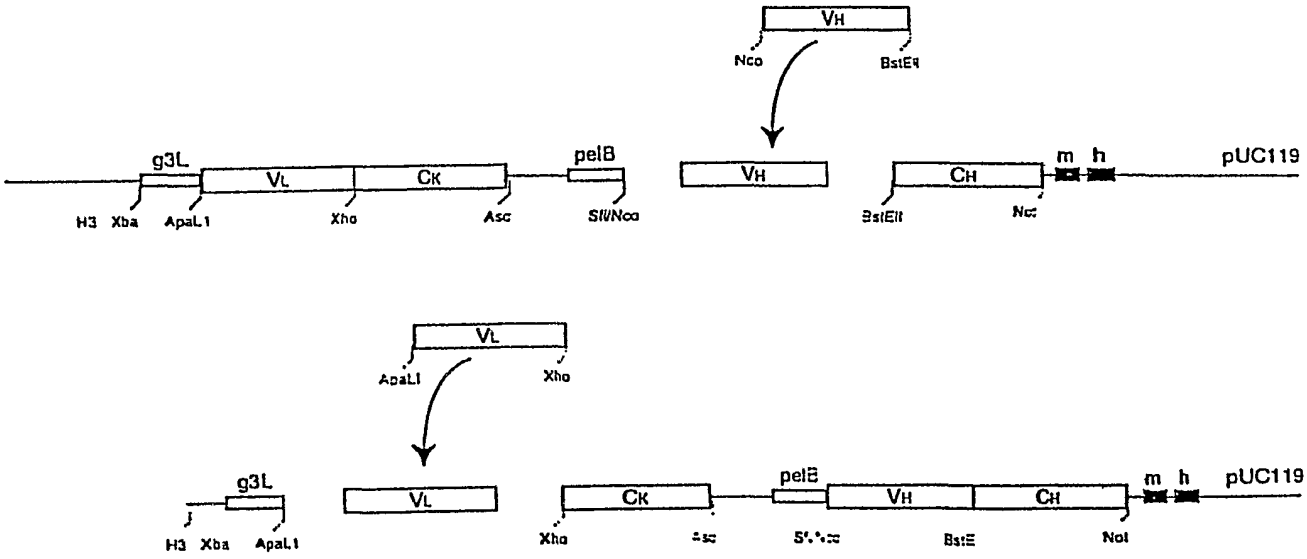
the method comprising culturing the host cell of claim 113 under conditions in which the nucleic acid molecule is expressed.

anti-DOX Fab



anti-E scFv





anti-E Fab

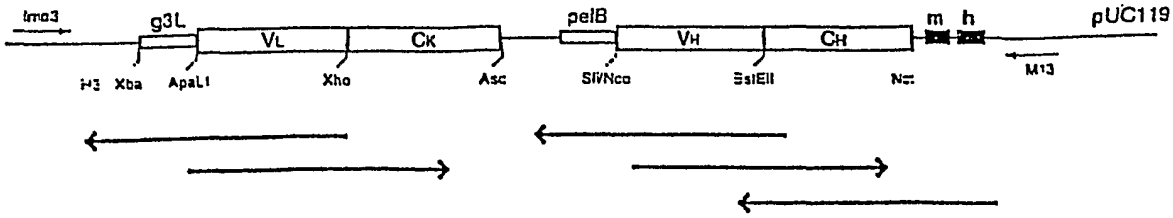


FIGURE 1

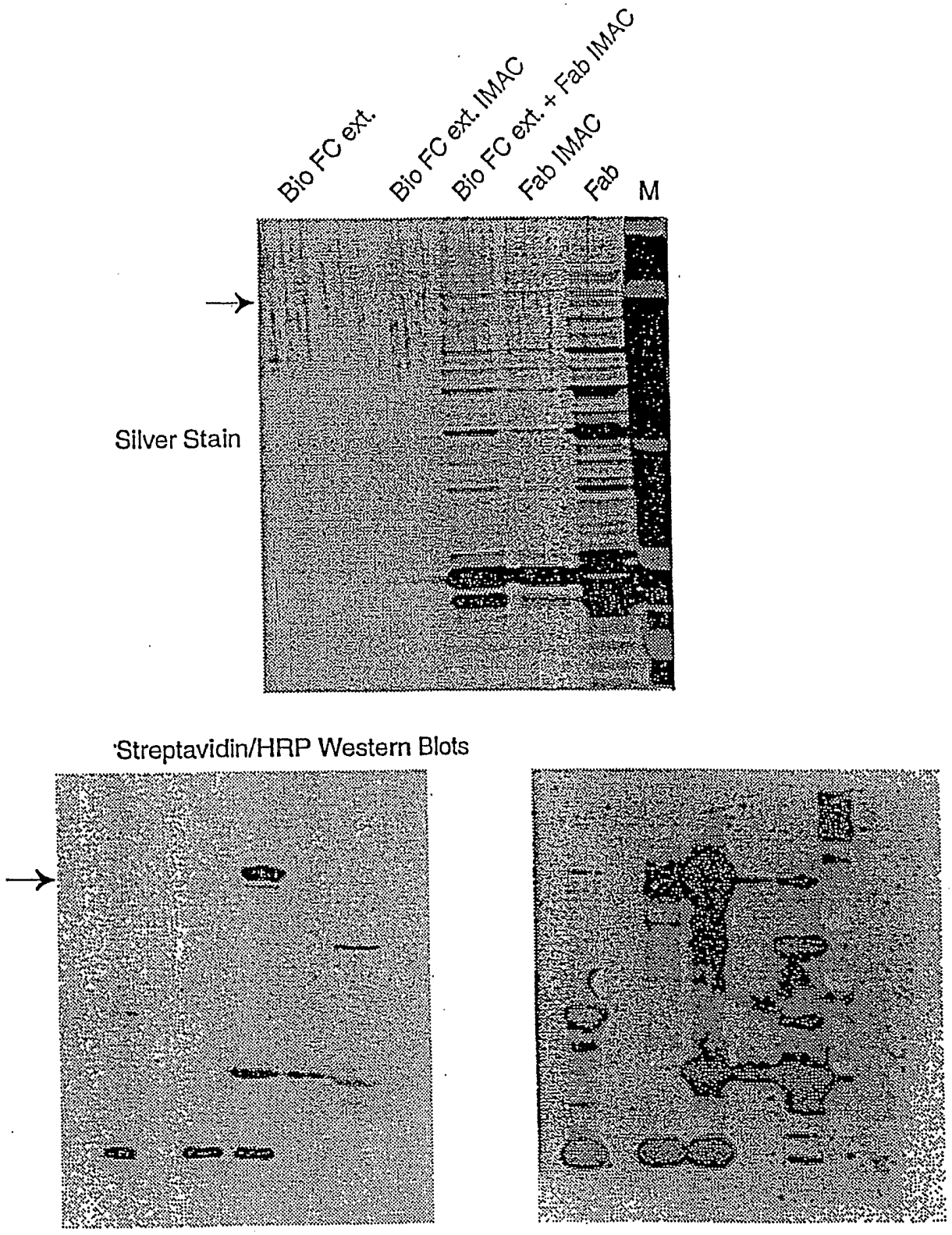
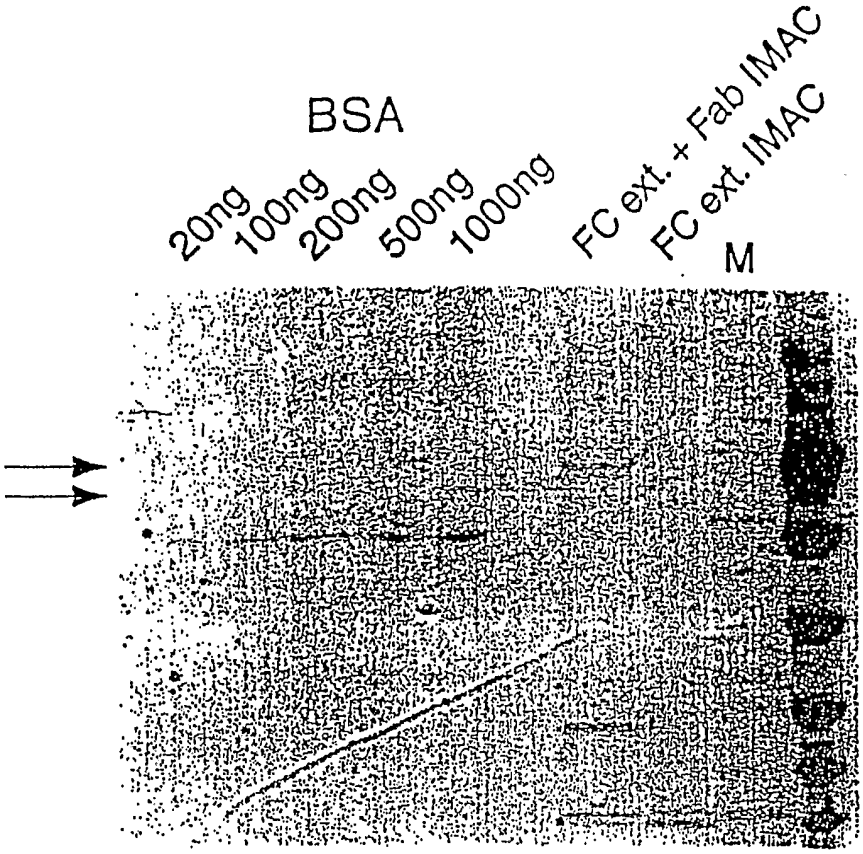

Bio FC ext.
Bio FC ext. IMAC
Bio FC ext. + Fab IMAC
Fab IMAC
Fab
M
Silver Stain
Streptavidin/HRP Western Blots

FIGURE 2


BSA
20ng
100ng
200ng
500ng
1000ng
FC ext. + Fab IMAC
FC ext. IMAC
M

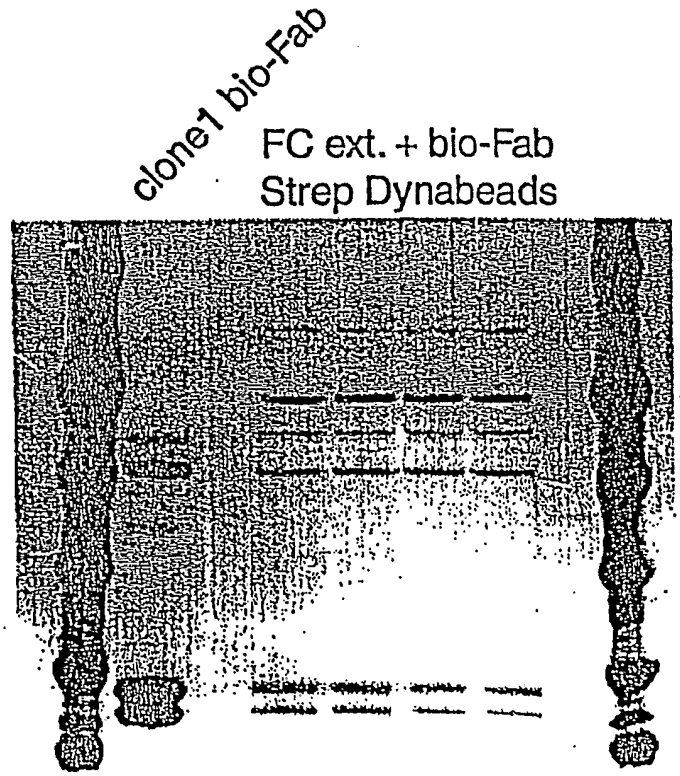
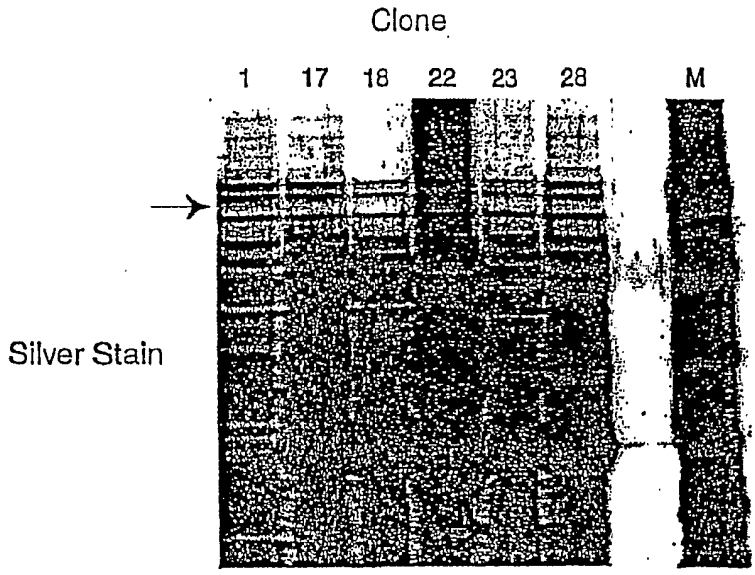
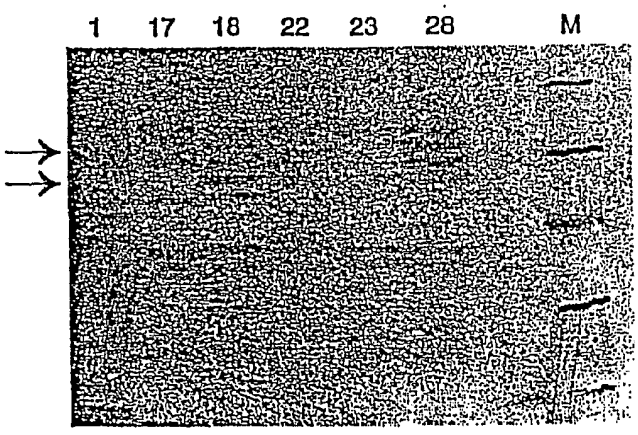
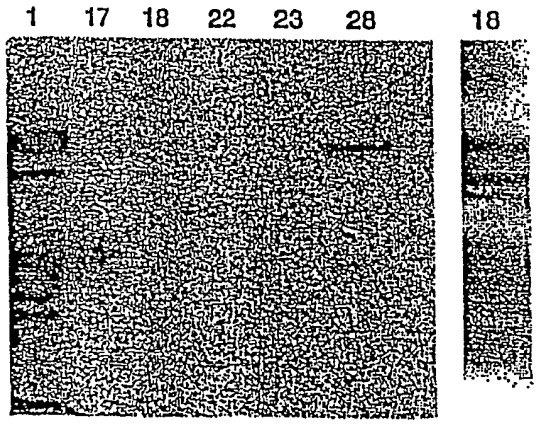

clone1 bio-Fab
FC ext. + bio-Fab
Strep Dynabeads

FIGURE 3





Streptavidin Western



Clone 1 Fab Western

FIGURE 4

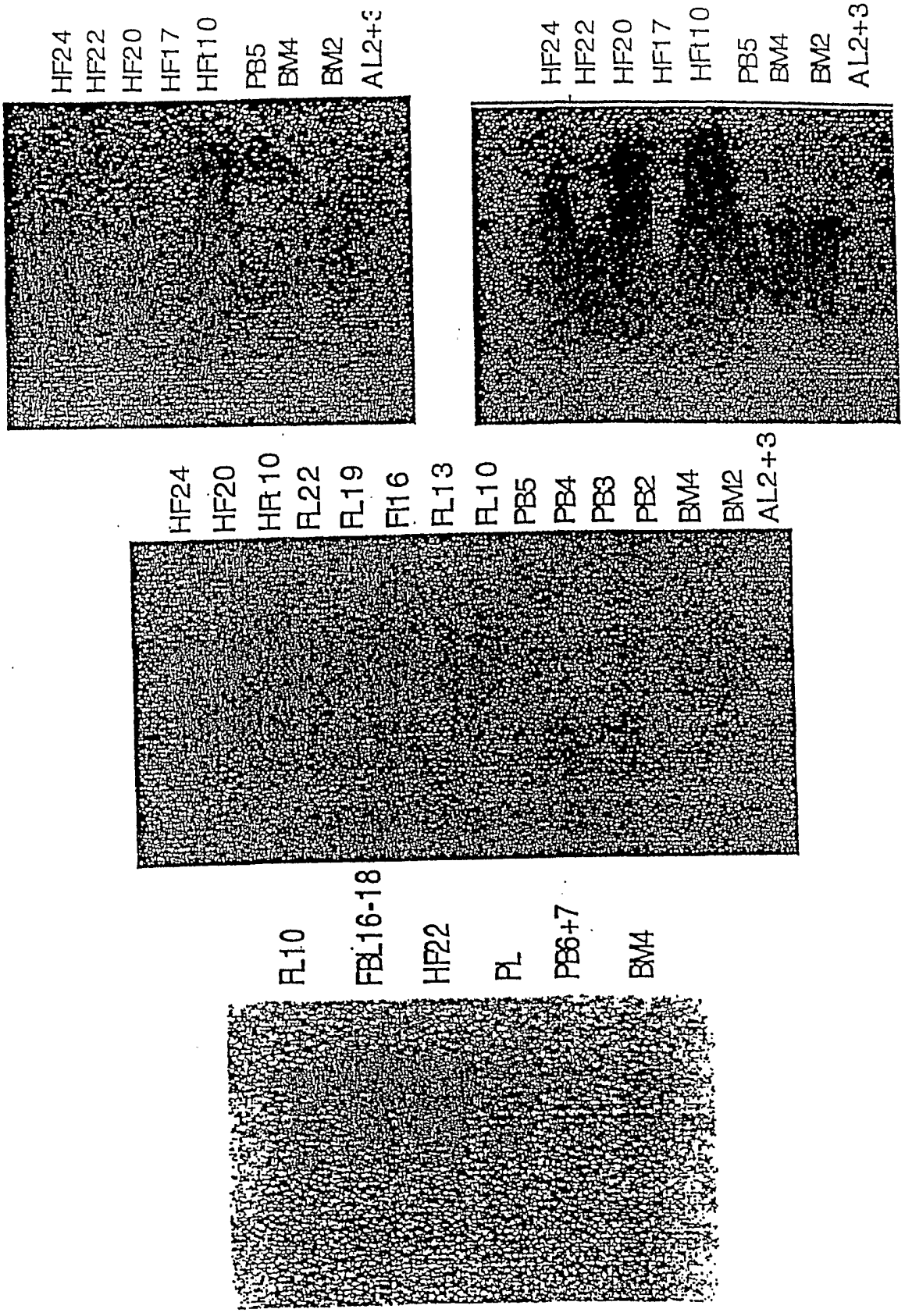
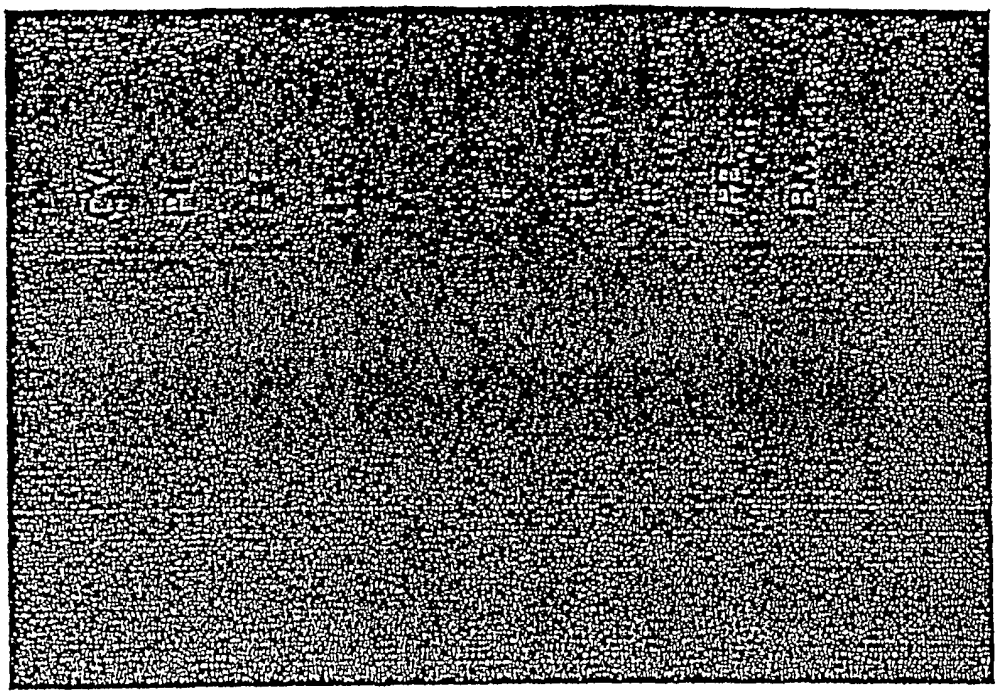

HF24
HF22
HF20
HF17
HFt10
PB5
BM4
BM2
AL2+3
HF24
HF22
HF20
HF17
HFt10
PB5
BM4
BM2
AL2+3
HF24
HF20
HR10
RL22
RL19
FI16
RL13
RL10
PB5
PB4
PB3
PB2
BM4
BM2
AL2+3
RL10
FBL16-18
HF22
PL
PB6+7
BM4

FIGURE 5



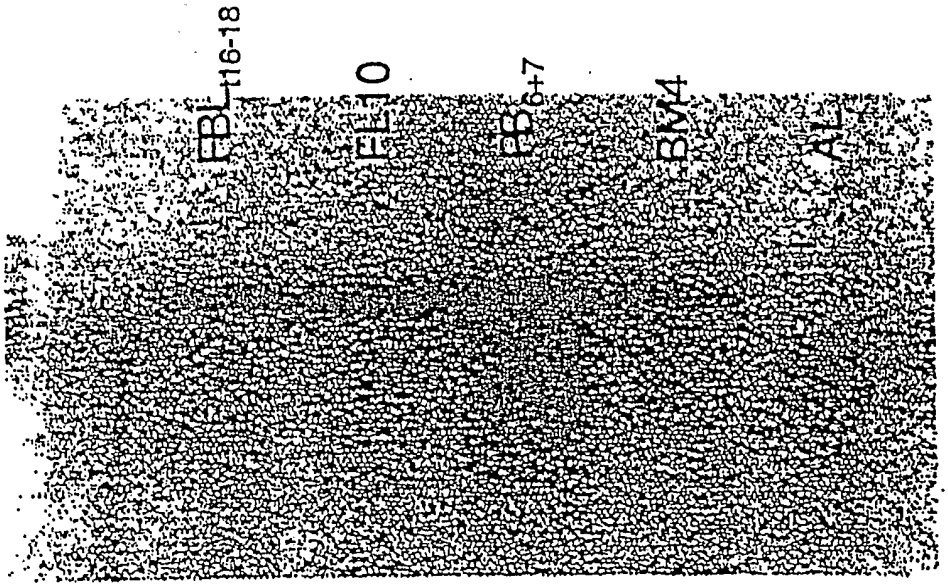


FIGURE 6

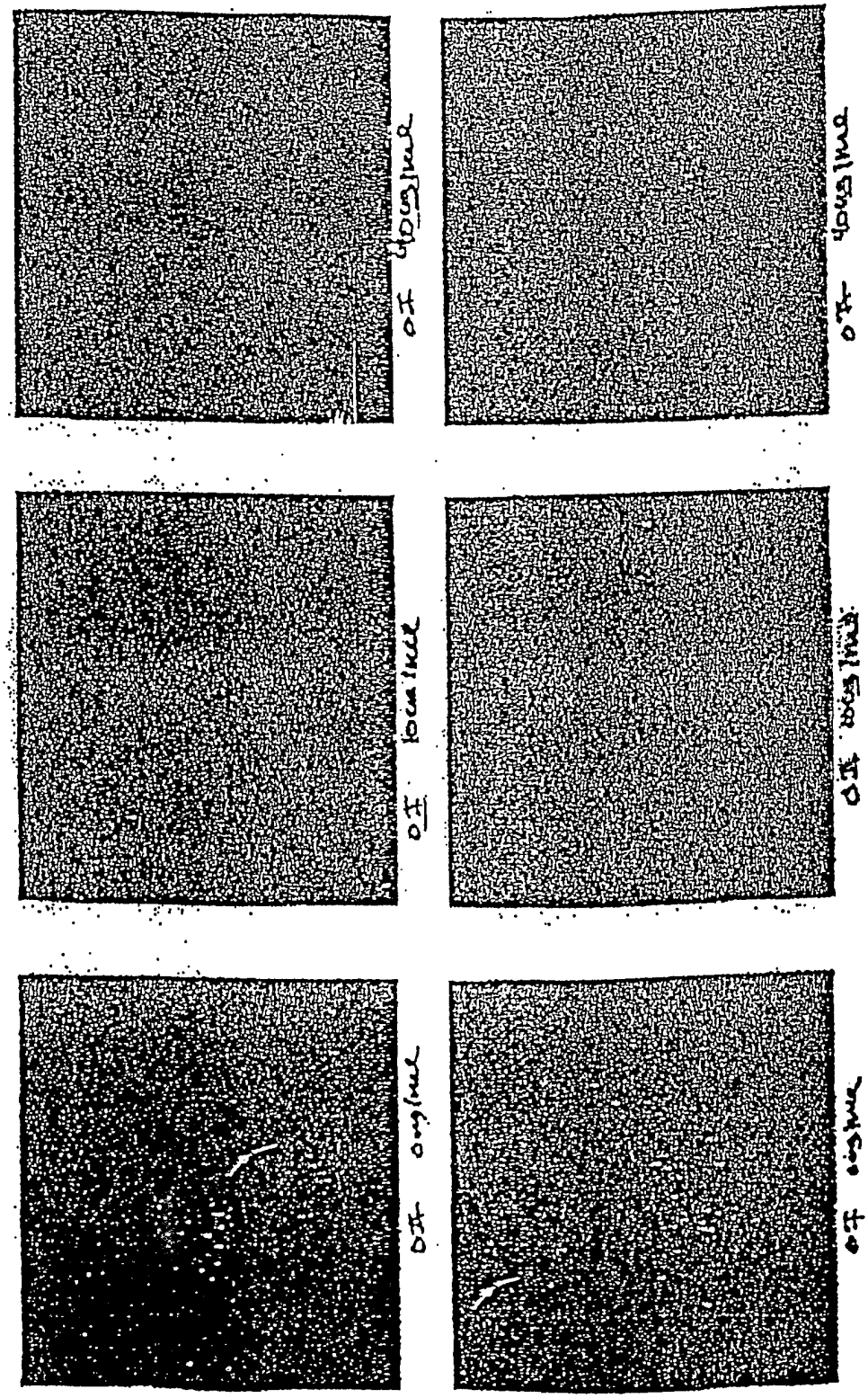


FIGURE 7

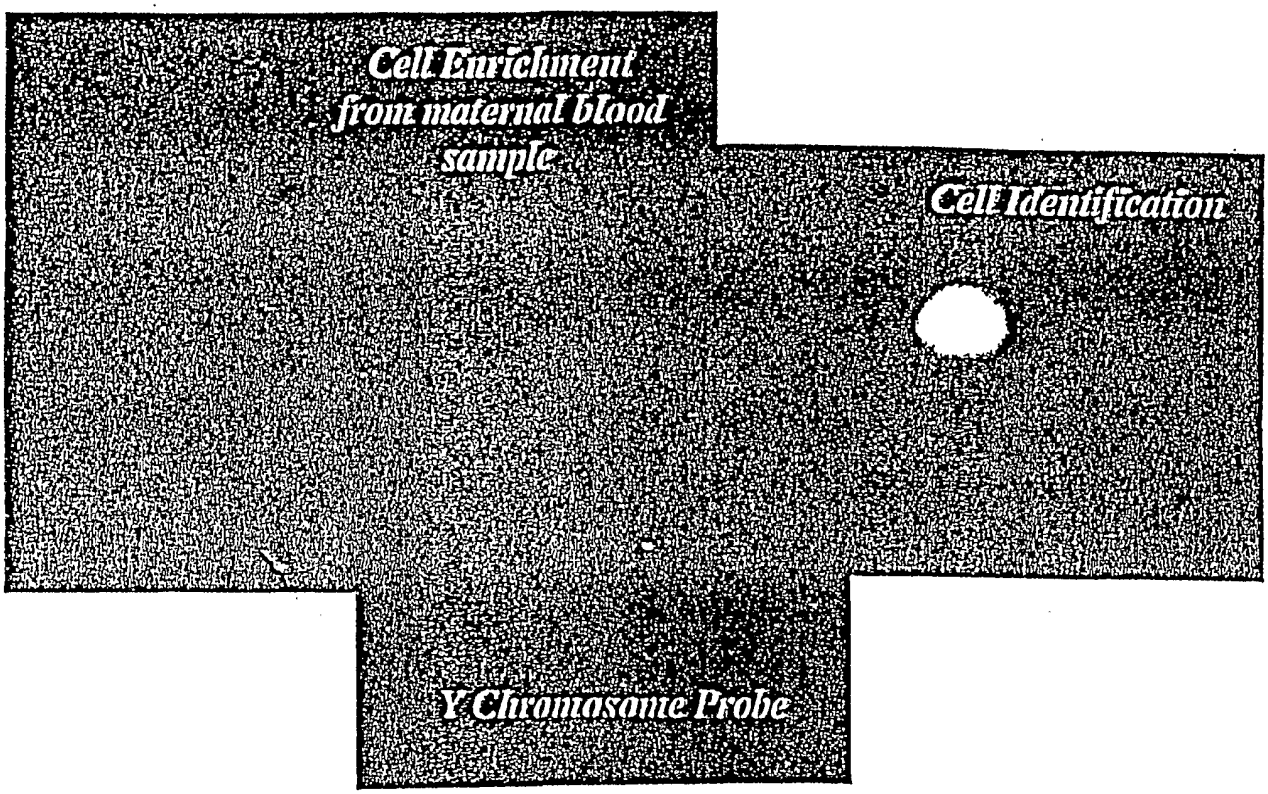

Cell Enrichment from maternal blood sample
Cell Identification
Y Chromosome Probe

FIGURE 8

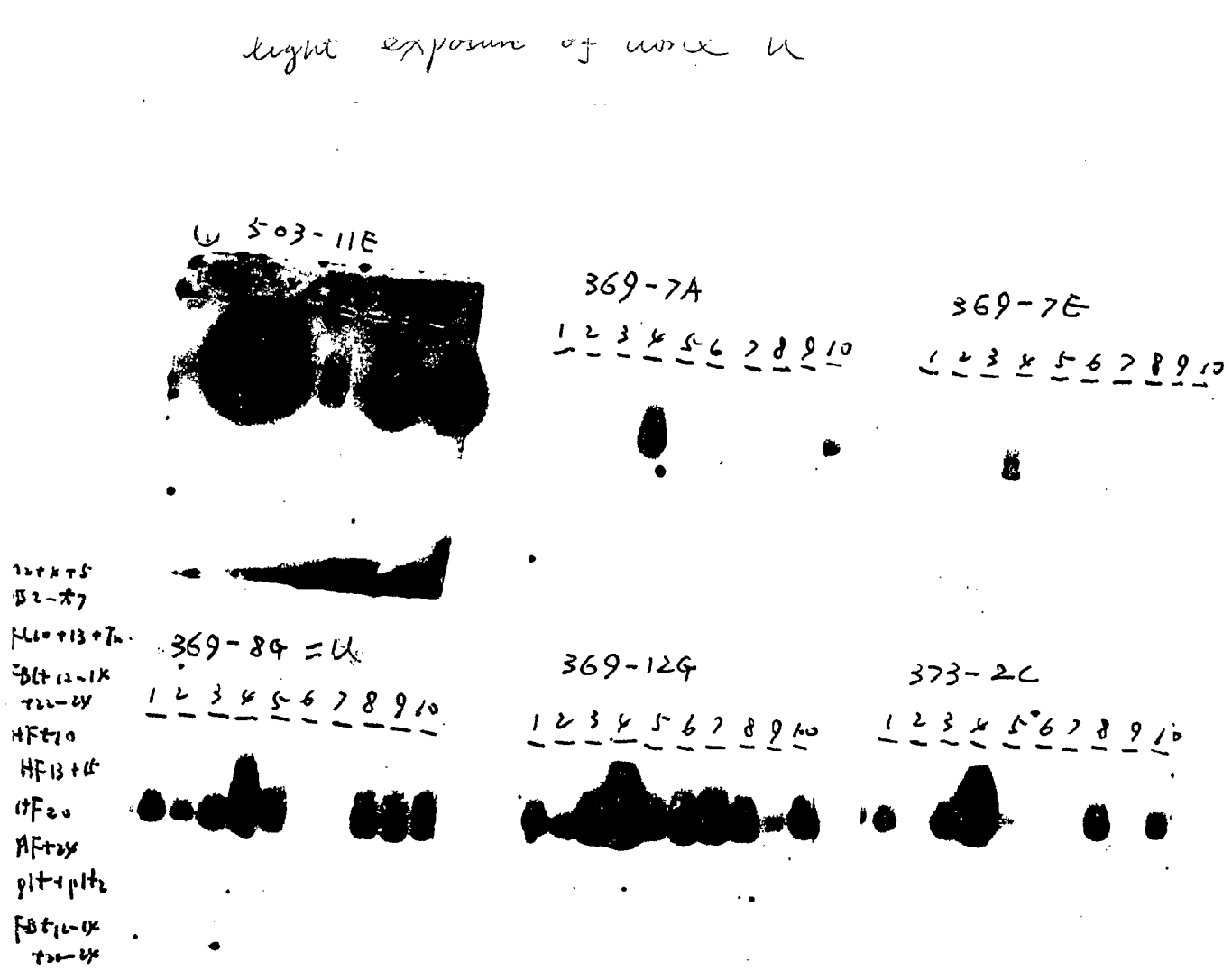
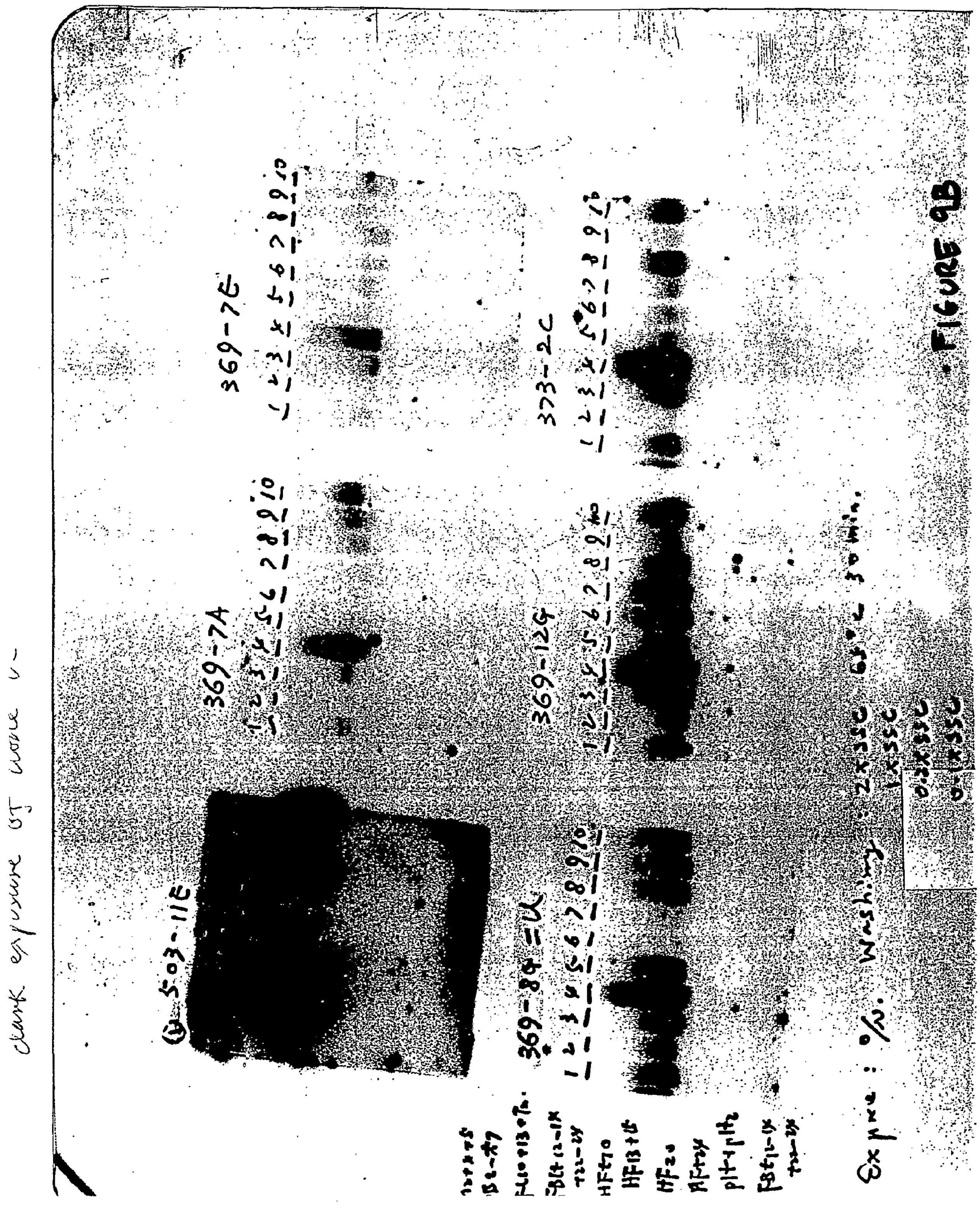

light exposure of work U
503-11E
369-7A
1 2 3 4 5 6 7 8 9 10
369-7E
1 2 3 4 5 6 7 8 9 10
369-8G = U
1 2 3 4 5 6 7 8 9 10
369-12G
1 2 3 4 5 6 7 8 9 10
373-2C
1 2 3 4 5 6 7 8 9 10
Expose : 5/20. Washing : 2×SSC 65°C 30 min.
1×SSC
0.2×SSC
0.1×SSC

FIGURE 9A



369-7A
369-7E
373-2C
369-12G
503-11E
FIGURE 9B

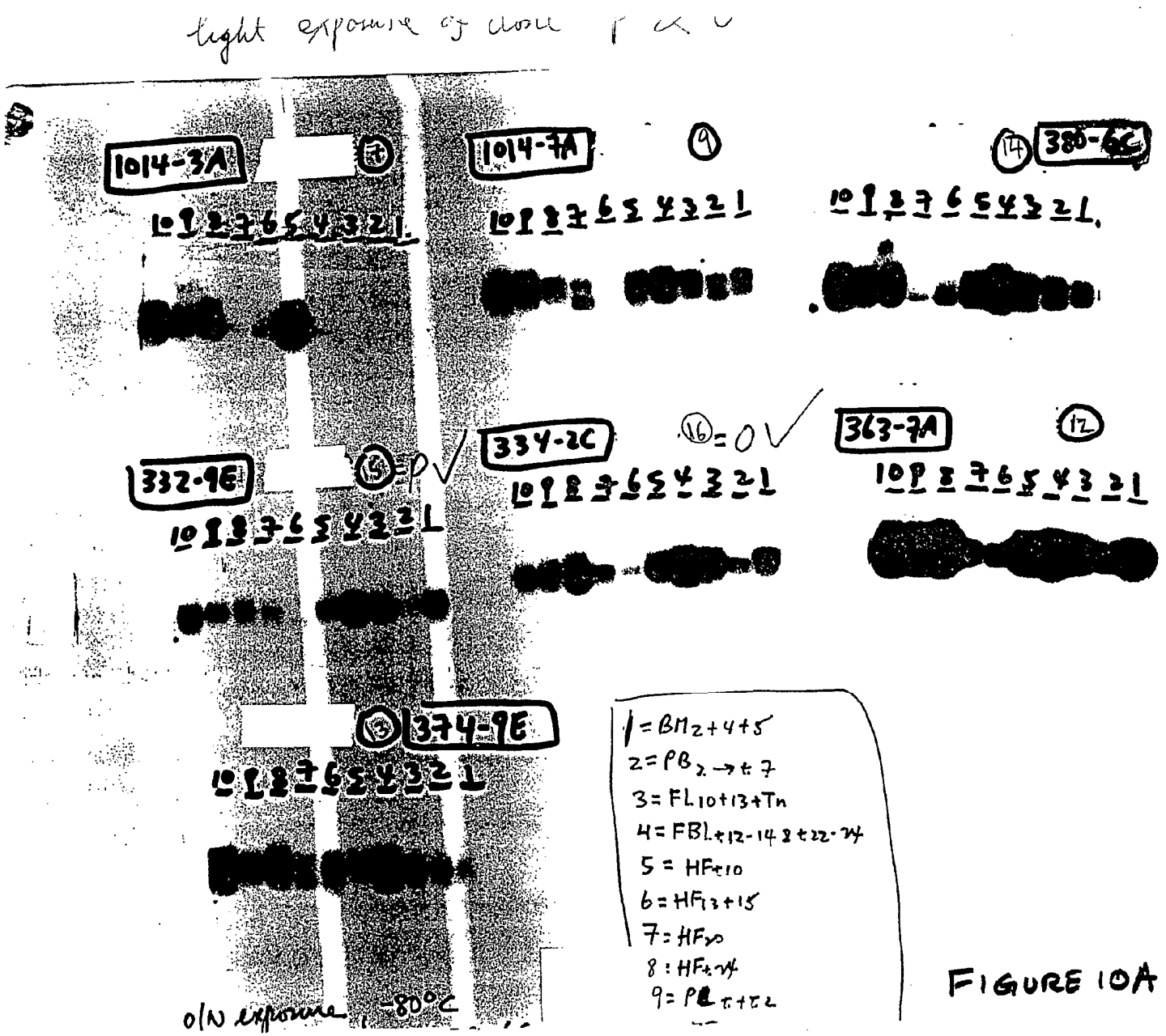

light exposure of short
1014-3A
1014-7A
380-6C
332-9E
334-2C
363-7A
374-9E
O/N exposure, -80°C
1 = BM2+4+5
2 = PB2→+7
3 = FL10+13+Tn
4 = FBL+12-14 & +22-24
5 = HF+10
6 = HF13+15
7 = HF20
8 = HF+24
9 = PBL+22

FIGURE 10A

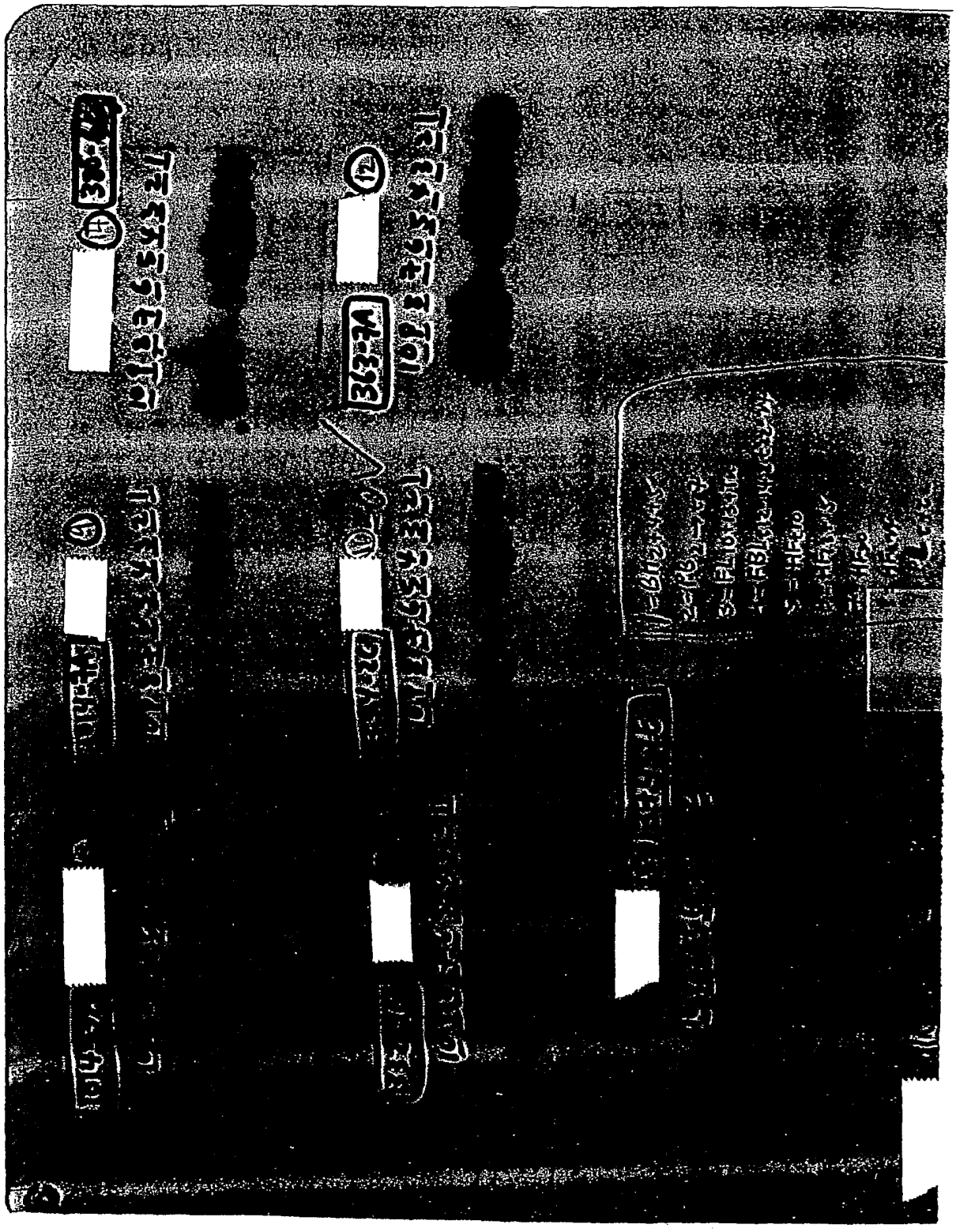


FIGURE 10B

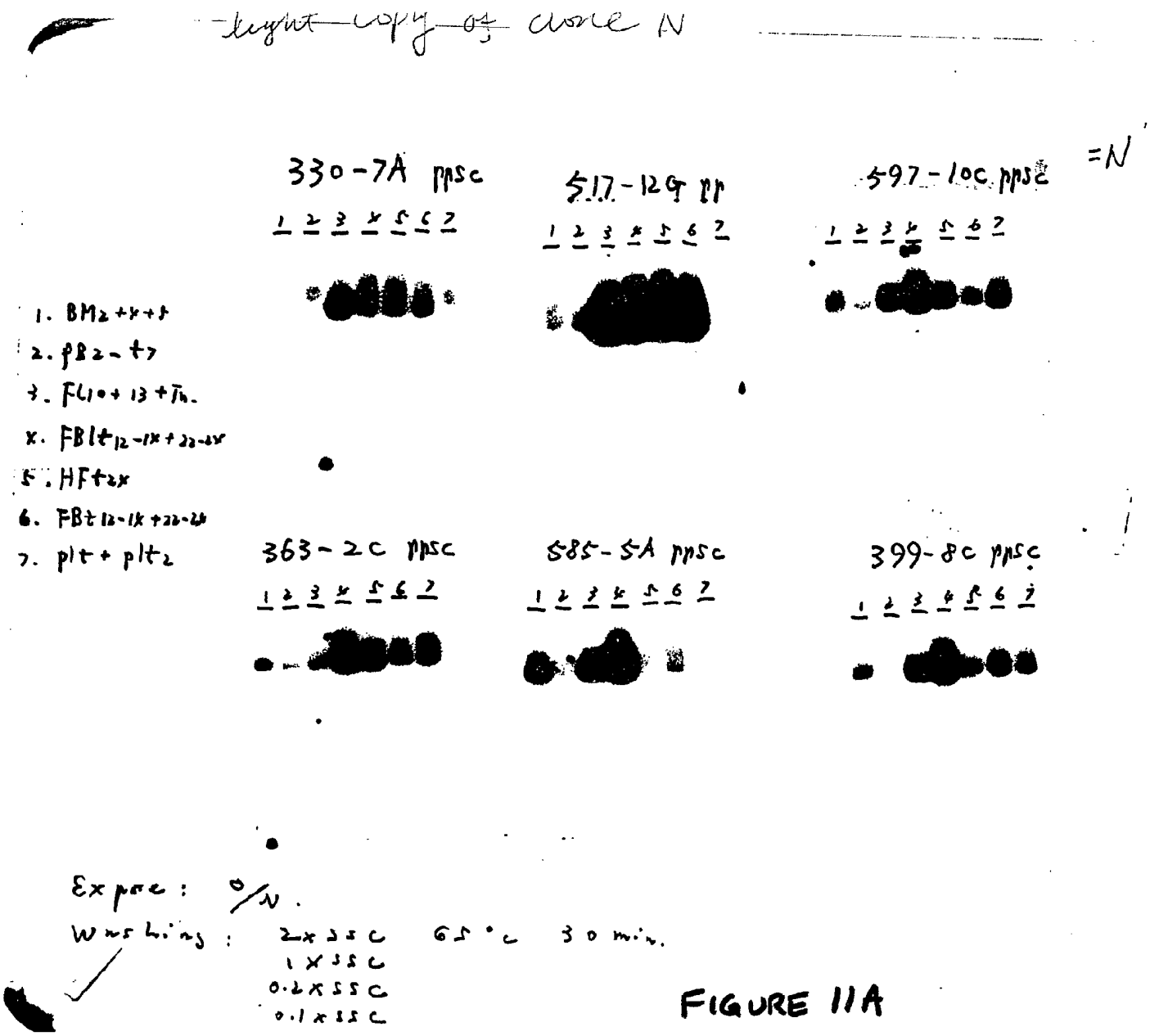

light copy of above N
330-7A pnsc
517-12G pn
597-10C pnsc
=N
1 2 3 4 5 6 7
1. BM2 + 4 + 3
2. pB2 - t7
3. FL10 + 13 + Th.
4. FBlt12-11 + 22-24
5. HF + 2x
6. FBt 12-11 + 22-24
7. plt + plt2
363-2C pnsc
585-5A pnsc
399-8C pnsc
Expose: O/N.
Washing: 2x SSC 65°C 30 min.
1x SSC
0.2x SSC
0.1x SSC

FIGURE 11A

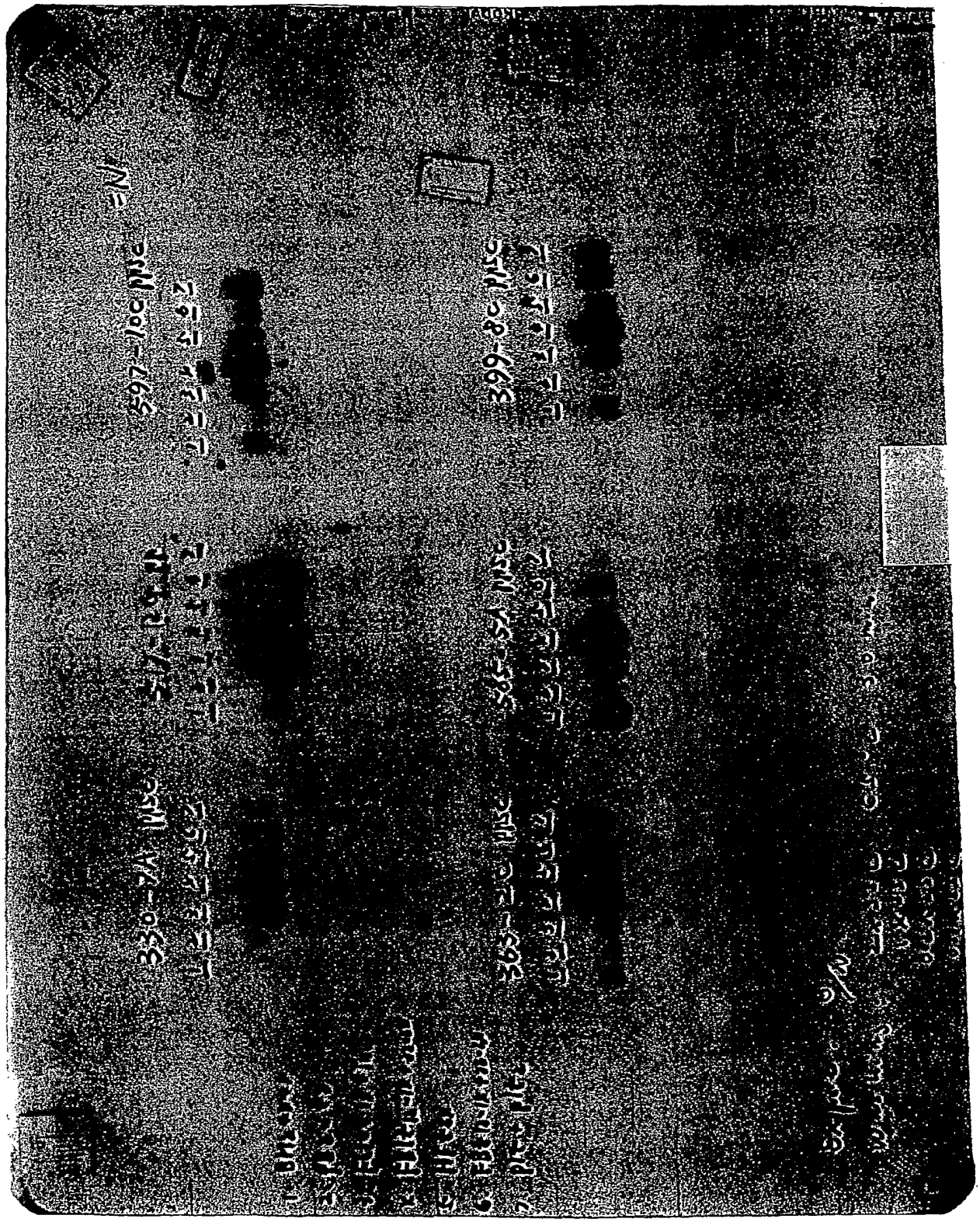


FIGURE 11B

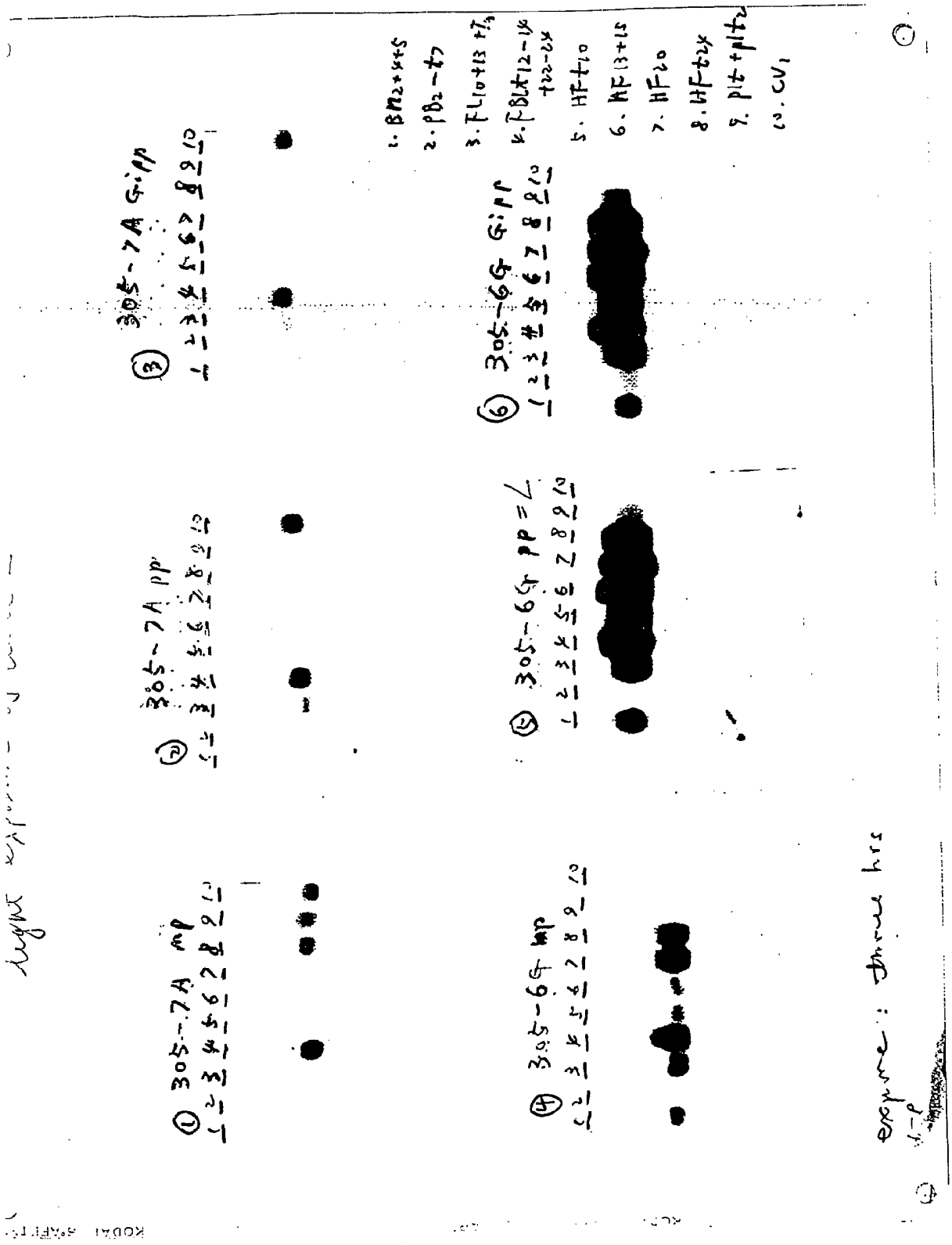


FIGURE 12A

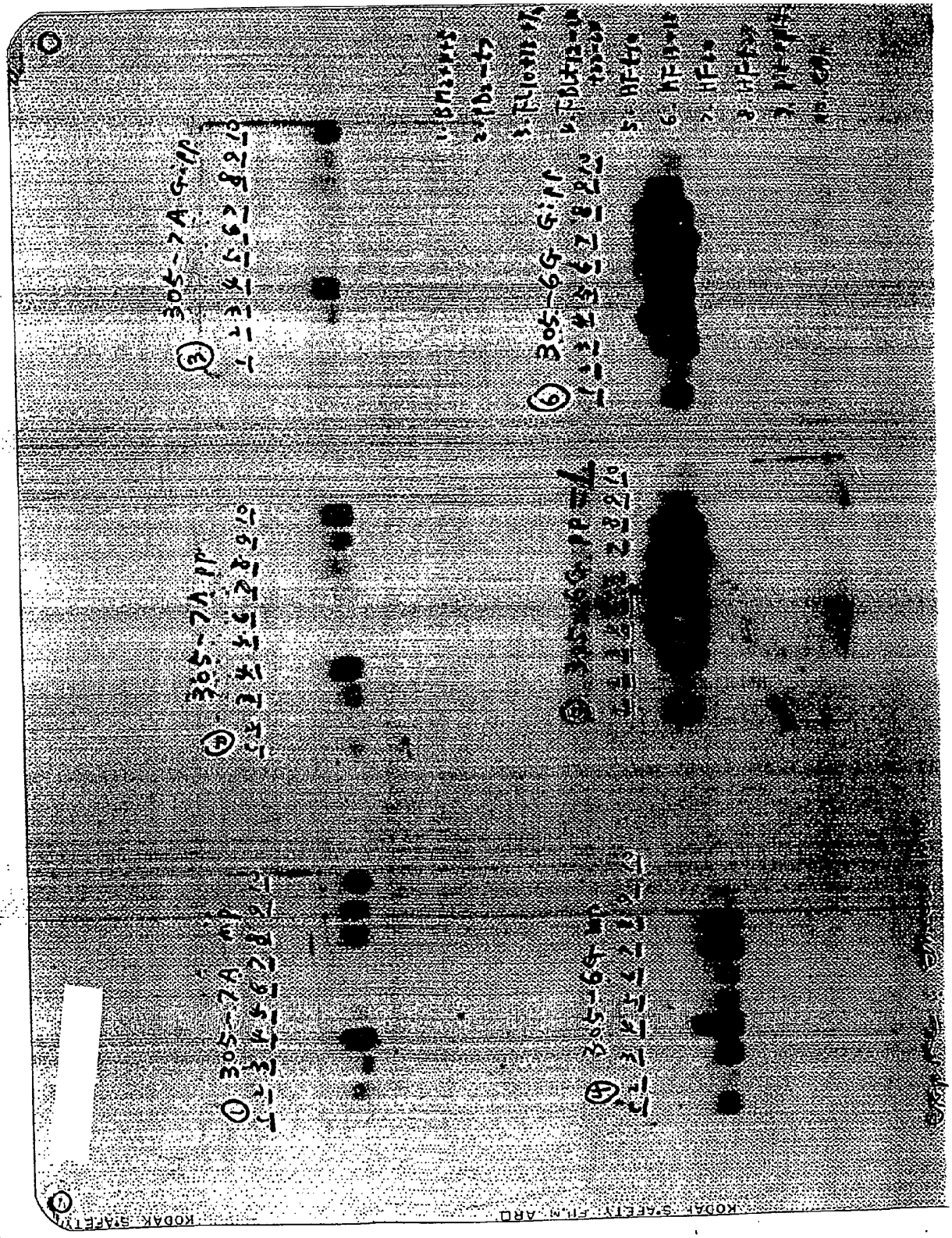


FIGURE 12B

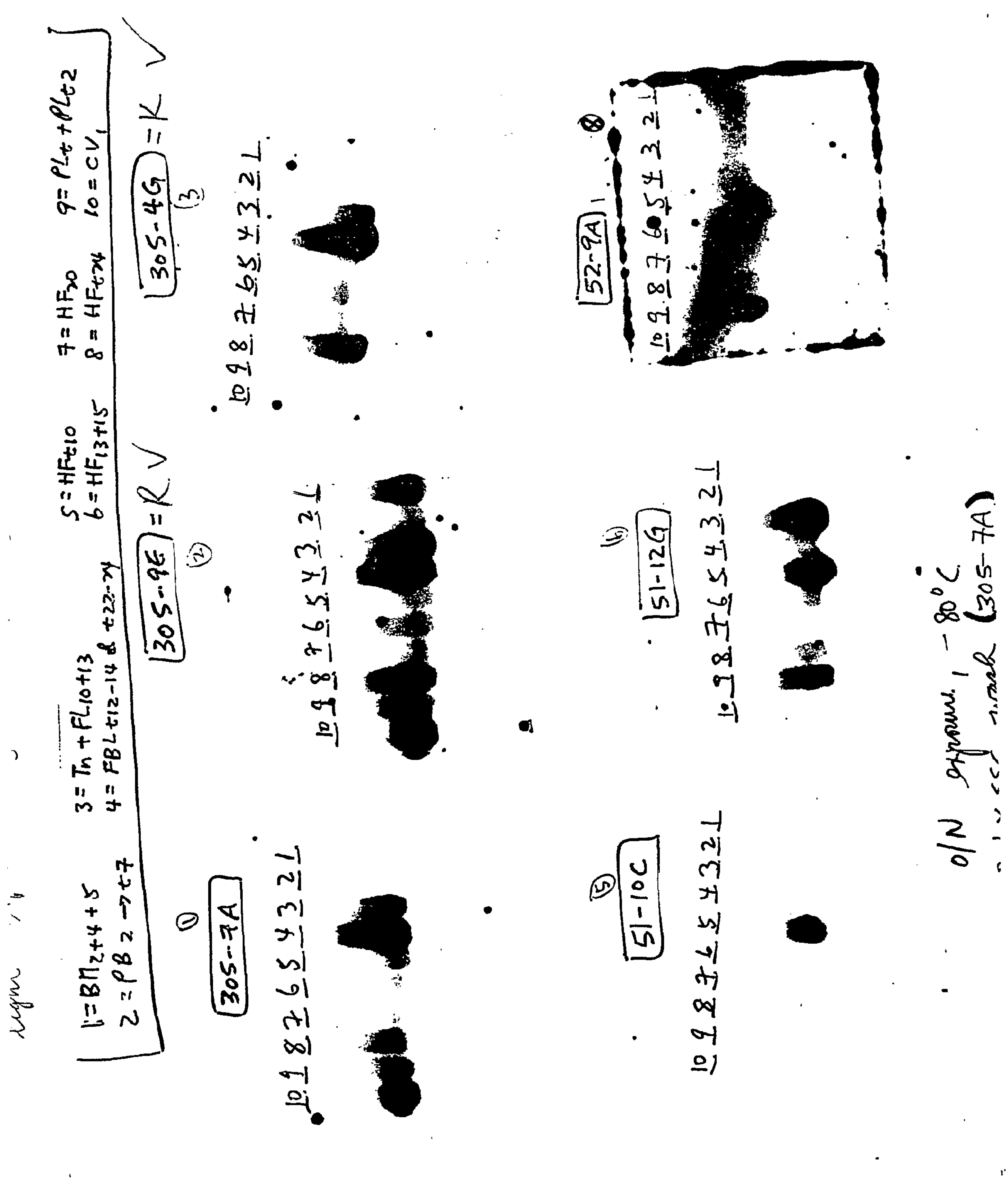


FIGURE 13A

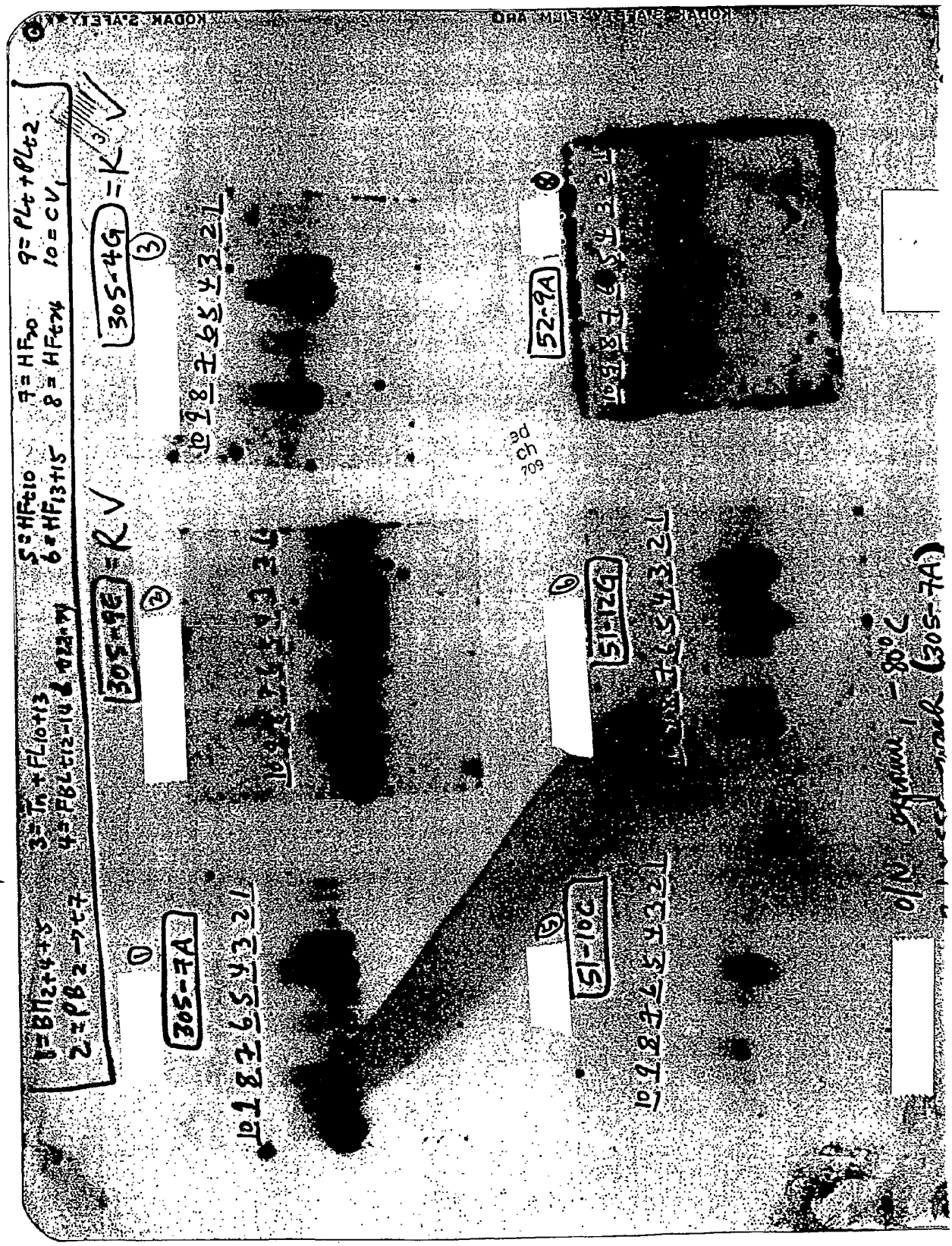

dark exposure of clone R & K
1=BM2+4+5 3=Tn+FL10+13 5=HF10 7=HF20 9=PL4+PL12
2=PB2 4=PB12+14 6=HF13+15 8=HF24 10=CV1
305-4G
305-4E
305-7A
52-9A
51-10G
51-12G
305-4G (305-7A)

FIGURE 13B

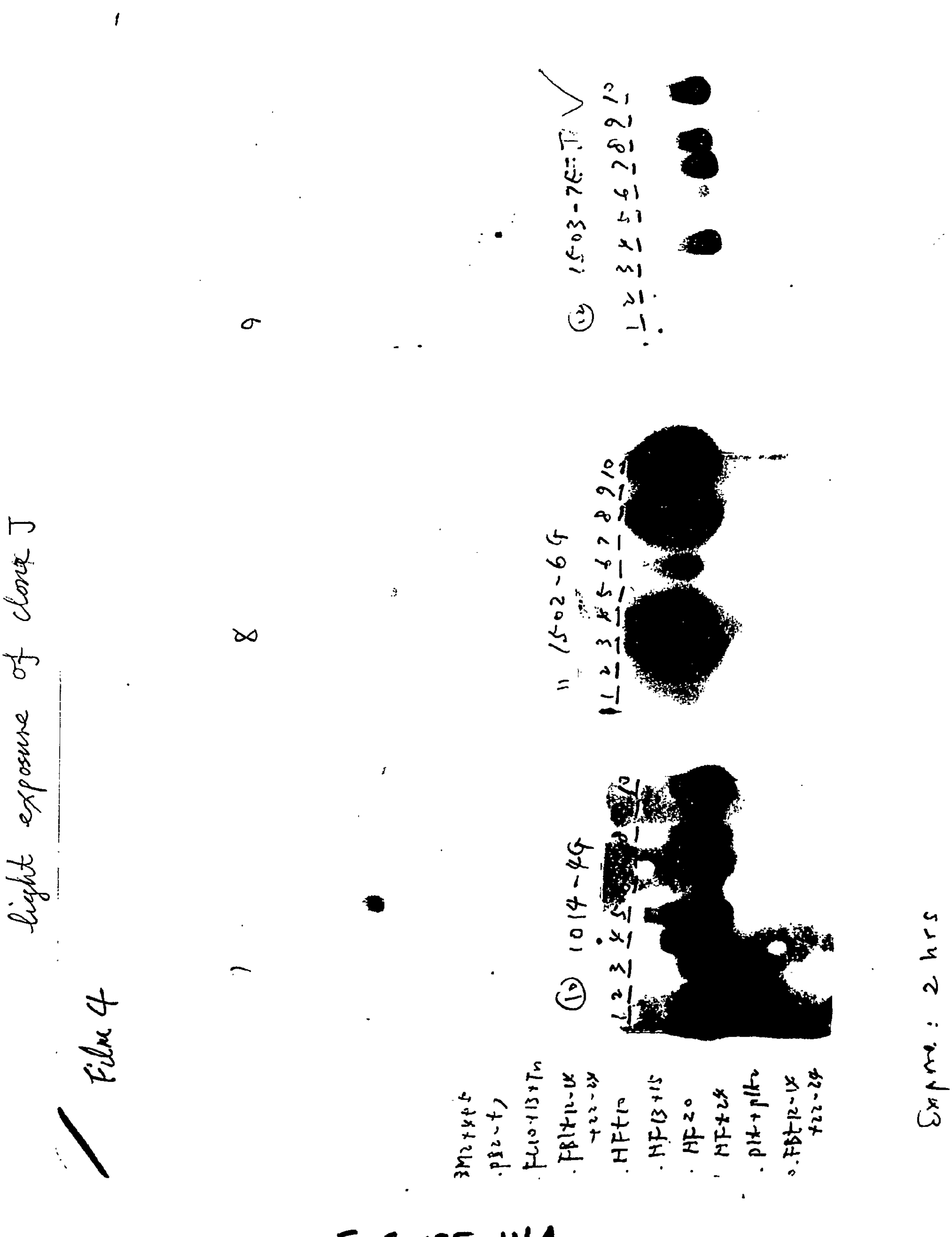


FIGURE 14A

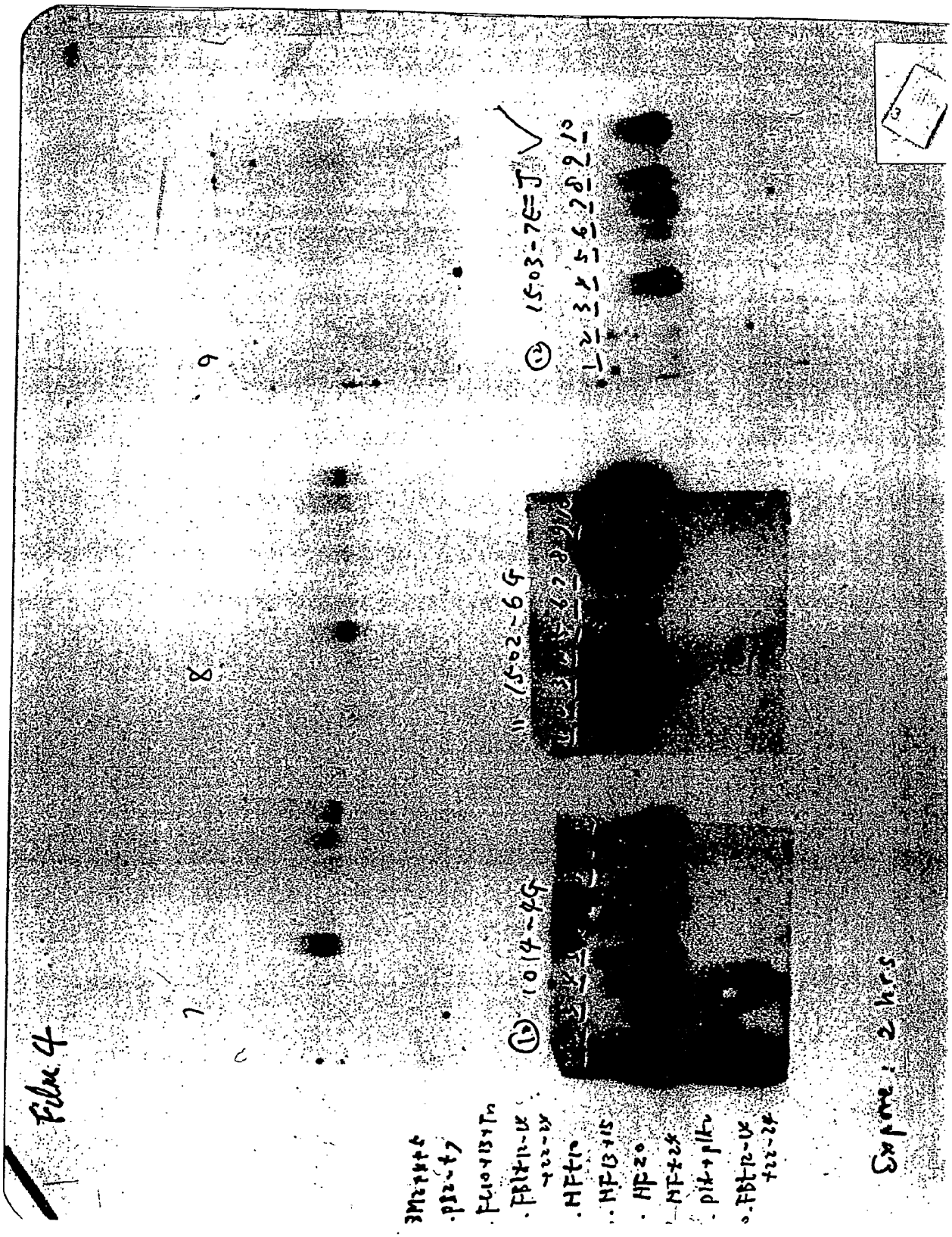


FIGURE 14B

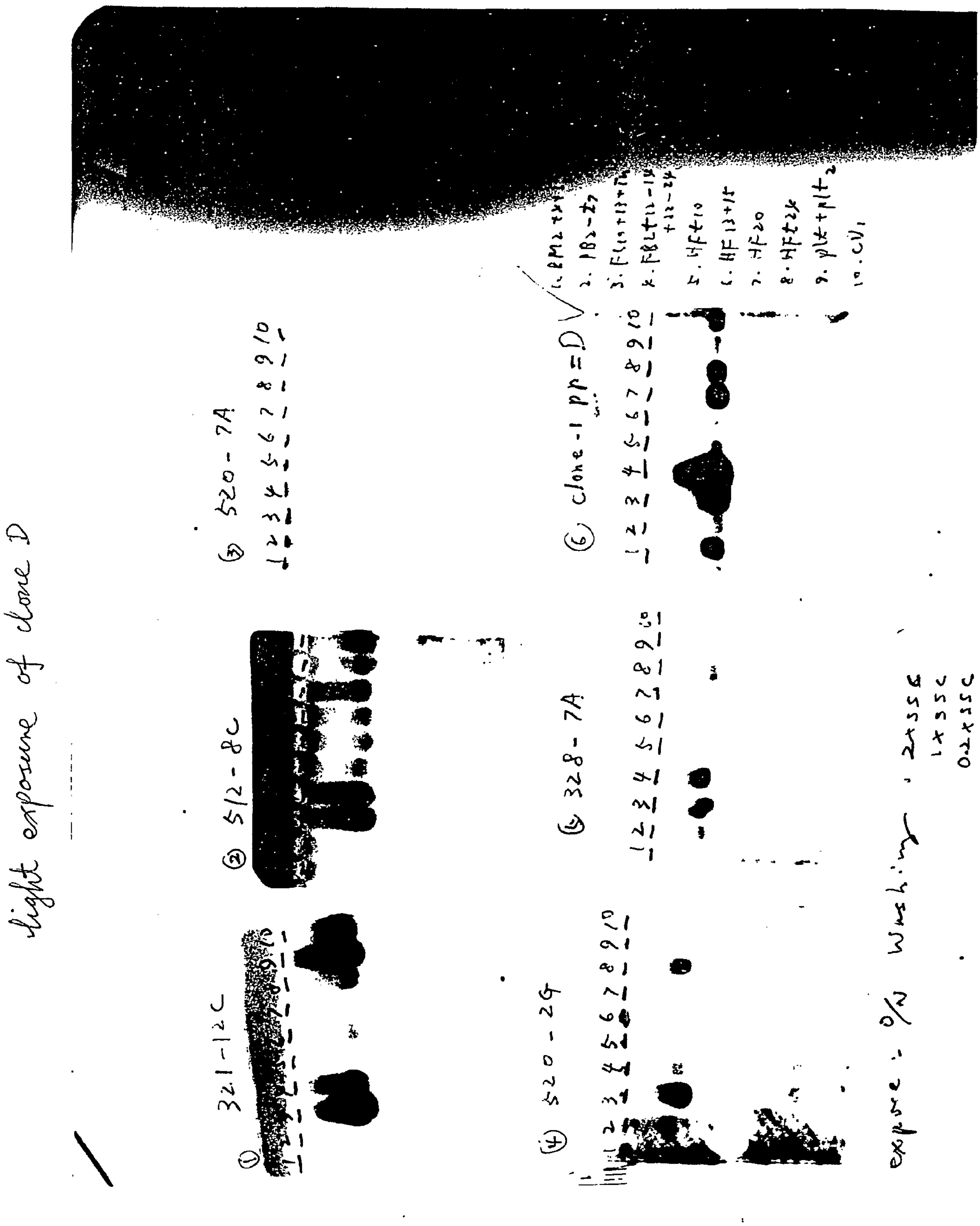

light exposure of clone D
① 321-12C
② 512-8C
③ 520-7A
④ 520-2G
⑤ 328-7A
⑥ clone-1 pp = D
expose: O/N Washing: 2xSSC
1xSSC
0.2xSSC

FIGURE 15A

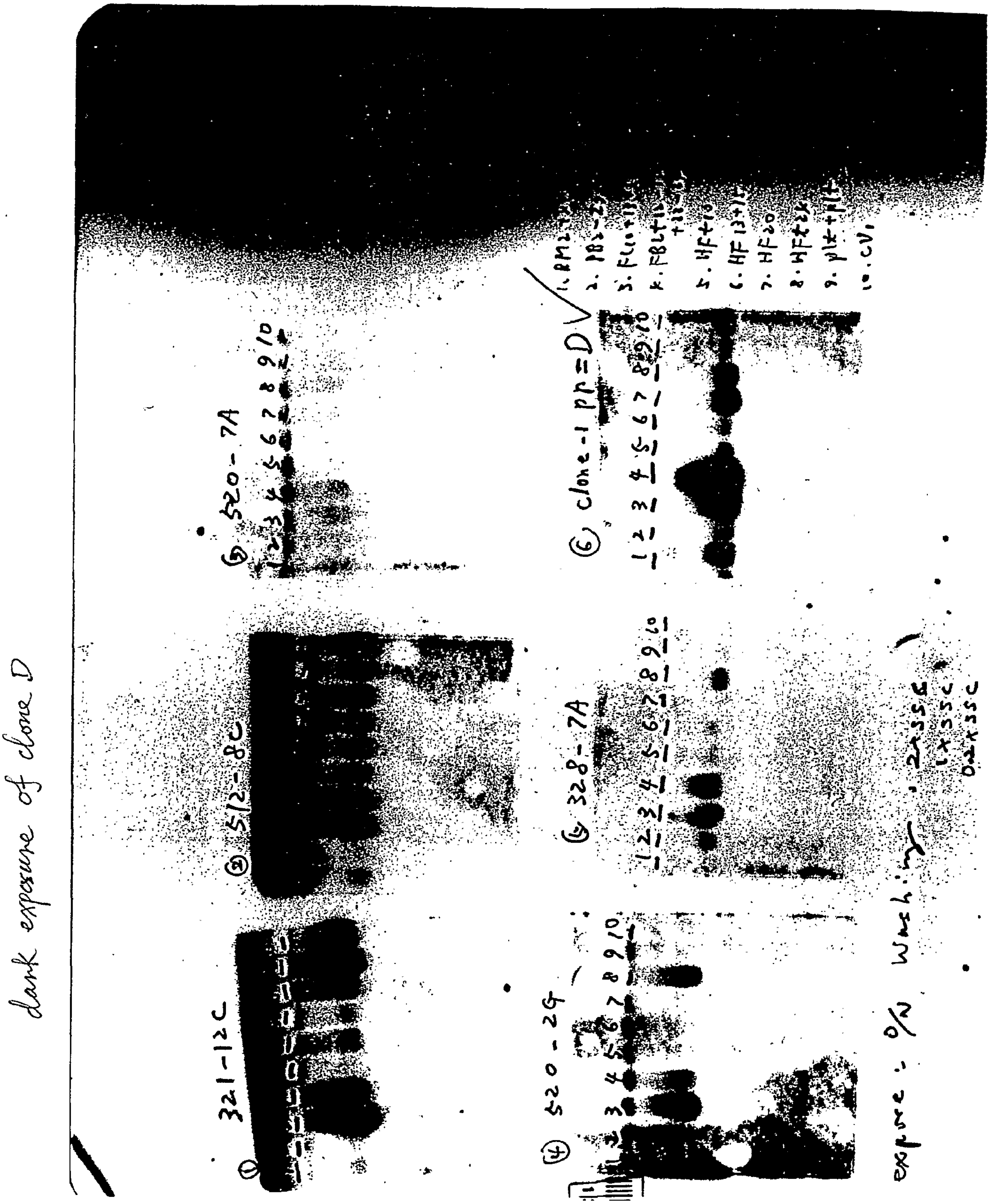
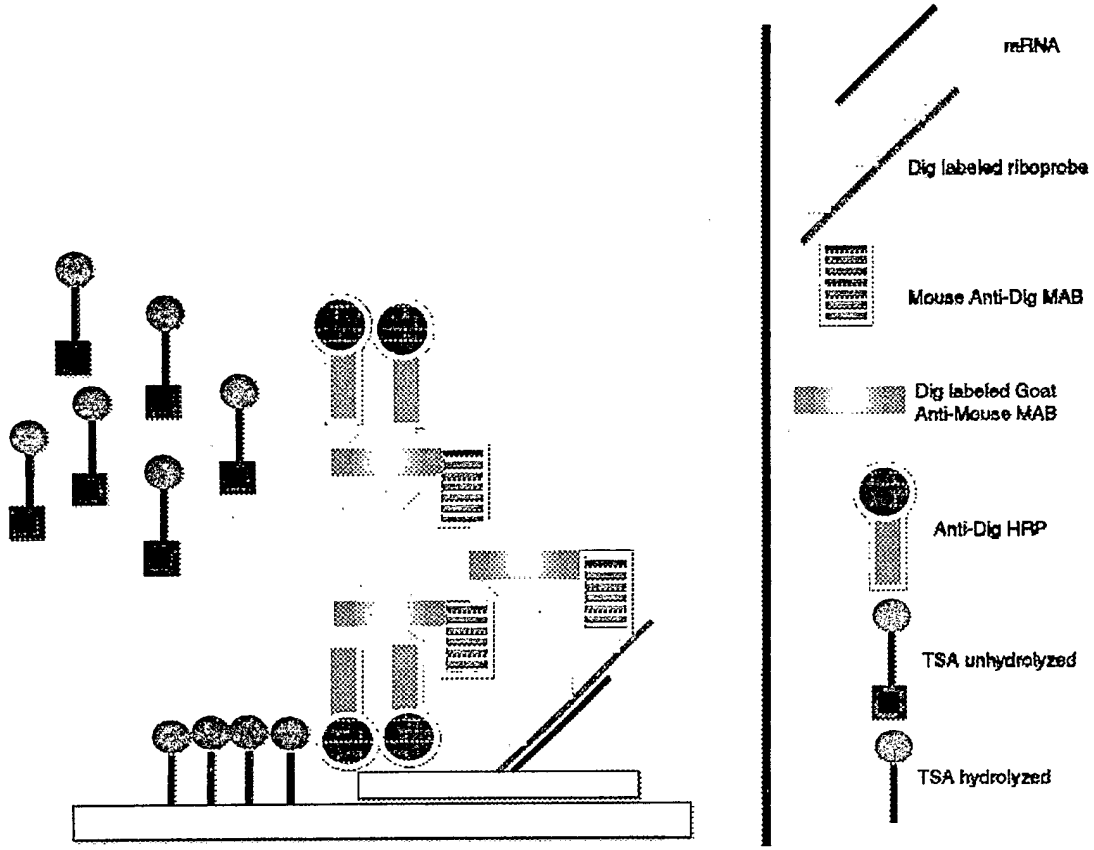

dark exposure of clone D
321-12C
512-8C
520-7A
520-2G
328-7A
clone-1 pp=D
1 2 3 4 5 6 7 8 9 10
expose: O/N Washing

FIGURE 15B

Fetal Cell Analysis

In Situ Hybridization Protocol



Evaluation of fetal cell erythroblast marker candidates

FIGURE 16A

WO 02/055985 PCT/US01/45340

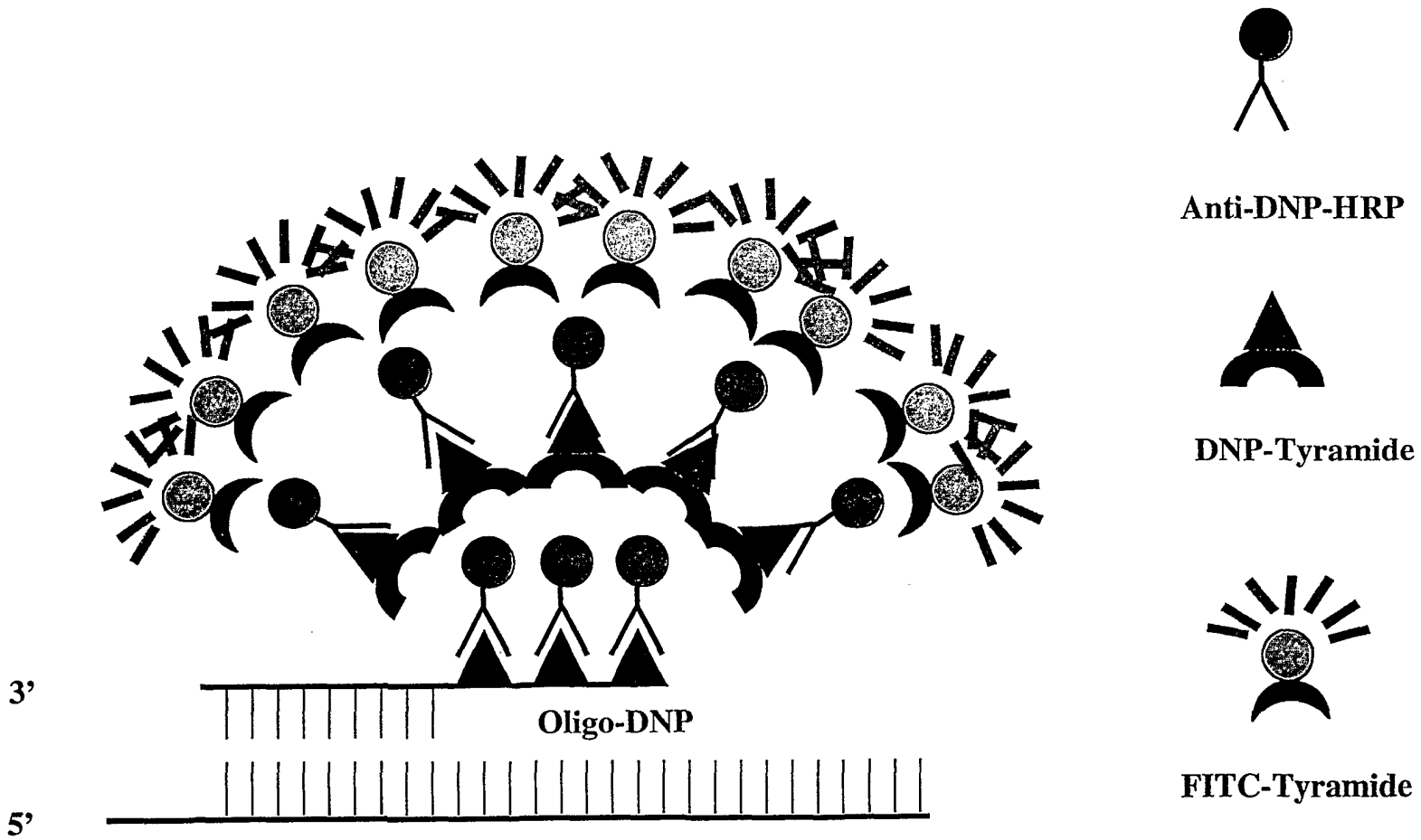


FIGURE 16B

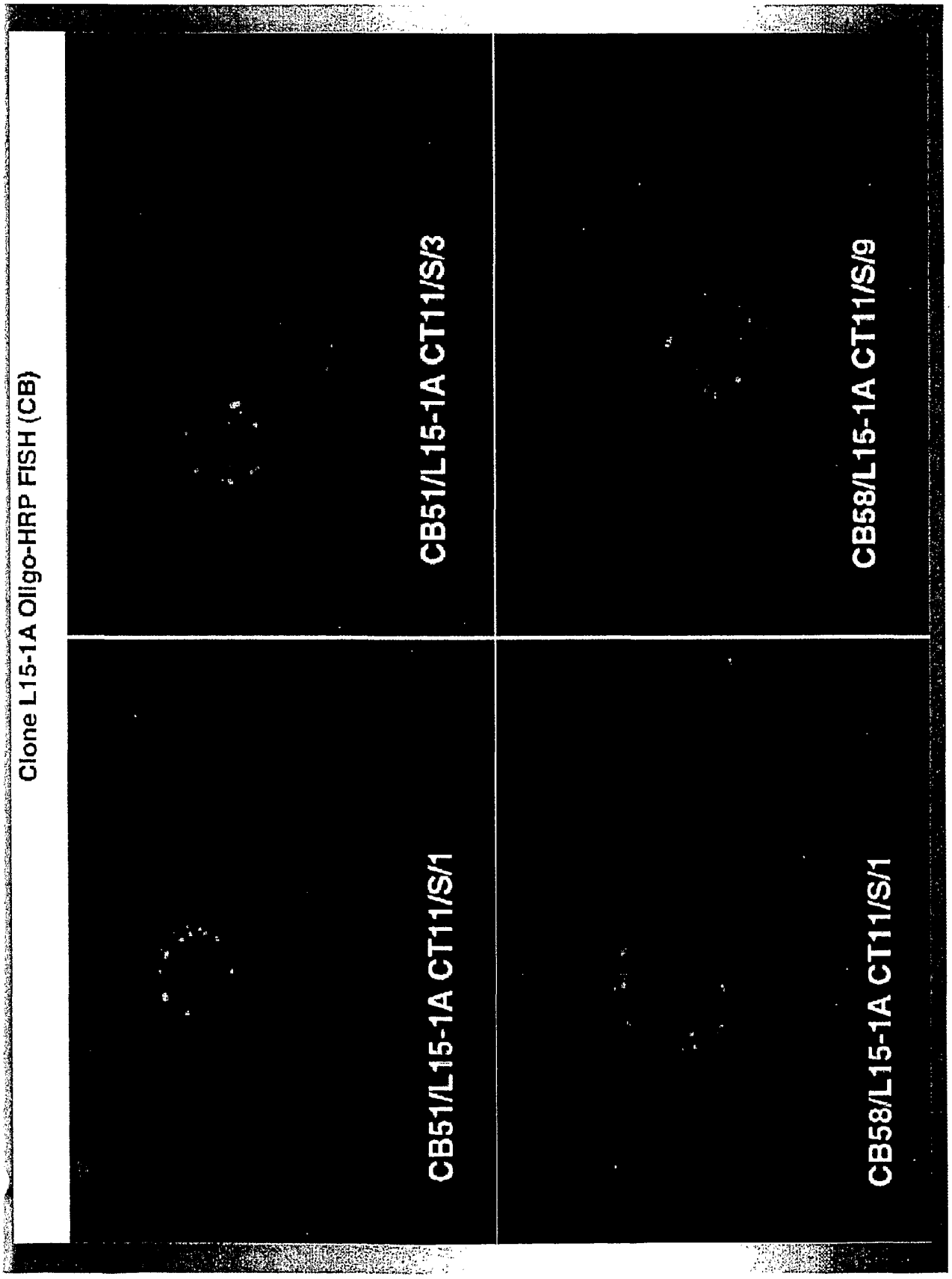

Clone L15-1A Oligo-HRP FISH (CB)
CB51/L15-1A CT11/S/1
CB51/L15-1A CT11/S/3
CB58/L15-1A CT11/S/1
CB58/L15-1A CT11/S/9

FIGURE 17

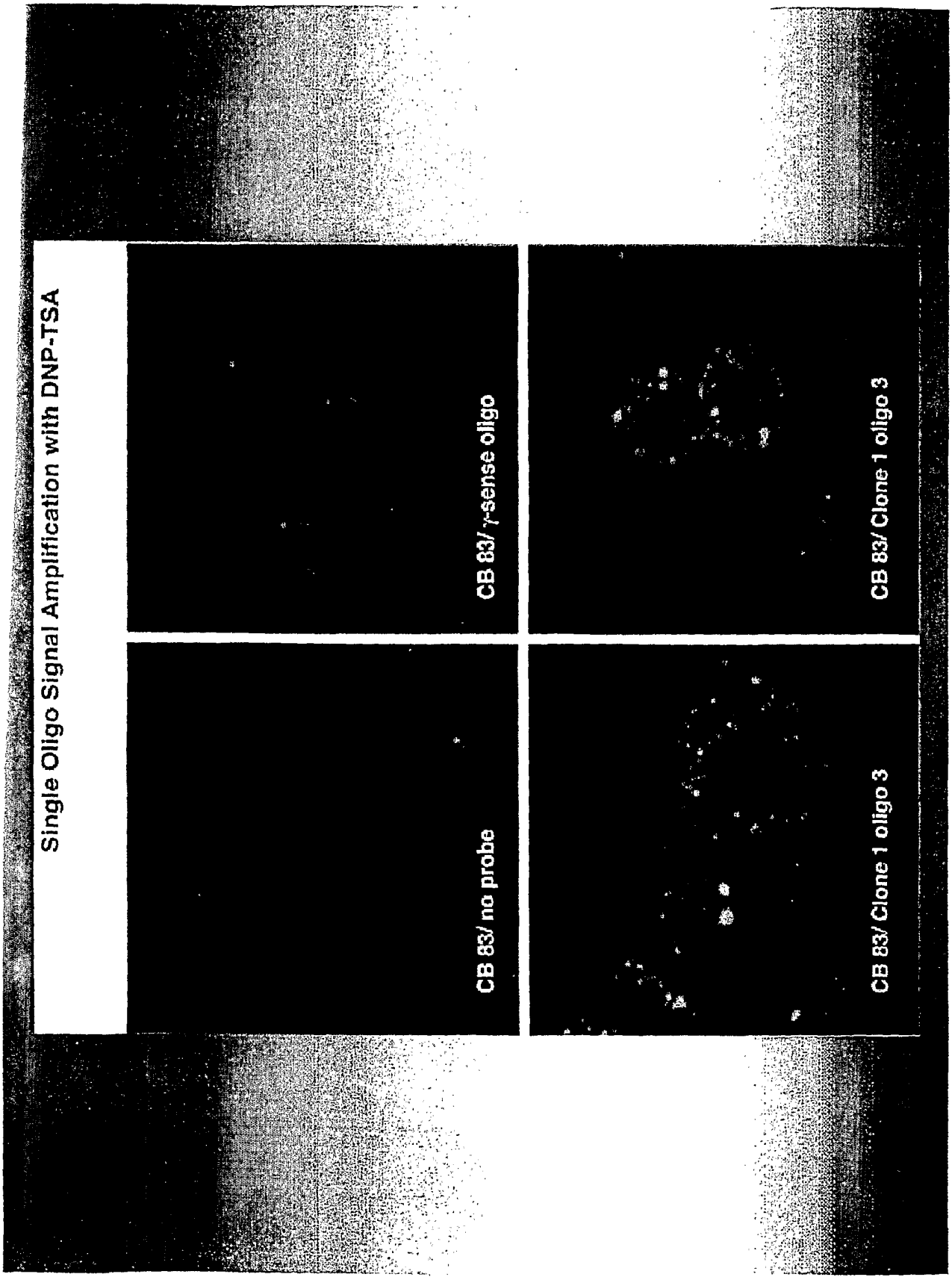

Single Oligo Signal Amplification with DNP-TSA
CB 83/ no probe
CB 83/ γ-sense oligo
CB 83/ Clone 1 oligo 3
CB 83/ Clone 1 oligo 3

FIGURE 18

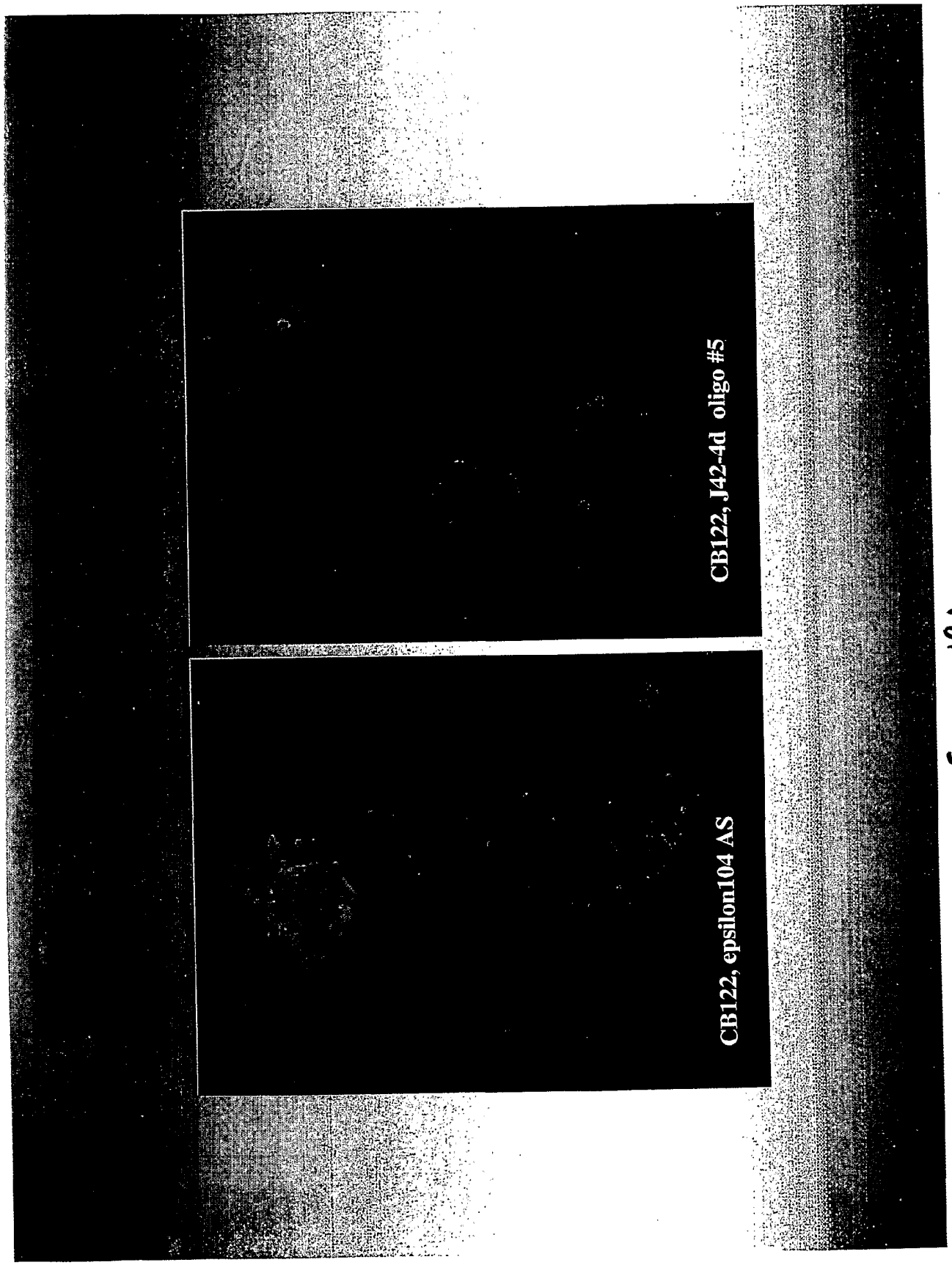


FIGURE 19A

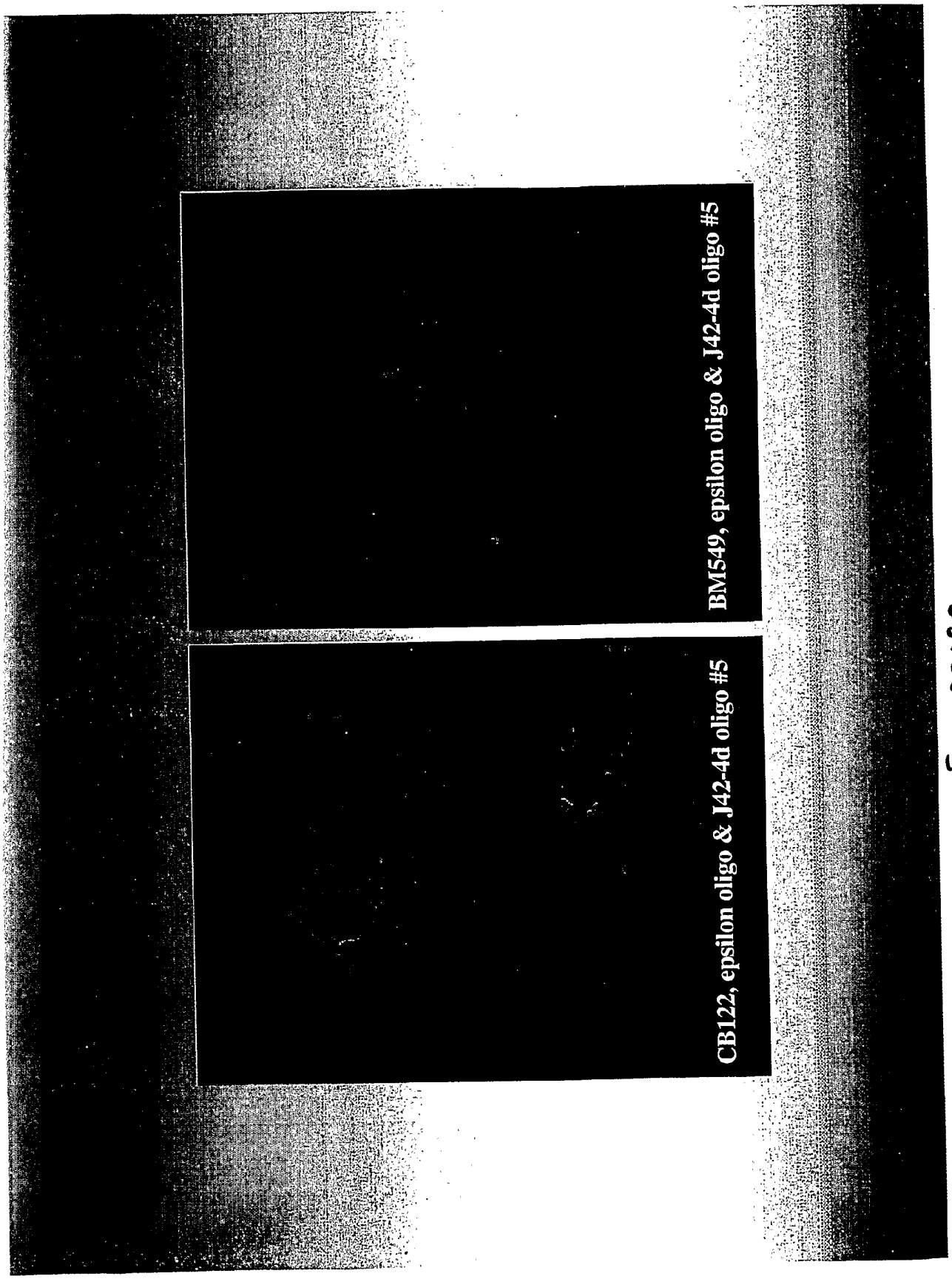

CB122, epsilon oligo & J42-4d oligo #5
BM549, epsilon oligo & J42-4d oligo #5

FIGURE 19B

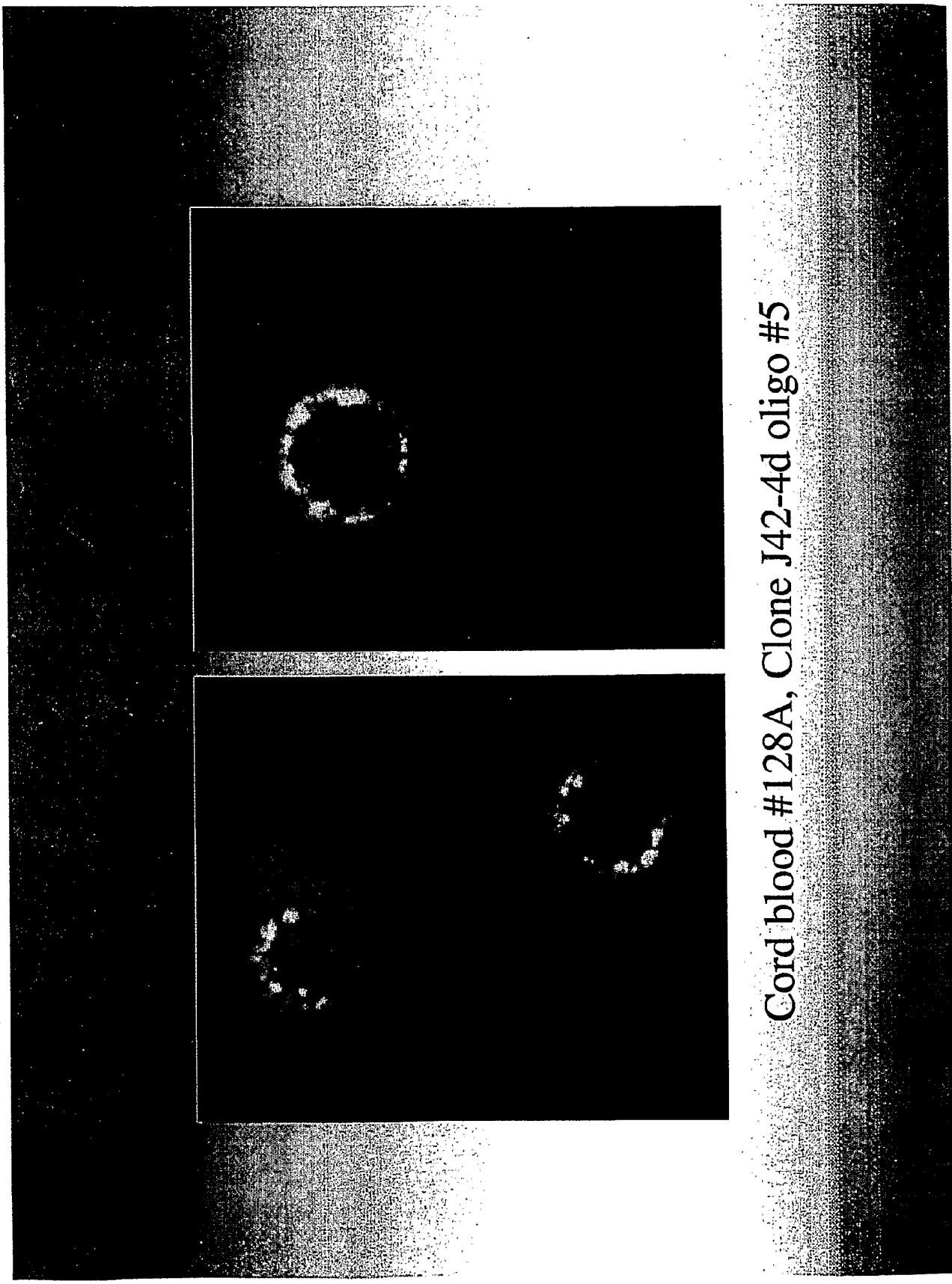
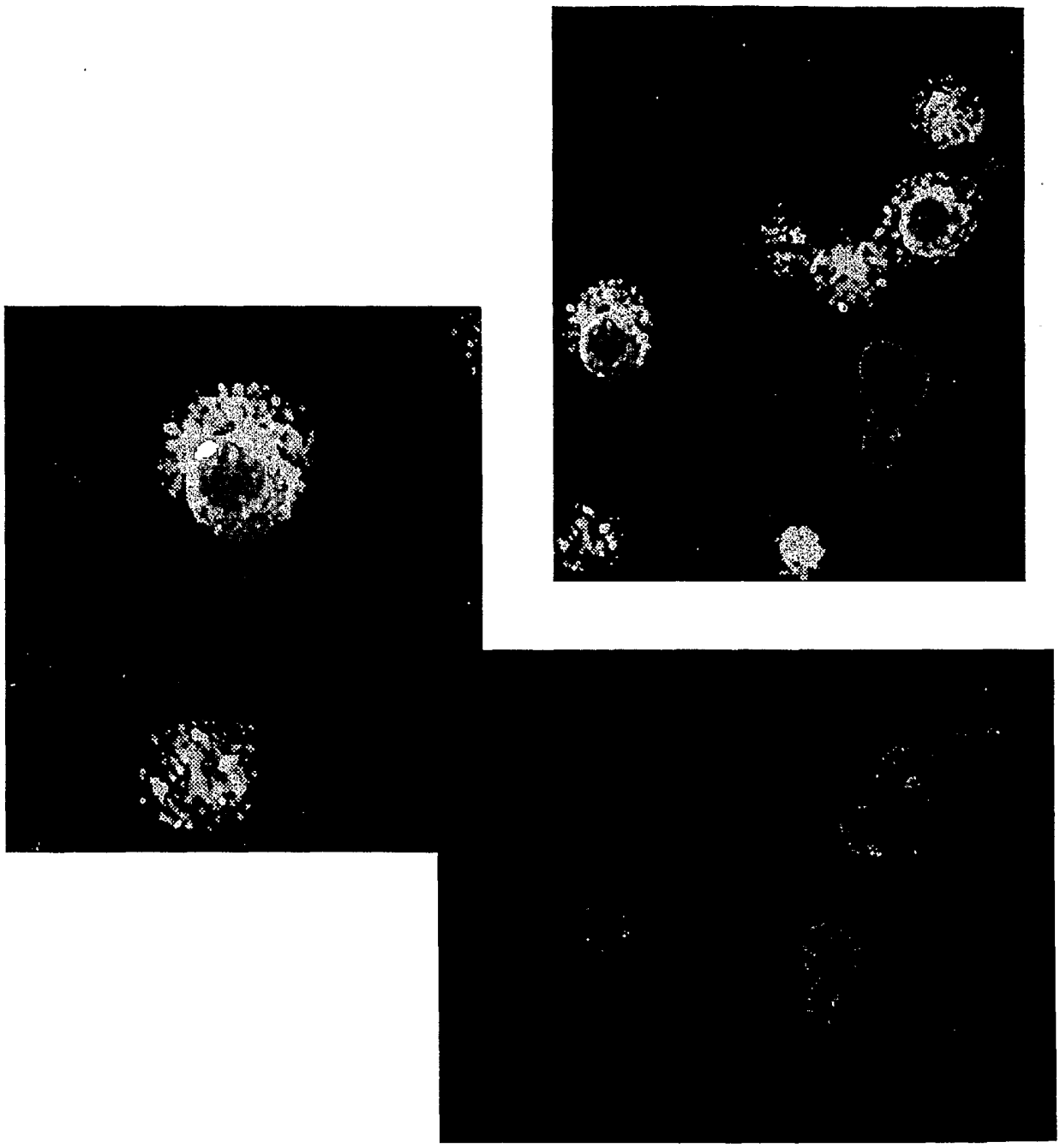


FIGURE 19C

Oligo-HRP-FISH Combined with X/Y DNA FISH in GPA-Coated CB Preparation



Probes: Gamma+ Epsilon Globin AS FITC (green)
X-red
Y-aqua

FIGURE 20

SEQUENCE LISTING

<110> Roche Diagnostics Corporation

<120> METHODS AND REAGENTS FOR IDENTIFYING RARE FETAL CELLS IN THE MATERNAL CIRCULATION

<130> 11012-004-228

<140> To be assigned

<141> 2001-11-15

<150> 60/248,882

<151> 2000-11-15

<160> 78

<170> PatentIn version 3.0

<210> 1

<211> 48

<212> DNA

<213> Homo sapiens

<400> 1
gagtcattct cgacttgcgg ccgcacgttt gatttccasc ttggtccc 48

<210> 2

<211> 48

<212> DNA

<213> Homo sapiens

<400> 2
gagtcattct cgacttgcgg ccgcacctag gacggtcagc ttggtccc 48

<210> 3

<211> 24

<212> DNA

<213> Homo sapiens

<400> 3
gcgatggttg ttgtcattgt cggc 24

<210> 4

<211> 17

<212> DNA

<213> Homo sapiens

<400> 4
caggaaacag ctatgac 17

<210> 5

<211> 21

<212> DNA

<213> Homo sapiens

<400> 5
ggtgctcttg gaggagggtg c 21

<210> 6

<211> 18

<212> DNA

<213> Homo sapiens

<400> 6
caggaaacaa gctatgac 18

<210> 7

<211> 17

<212> DNA

<213> Homo sapiens

<400> 7
gaattttctg tatgagg 17

<210> 8

<211> 678

<212> DNA

<213> Homo sapiens

<400> 8
caggtgaagc tgcagcagtc aggggctgaa ctggtgcagc ctggggcttc agtgaatttg 60

tcctgcaagg cctctggctt caccctcacc agctactata tgtactggtt gaagcagagg 120

cctggacaag gccttgagtg gatcggagag atcaacccta gcaatggtgt tactaatttt 180

aatgagaagt tcaagagcaa ggccacactg actgtagaca agtcctccag cacagcatac 240

atgctactca gcagcctgac atctgaggac tctgcggtct attactgtac aagatcacat 300

tactacggcc actgggatgt tatggactac tggggccaag ggaccacggt caccgtcacc 360

gtctcgagtg cctccaccaa gggcccatcg gtcttccccc tggcaccctc ctccaagagc 420

acctctgggg gcacagcggc cctgggctgc ctggtcaagg actacttccc cgaaccggtg 480

acggtgtcgt ggaactcagg cgccctgacc agcggcgtgc acaccttccc ggctgtccta 540

cagtcctcag gactctactc cctcagcagc gtggtgaccg tgccctccag cagcttgggc 600

acccagacct acatctgcaa cgtgaatcac aagcccagca acaccaaggt ggacaagaaa 660

gttgagccca aatcttgt 678

<210> 9

<211> 648

<212> DNA

<213> Homo sapiens

<400> 9
gacatcgagc tcactcagtc tccaacagtc atgtctgcat ctccagggga gaaggtcacc 60

ataacctgca gtgccagctc aagtgtaagt tacatgcact ggtaccagca gaagtcaggc 120

acctccccca aaagatggat ttatgacaca tccaaactgg cttctggagt ccctgctcgc 180

ttcagtggca gtgggtctgg gacctcttac tctctcacaa tcagcagcat ggaggctgaa 240

gatgctgcca cttattactg ccagcagtgg agtagtaacc cacccatgta cacgttcgga 300

gggggggacaa agttggaaat aaaacgggga actgtggctg caccatctgt cttcatcttc 360

ccgccatctg atgagcagtt gaaatctgga actgcctctg ttgtgtgcct gctgaataac 420

ttctatccca gagaggccaa agtacagtgg aaggtggata acgccctcca atcgggtaac 480

tcccaggaga gtgtcacaga gcaggacagc aaggacagca cctacagcct cagcagcacc 540

ctgacgctga gcaaagcaga ctacgagaaa cacaaagtct acgcctgcga agtcacccat 600

cagggcctga gctcgcccgt cacaaagagc ttcaacaggg gagagtgt 648

<210> 10

<211> 711

<212> DNA

<213> Homo sapiens

<400> 10
gggggggtttg ttgggctggt gagttttcga ggtgacggtg ctgtgctcgg tgaggggacg 60

cccagagagc cctgacccgg ggtcactccg tcgccgttct cctcttgtct acgtgctgga 120

cccggtgcta ccttttttacc cacacttaag tgacgcaaaa tgcccttcaa tggcgagaag 180

cagtgtgtgg gagaggacca gccaagcgat tctgattctt cccggttttc cgaaagcatg 240

gcttcgctca gtgactatga atgctccagg cagagcttta caagtgactc ctccagcaaa 300

tccagctctc ctgcttcaac aagccctcca agggttgtaa catttgatga agtgatggct 360

acagcaagga acttatcaaa cttgactctt gctcatgaga ttgctgtaaa tgagaacctt 420

caattgaaac aagaggctct cccagaaaag agtttggctg gtcgagtgaa gcacattgtt 480

caccaggcct tctgggacgt cttggattca gaactaaatg ctgaccctcc tgagattgaa 540

catgccatca aactgtttga agaaatcaga gagattcttc tctcttttct cactcccggt 600

ggcaaccggc ttcgcaacca aatctgtgaa gttttggaca cagacctcat taggcagcag 660

gctgagcaca gtgctgttga catccaaggc ctggccaact atgtcatcag t 711

<210> 11

<211> 3194

<212> DNA

<213> Homo sapiens

<400> 11
gggggggtttg ttgggctggt gagttttcga ggtgactgtg ctgtgctcgg tgaggggacg 60

cccagagagc cctgacccgg ggtcactccg tcgccgttct cctcttgtct acgtgctgga 120
cccggtgcta ccttttttacc cacacttaag tgacgcaaaa tgcccttcaa tggcgagaag 180
cagtgtgtgg gagaggacca gccaagcgat tctgattctt cccggttttc cgaaagcatg 240
gcttcgctca gtgactatga atgctccagg cagagcttta caagtgactc ctccagcaaa 300
tccagctctc ctgcttcaac aagccctcca agggttgtaa catttgatga agtgatggct 360
acagcaagga acttatcaaa cttgactctt gctcatgaga ttgctgtaaa tgagaacttt 420
caattgaaac aagaggctct cccagaaaag agtttggctg gtcgagtgaa gcacattgtt 480
caccaggcct tctgggacgt cttggattca gaactaaatg ctgaccctcc tgagtttgaa 540
catgccatca aactgtttga agaaatcaga gagattcttc tctcttttct cactcccggt 600
ggcaaccggc ttcgcaacca aatctgtgaa gttttggaca cagacctcat taggcagcag 660
gctgagcaca gtgctgttga catccaaggc ctggccaact atgtcatcag tacgatggga 720
aagctgtgtg ctcccgtgcg agataatgat atcagagagt taaaggctac tggcaacatc 780
gtggaggtgc tgagacaaat attccatgtc ctggacctca tgcaaatgga catggccaat 840
tttacaatta tgagtctcag accgcacctt caacgccagt tggtggaata tgagagaacc 900
aagttccagg aaattttgga agaaactcca agtgagtata atataatgtg tatttatatt 960
gaaattaggt taaaatgatg attttaaatg ttaaaattta aattaaatta aattaaatta 1020
aattaataaa taaataaaat ttaatttaat ttaaattaat taaaatttaa aatgttaaag 1080
gttaaaatga aaattaggtg atgttttcat ttttgacttc atatatagta tcactaaaga 1140
attttggcct ttctattaaa catttttta ccttccaaaa ttgacatcac aaacttacag 1200
ttgtttattc tgtccaggaa ctgtgctaag cactattcgt ttcttcattc attcaagata 1260
tatttattga gtgcctcctg tgtttcagac actgcttagg tgcagtggat actgcagcga 1320
acagaatgga atacagttgc cgtccccaag gagctgacat tctagtaggc atgacaaata 1380
aataaacata tattgtcagg tgtattatta agtggtatga agaaaaatag gataaaggga 1440
cagtaatggg tcagaagtga ggacagcatt gtttttaga tgtgatggac tgggaaggtc 1500
tctgtaagta acatttgagt aaagtttgca aagaaatgag gaaggggcat gtaaagatgg 1560
ggaaaagaac gttctagcct gagggactgc aggtacaaag gcactgaagc agaaactgct 1620
tatgagtttg acaaacacag ggcagacatt gtgacgggag tagagttggt aaggggagag 1680
tggtaagaga tgaggctgaa aaggagataa aggcaggtca catagagcat tgttagtaaa 1740
tgagaggagt ttggatttca ttccatgtgc ggtggaaact catcagagag ttttaatcag 1800
gggagcaaca tctgattttc atttttgaag aatattcttg ttcctggcta cagaataggt 1860
catgtgggaa caaagaggga aacagaggta ccaattaggt ggcattacag ttgctcaggc 1920
aagaaatgat ggtggcttgt tccagatggt aatggtggaa ggagtgagaa gagaccagat 1980

tctggataca ttttgaaagt ggagtaatca gatttgctga agattggata agggagaaaa 2040

gagttgaggg tgactctaag gtttttggcc tcgtacaacc gggtgaatgg tcatacagtt 2100

tgagtgaaac tggtgagacc aggagaaaga aggttgggag atgacatgga atcaagagtt 2160

cagtttggtg atgtaaactt ttgagatttc atttgacatc aatgtggaga tgtcaggtag 2220

gcatttagat atgaattacg agtttacagg aataatggaa actgtatatg tatttgggaa 2280

tcatcagtat atgtatgata tataaagcca tggaactggg tactagttaa cacttagaaa 2340

atccttatca cactctaagt actttacctg tacaaactaa tttaatcctt gcaacaacta 2400

tatgaagtta ccccccattt tatgggtgcg gaaactgagg cactgtgagg ttaagtcact 2460

ggcctatgta acacacttaa aagtatagga gctcacctag gaaggcagct tggctccata 2520

gtccatgttc ctgtccatta tattacactg ggtgagacta cctaggggga gtgagtgtgg 2580

actctgaaag aaactgacgt ggggcactcc agcagctaga agctgaagtt gagatgagga 2640

ggagggtcta acaatggaga ctatgaggga gtagccagag aggtagtgtg acagtcagga 2700

gaggcgggtg tcctgtaaga tacgtgaaga aggtatttgg agaaggaagg catgtttgga 2760

tttgacaagc attgctaagt ggtcaagaaa gacacaggcc aacttgacca ttggatttga 2820

ctgtttgatg atggtcagta accttgatga gagtggtttc agtggatttg tgtgcattaa 2880

agccagactg ggcagggcac agtggctcac acctgtaatc ctagcacttt gggaggccga 2940

ggcaggcaca tcacctgaga tcaggagttc aagaccagcc tggccaacat gctgaaaccc 3000

cttctctact aaaatctaaa aattagccgg gatgatggca ggtgcctgta atcccagcta 3060

ctcgggaggt tgagatggga gaatcgcttg aacccaggag atggtggttg cagtgagcca 3120

aggtcgcatc attgcactcc agcctggggg gctgagcaag actccgtctc ggaaaaaaaa 3180

aaaaaaaaaa aaaa 3194

<210> 12

<211> 827

<212> DNA

<213> Homo sapiens

<400> 12

gggtgaggtc gtgcagactg ctgctgcgtc cccttgccgg gacctttggt taaggggaag 60

aggtggagaa aagccgattg ggcctgggct cgaacgctga accctcaggt gttgcttgtg 120

cttcccaaag ctgctactgt aggtggcacc actgggggta atagcaatga tttgaggtgg 180

catatggatg aacatagtta cttgttttcg ttactattaa gaaaatctgt tccattaagt 240

aagcagtttc acaatgatag agtttccttt taaaataaat atatttaaat gcaaaaatct 300

agtcaagata gagacgcggt tactaaatta tagtaaaggt aaggagggag acgtagaaat 360

agtcatgaag agtggcactc aaacgactta agtttgggaa atcttagatg gttgtcctgc 420

taccttaact gtgaagaaac tggacaagtt gtccttgtcc ccagtgggtg tctcagcacc 480

catggtcacc agcccagtca gggactttgc tcttcttacc tgccaacgga taatgagtta 540

atctcagtgg ttcccacttg aggaggctcc ctgccgtggt ccttttcctc tcttctctag 600

ggatttcatg caccaccctg acagccagtg gaagttgcct gatttgaaga ggcccagggc 660

cgaagatacc cagtcatgtg tggaagcagc ttcaatgcca ttttgtgaga ggttgagggc 720

cccagagggg gaagcacagg tgctaccttt ttacccacac ttaagtgacg caaaatgccc 780

ttcaatggcg agaagcagtg tgtgggagag gaccagccaa gcgattc 827

<210> 13

<211> 153

<212> DNA

<213> Homo sapiens

<400> 13

acgcacgcca agtcttcaga gggcctggat gggtgtgact ccaaggagct aatcgtcccg 60

tgcaggtgct acctttttac ccacacttaa gtgacgcaaa atgcccttca atggcgagaa 120

gcagtgtgtg ggagaggacc agccaagcga ttc 153

<210> 14

<211> 544

<212> DNA

<213> Homo sapiens

<400> 14

tggcactcaa acgacttaag tttgggaaat cttagatggt tgtcctgcta ccttaactgt 60

gaagaaactg gacaagttgt ccttgtcccc agtgggtgtc tcagcaccca tggtcaccag 120

cccagtcagg gactttgctc ttcttacctg ccaacggatg atgagttaat ctcagtggtt 180

cccacttgag gaggctccct gccgtggtcc ttttcctctc ttctctaggg atttcatgca 240

tcaccctgac agccagtgga agttgcctga tttgaagagg cccagggccg aagataccca 300

gtcatgtgtg gaagcagctt caatgccatt ttgtgagagg ttgagggccc cagaggggga 360

agcacagtga cctgctatcg tgttggaggc tttgcttgga gtcatcttca gtcttttctg 420

ttggctttac ctttgaccag tgattaaatc ctccaggtgc tacctttta cccacactta 480

agtgacgcaa aatgcccttc aatggcgaga agcagtgtgt gggagaggac cagccaagcg 540

attc 544

<210> 15

<211> 378

<212> DNA

<213> Homo sapiens

<400> 15
ctcgtggggg ggcgcgcgat tatttgaaga cgctcacgga gcggctggct aggctgagga 60

gagctcgccg ggctctgagg cgcaggaatt caataaagaa aatggcagct cttactccaa 120

ggaagaggaa gcaggattct ttgaagtgtg acagcctttt acacttcact gaaaatctgt 180

ttccatcacc taataaaaag cactgttttt atcaaaacag tgataaaaat gaagaaaacc 240

tgcattgctc tcaacaagag cattttgttt taagtgcgct caaaacaact gaaataaata 300

gactgccatc agcaaatcaa ggctcaccat ttaaatctgc gctctccact gtatcttttt 360

acaaccaaaa taagtggt 378

<210> 16

<211> 2256

<212> DNA

<213> Homo sapiens

<400> 16
gggcgcagga attcaataaa gaaaatggca gctcttactc caaggaagag gaagcaggat 60

tctttgaagt gtgacagcct tttacacttc actgaaaatc tgtttccatc acctaataaa 120

aagcactgtt tttatcaaaa cagtgataaa aatgaagaaa acctgcattg ctctcaacaa 180

gagcattttg ttttaagtgc gctcaaaaca actgaaataa atagactgcc atcagcaaat 240

caaggctcac catttaaatc tgcgctctcc actgtatctt tttacaacca aaataagtgg 300

tacctcaatc cactggagag aaagctgata aaagagagta gatctacttg tctaaaaact 360

aatgatgaag ataaatcttt tcccattgtg acagaaaaaa tgcaaggaaa accagtctgc 420

tccaagaaga acaacaaaaa accacagaag agtttaactg ctaagtatca accaaagtat 480

agacacatca agcctgtatc aaggaattct agaaattcca agcaaaatcg agtgatctat 540

aagccaattg tggagaagga aaataattgt cattcagctg aaaataattc caatgctcct 600

cgggttctga gccaaaaaat aaaaccacaa gttacactcc agggtggagc agcatttttt 660

gttagaaaaa aatcttctct tagaaaatcg tccctggaaa atgagccgtc actgggacgc 720

acccaaaaga gtaaatcaga agtcattgaa gattctgatg tagagactgt cagtgaaaaa 780

aaaacttttg cgacaaggca agtgccaaag tgcttggtcc tagaagagaa attgaaaatt 840

ggactactga gtgcaagcag taaaaataaa gagaaattaa taaaggtaaa gctaaatata 900

tcactttaaa aatggctgta taacaaaact tcagtataaa tgacatagtt gaataaaatt 960

ttattttctg gactgctttt ataaagccag atagcatagt gtttagttac tgtaaggagt 1020

ttatttattt atttatttta agacagggtc tttctttatg acccaggctg tagtgcagtg 1080

gcacactgct cactacagcc tcaacctcct aggctcaagc aatcccacct cagcctccca 1140

agtagctggt actatagatg cacaccacct cacccggcta attcttgcat tttttgtaga 1200

ggctggggtt tcaccatgtt gcccaggctg gtctcgcact gctgggctat agtgatccac 1260

ctgcctcacc ctcccaaatt gctgggatta tgggtgtgaa ccactgcacc tggcccagcc 1320

acattttaag ttctcaatag ttacatgtag ctggtggcta ccatattaga ccatgctggt 1380

ctagaacttc tgttgctgac ctcctgttta tatccttcat cagaaaaaaa gttgagtttt 1440

gtggggcagc taggggcaga aagccattga agcaacctta agagaggttg tctccagatc 1500

agggcttcca tgtatggtgg cttgtttata ttcatcacaa cccttgtagc tggagaaatg 1560

ataggcttat atgataccca gtacactcag gcttatagtc atgtctttcc taacccacct 1620

ctcatcattg ccttcctcac tagtttctct caaattcaac cccagtccat atctctcgta 1680

atcacctaaa tttcaaacct cagtggcaaa gataataaat tggtggccag taaaccaaat 1740

ctcacccata aatgttttgc ttggcctaca taatttaaaa gttggagcca aaattatttt 1800

taaatacaag aaaaacattt ctggtgtctc atgaaatatt gaaagctgag ccttcatttc 1860

tacatggcaa taatcagagc tgaatatcag gctgagtatc ataccatttg acaaatatgg 1920

gttcttcatt tctctaagtc agtgtatatc tatctattta gatcagttta gttttgatca 1980

ttttacttac tgctgccctt gtaaggcatt taagttttta gtcctcactg taaatgtttg 2040

tttttttcc tgtttatcga gtatggccaa gtagattgtt aaatgagata cgttcttgct 2100

agctaagcca tgacagtcta attttactaa ttcattttct gtttatatca acaaaacact 2160

tctgtgtctt gatagggatt ttgattataa gagtaagctt tatcaaaatg tattaaactg 2220

tgctttatac ttaaaaaaaa aaaaaaaaaa aaaaaa 2256

<210> 17

<211> 2170

<212> DNA

<213> Homo sapiens

<400> 17
agatttttac ctcaccaggg cgcgagaatc ttggaaacag ccaccagggt gggcggggca 60

agggctggtg ctccaccaat caccgagcca gcccctggaa gcgctgagcc gtggccaatc 120

ggaacgcagc ggcttcctcc tagcctggcg cgcgattatt tgaagacgct cacggagcgg 180

ctggctaggc tgaggagagc tcgccgggct ctgaggcgca ggtaacctct ggagtaggct 240

gaggcggggg gctgtggaag gctgaggcga gcgggagacc ctgtggaacg tgggggcgaa 300

tgtcacgggg aaagagcttt tgctcggcgc tcggcaggtc cgcaggcccg agaccgaaga 360

agccactgac tcgttctcgg cgctcacagg cgccgggttt ccctcttccc gacgccggcg 420

tttccttgct ttgggtttag ggtttgtttg tttgtttgtt tccctgaccg ggttcagggt 480

cgggacgctt tcctctccta cctttcctcg cttagcccca gaggggtccc cccttaagcc 540

gcagctcgac ctgaattccc agagaccctg gcctttcagt ttgttcattc attaatagct 600

ttgaaataac gcttgtgtgc tgctttgggg ggaatttccc tgatcccagg gaaggggaag 660

gaaatgaatt ctgacgagcc acgtagccgc ctgctcgtgc attctagcat ttgttttcag 720

cagccttacg acttcaagag atttttataa aaagggaagc aaacggcttg ctccgggctc 780

atcagctagg aaaatggcag aggactgtgc ttcacccgtt ttcttgctgg gctctgatga 840

aaagttacag acggtctctg ccttcattga aggctgttga gatggggaaa caagaaaatg 900

tcttaaatgg actattaggg aaatacaggg tgccatggaa gcagaacaag cgcgtgtaac 960

cctgggaggg gtggaaggag agcaatgtcg aggaagacct cccagaaaat agggttttaa 1020

cttggtgtga aatagactaa ttaaaggggg tataattttg ataaagcgat ttttaatatt 1080

ttgatgaatg tggttattgt catttctttt aggaattcaa taaagaaaat ggcagctctt 1140

actccaagga agaggaagca ggattctttg aagtgtgaca ggtgaatctc agcctgtgaa 1200

tagaaactct tagaaaaatc caccttcttg tctctcttgt tctctcctat tttctaaaat 1260

tttcgttctt ccactagcct actccttgtg gctacagtga taacctgata acctatattc 1320

taattctagc ttggaaatac attaagtttt cataatgaat tatttagtcc ttttcaaggg 1380

aatcaaagct gtttttttagt ttttgttttt ttcttttgac atagaagcag cacattgaag 1440

aaagctgttt tctatagcta ggaaatatac caaatcagct attacctttt tcctttctgt 1500

tcccctataa tattttacag tgttttgcat acagtagatg cttgtctgga aacaattcta 1560

gaaagactgt gtactcagat attaattact ctacccagga aaagccccag ggggtcagtt 1620

cattacagac agttattata gggggcttgt agacaataga gcaaaatata tgagctataa 1680

tttatatatg gccaacacat agaaatcgat tttaaaacat acatgcatat atgtctatca 1740

tagcttcttt gtatatagta tttgttttca tgaactcttt gggaagatat ttgaatgttt 1800

tatttgataa agacagaatt tgagaatcct actgttaaac taatagaaaa ttgtgtgtat 1860

gggggggtacg tgtattgttt ttaatatgac ctacaagtag agcttactta gcagtagatt 1920

tatgtaaatt ttgacgcaaa ataatcttat caatggactt tgtttctttt tatagccttt 1980

tacacttcac tgaaaatctg tttccatcac ctaataaaaa gcactgtttt tatcaaaaca 2040

gtgataaaaa tgaagaaaac ctgcattgct ctcaacaaga gcattttgtt ttaagtgcgc 2100

tcaaaacaac tgaaataaat agactgccat cagcaaatca aggctcacca tttaaatctg 2160

cgctctccac 2170

<210> 18

<211> 487

<212> DNA

<213> Homo sapiens

<400> 18

agatttttca cctcaccagg gcgcgagaat cttggaaaca gccaccaggg tgggcggggc 60

aagggctggt gctccaccaa tcaccgagcc agcccctgga agcgctgagc cgtggccaat 120

cggaacgcag cggcttcctc ctagcctggc gcgcgattat ttgaagacgc tcacggagcg 180

gctggctagg ctgaggagag ctcgccgggc tctgaggcgc aggaattcaa taaagaaaat 240

ggcagctctt actccaagga agaggaagca ggattctttg aagtgtgaca gccttttaca 300

cttcactgaa aatctgtttc catcacctaa taaaaagcac tgtttttatc aaaacagtga 360

taaaaatgaa gaaaacccgc attgctctca acaagagcat tttgttttaa gtgcgctcaa 420

aacaactgaa ataaatagac tgccatcagc aaatcaaggc tcaccattta aatctgcgct 480

ctccact 487

<210> 19

<211> 3375

<212> DNA

<213> Homo sapiens

<400> 19

atttgaagac gctcacggag cggctggcta ggctgaggag agctcgccgg gctctgaggc 60

gcaggaattc aataaagaaa atggcagctc ttactccaag gaagaggaag caggattctt 120

tgaagtgtga cagcctttta cacttcactg aaaatctgtt tccatcacct aataaaaagc 180

actgttttta tcaaaacagt gataaaaatg aagaaaacct gcattgctct caacaagagc 240

attttgtttt aagtgcgctc aaaacaactg aaataaatag actgccatca gcaaatcaag 300

gctcaccatt taaatctgcg ctctccactg tatcttttta caaccaaaat aagtggtacc 360

tcaatccact ggagagaaag ctgataaaag agagtagatc tacttgtcta aaaactaatg 420

atgaagataa atcttttccc attgtgacag aaaaaatgca aggaaaacca gtctgctcca 480

agaagaacaa caaaaaacca cagaagagtt taactgctaa gtatcaacca aagtatagac 540

acatcaagcc tgtatcaagg aattctagaa attccaagca aaatcgagtg atctataagc 600

caattgtgga gaaggaaaat aattgtcatt cagctgaaaa taattccaat gctcctcggg 660

ttctgagcca aaaaataaaa ccacaagtta cactccaggg tggagcagca ttttttgtta 720

gaaaaaaatc ttctcttaga aaatcgtccc tggaaaatga gccgtcactg ggacgcaccc 780

aaaagagtaa atcagaagtc attgaagatt ctgatgtaga gactgtcagt gaaaaaaaaa 840

cttttgcgac aaggcaagtg ccaaagtgct tggtcctaga agagaaattg aaaattggac 900

tactgagtgc aagcagtaaa aataaagaga aattaataaa ggattcatca gatgacagag 960

tttcttcaaa ggaacataaa gttgataaaa atgaggcttt ttcttcagag gattctcttg 1020

gtgagaataa gacaatttct cctaagtcca ctgtctatcc aatcttcagt gcatcttcag 1080

tcaattcaaa aagatcttta ggtgaagaac agttttctgt gggatctgtc aacttcatga 1140

aacagaccaa tatccagaaa aatactaata ccagagatac aagtaaaaaa acaaaagacc 1200

agctcatcat cgacgctggt cagaaacatt ttggggctac tgtgtgcaag tcttgtggta 1260

tgatatatac tgcttccaac cctgaagatg aaatgcagca tgtacagcat caccacaggt 1320

ttctggaagg aatcaaatat gtgggttgga agaaagaacg tgtagtagca gagttttggg 1380

atgggaaaat cgtgttggtt ctgccacatg atccaagctt tgctaccaaa aaggtagaag 1440

atgtccaaga acttgttgat aatgaattgg gcttccagca agttgttcct aaatgtccaa 1500

acaaaataaa aacttttctt tttatatctg atgaaaagag agtagttggg tgtttaattg 1560

cagaacccat caaacaggca tttcgtgtcc tgtctgaacc aattggtcca gaatccccaa 1620

gctctacgga atgtcctagg gcttggcaat gttcagatgt accagaacct gcagtctgtg 1680

ggataagtag aatctgggtt ttcagactga agagaagaaa gcgcattgca agacgactgg 1740

ttgataccct caggaattgc ttcatgtttg gctgttttct cagcactgat gaaatagcat 1800

tttctgaccc aacaccagat ggcaagttat ttgcaaccaa gtactgcaac acccctaatt 1860

tcctcgtata taattttaat agttaaagct gatttcagtt ataaaggagt tactatctgg 1920

ataagttcaa agagctcctt attataaaat acaaactatt taatatcaaa ataaaaaata 1980

ccgagactca cactcataca cacacacaca cacacacgca cacacacata tcacagtttt 2040

gttccttatg agttgaaaag tcaggaataa atttgttgaa aattatctgg ggattcaaag 2100

gaaaaatctt tgggtgattc cctgattagc actctgaatg tttaattatg aaactttgta 2160

gctataactg gaaaattacc tgactctttg taagagtatt aaatacaaag tgatttttct 2220

ctagaaatgt gacctggtct tttataaagc ccactcttag accaggatta tctaatgcca 2280

catcagaagc aaacaggcaa atttaaactt gggcaagtaa tttctgtgcc caatttgtaa 2340

agggaattcc tgaatttttt ttttttttta atagaggcat gggtctcact gtgttgccca 2400

ggctggtctg aaacttttgg gctcaagcga tcctcccaaa acgctgggat tacagtcatg 2460

agccaccgtg cccagcctaa ttcctgactt ctctatacag agtcttcact tgataggcac 2520

tcgtctgtag taactcagtt tgaatatctt tagaaaatgt ttagaattta tttgtaacaa 2580

gatggtaagg aataagatta tcccatatgc atttctgtag agcagaattt gatagcttag 2640

tgttcaatct ttttgaaaat aaatgtttac ctgtcatcag atttaattaa aattatactt 2700

agtaattgca ctattactta gttaattttt gttgtatgga aatattggta gtactacttt 2760

gggaacctgt tactgacaat tgatgtcatt aacaaaatgc ctagttggat tagatgtttt 2820

cattttctaa ttttttgctt gtttaaaatg caccttactt gttctgagat acctggcaaa 2880

agtctttaca aaatgtatgg taatagaacc aaggttagta aatatacata ggctggtgga 2940

tgagagacca tggaactgtg taaatacact taaatgttca cacatttttc tagtgtaatt 3000

cttggatact ttaaaaagca aaacattgtt caaattgttt tgattctgaa aaatcattca 3060

actgctaact ggcaataaga ctctaggcaa gtcgttttcc agattgtaat tatatgtaga 3120

aactattcat ctgcattcat tttatttgcc tgtaagttaa catgtttcca aaatttaaaa 3180

gcctgggtcc ccaaaagaat gtggaagtat taaaatgtat gtaattatgc aaacatttta 3240

atgctatttt ctgcacttat ttcttttaaa tattttattt aaaattttta attaacattt 3300

tgtttgctta atgcttttgt tatgaatcaa ttaaaattct ttattttata caactaaaaa 3360

aaaaaaaaaa aaaaa 3375

<210> 20

<211> 1742

<212> DNA

<213> Homo sapiens

<400> 20

cgttcttgct agctaagcca tgacagtcta attttactaa ttcattttct gtttatatca 60

acaaaacact tctgtgtctt gatagggatt ttgattataa gagtaagctt tatcaaaatg 120

tattaaactg tgctttatac ttatacattt gtatgaaaat taagtttttt gaaagtttgg 180

ggcgtaactc aagctaaaat tgtctcaatc cagtaaaaca aaggaagagt ctgcagtttc 240

cccatgcatc tcctgatgat tattcttatc tgtttggtac agcctggtgg tgtaccaaat 300

gaatggtgac ctgatgaatt ctctgtaacc agtgtaacag agaattctgt aattctcttt 360

tctgtcttat tttctttttc ttagcttcac tttacgggtt tggtattgag ttagttttca 420

gtagtttggg aagctgggct tttgtgcatt ttaaatttgt ggaacggggc tatgctcttt 480

caagattttg ctaaataccc atgtctggga ttcgtctctg ctagctaagg gcgaatctga 540

ttttctgttt ttttattttg agatggagtc tcactccgtt gcccaggctg gagtgcagtg 600

gcatgatctc aactcactgc aacctccgcc tcccaggttc aagcaattct ctgcctcagc 660

ctcccaagta gctggtattg taggcaccgg ccaccacacc tggctaattt ttgtattttt 720

agtaggaacg gggtttcacc ctctcagcca ggctggtctt gaactcctga ctttgtgatc 780

cacctgcctt ggccacccaa agtgctggta ttacaagtgt gagccaccgc gcctggctgg 840

tgaatctgat ttttgtgagg ataccctttc cccaactcat ggagccctga aaagtgttaa 900

actgatgggt tatcccacct ctgctgggat attatggttt tgagtagaag ttctacttta 960

accaagagtt tttttccttt ttccccaagc tatatattcc ctagtaattc ccaaattagt 1020

atgagggaaa tctgtcagga tttttttcta ctttgctttt cccacattcc tgctactctc 1080

cagtgactct agaatgttct ttccttagct gctgattttt aaagtttatg gcaaaaagtt 1140

gtccctttcc agtttggtgg ctgctgggga aacttttgcc atttttactg gctttttttgt 1200

tcatttcatt atggaagtga tggtctgtaa tcattctgcc atctctgtca ggaagtataa 1260

cacatttttg agaggatgac agattacaac acgtgtactc tggctttttt ttcttttgag 1320

acagtcttgc tctgttgcca ggctggagtg cagtggcacg atcttggctc actgcaacct 1380

ccgcctcctg ggttcaagca attctcctgc ctcagcctcc cgagtagctg ggactgcagg 1440

tgccaccacg cccagctaat ttttgtattt ttagtagaga tagggtttca ccatgttagc 1500

caggatggtc tcaatccctt gacctcgcaa tctacccacc tcggcctccc aaagtgctgg 1560

cattacaggc atgagccacc gggccctgct gtactctggc tttaaaggta atgtacgctt 1620

agtgaataca ggggaagttt tcctaaaaat agtttgaaaa ttgattgaaa atcacattat 1680

gaaatgtaat tcataataaa ccaaatcttt tgtttttgaa aaaaaaaaaa aaaaaaaaaa 1740

aa 1742

<210> 21

<211> 1068

<212> DNA

<213> Homo sapiens

<400> 21

acgggtcgcg agaggttgtt cgcgccttga gagttaagcg aagtgtggtg gcttccaagg 60

aatacaaaca taaaggcctt cgaccgttgc aaatagacta aagtgaaaac aaatctgaat 120

gaagatgaag ttatttcaga ccatttgcag gcagctcagg agttcaaagt tttctgtgga 180

atcagctgcc cttgtggctt tctctacttc ctcttactca tgtggccgga agaaaaaaag 240

tgaacccata tgaagaagtg gaccaagaaa aatactctaa tttagttcag tctgtcttgt 300

catccagagg cgtcgcccag accccgggat cggtggagga agatgctttg ctctgtggac 360

ccgtgagcaa gcataagctg ccaaaccaag gtgaggacag acgagtgcca caaaactggt 420

ttcctatctt caatccagag agaagtgata aaccaaatgc aagtgatcct tcagttcctt 480

tgaaaatccc cttgcaaagg aatgtgatac caagtgtgac ccgagtcctt cagcagacca 540

tggcaaaaca acaggttttc ttgttggaga ggtggaaaca gcggatgatt ctggaactgg 600

gagaagatgg ctttaaagaa tacacttcaa acgtcttttt acaagggaaa cggttccacg 660

aagccttgga aagcatactt tcaccccagg aaaccttaaa agagagagat gaaaatctcc 720

tcaagtctgg ttacattgaa agtgtccagc atattctgaa agatgtcagt ggagtgcgag 780

ctcttgaaag tgctgttcaa catgaaacct taaactatat aggtctgctg gactgtgtgg 840

ctgagtatca gggcaagctc tgtgtgattg attggaagac atcagagaaa ccaaagcctt 900

ttattcaaag tatatttgac aacccactgc aagttgtggc atacatgggt gccatgaacc 960

atgataccaa ctacagcttt caggttcaat gtggcttaat tgtggtggcc tacaaagatg 1020

gatcacctgc ccacccacat ttcatggatg cagagctctg ttcccagt 1068

<210> 22

<211> 2186

<212> DNA

<213> Homo sapiens

<400> 22

acgggtcgcg agaggttgtt cgcgccttga gagttaagcg aagtgtggtg gcttccaagg 60

aatacaaaca taaaggcctt cgaccgttgc aaatagacta aagtgaaaac aaatctgaat 120

gaagatgaag ttatttcaga ccatttgcag gcagctcagg agttcaaagt tttctgtgga 180

atcagctgcc cttgtggctt tctctacttc ctcttactca tgtggccgga agaaaaaaag 240

tgaacccata tgaagaagtg gaccaagaaa aatactctaa tttagttcag tctgtcttgt 300

catccagagg cgtcgcccag accccgggat cggtggagga agatgctttg ctctgtggac 360

ccgtgagcaa gcataagctg ccaaaccaag gtgaggacag acgagtgcca caaaactggt 420

ttcctatctt caatccagag agaagtgata aaccaaatgc aagtgatcct tcagttcctt 480

tgaaaatccc cttgcaaagg aatgtgatac caagtgtgac ccgagtcctt cagcagacca 540

tgacaaaaca acaggttttc ttgttggaga ggtggaaaca gcggatgatt ctggaactgg 600

gagaagatgg ctttaaagaa tacacttcaa acgtcttttt acaagggaaa cggttccacg 660

aagccttgga aagcatactt tcaccccagg aaaccttaaa agagagagat gaaaatctcc 720

tcaagtctgg ttacattgaa agtgtccagc atattctgaa agatgtcagt ggagtgcgag 780

ctcttgaaag tgctgttcaa catgaaacct taaactatat aggtctgctg gactgtgtgg 840

ctgagtatca gggcaagctc tgtgtgattg attggaagac atcagagaaa ccaaagcctt 900

ttattcaaag tacatttgac aacccactgc aagttgtggc atacatgggt gccatgaacc 960

atgataccaa ctacagcttt caggttcaat gtggcttaat tgtggtggcc tacaaagatg 1020

gatcacctgc ccacccacat ttcatggatg cagagctctg ttcccagtac tggaccaagt 1080

ggcttcttcg actagaagaa tatacggaaa agaaaaagaa ccagaatatt cagaaaccag 1140

aatattcaga atagggagca agttgctatt tgggaacatt cagcaccttc tcacagtttg 1200

ggaacatata ttgctgttta ctccagtgta aaaatgaggt gccactggat ctgagtgcta 1260

cacgaacaca agtagaagta ttaatttgtt gaaatgtgtt gttaccaaaa agactgaaaa 1320

gccccaaagt ctagatataa agacctagac ttcggcacgc gaaatcccag ctatgctacc 1380

tcttatttac ctgaaaggag gacacgcagg atgggcagtc atgctggtga ctcttgtact 1440

cccttgaggg acattggtgg gggggggggg gcgtggtccc aggcaggatg cccagtcttt 1500

gagctgagat tggaaggcag tgaggctgag ggtgccaaga tttccccagg gttcacccag 1560

aggggaaggg gctacatgcc cccagctgtg tgcagggagg acacatcagc ccactaccgc 1620

tgccaacacc aatgcctaaa acttgtttca tacattgggg ttttctatat atttcagttg 1680

ggaaaagctt acatttaacc ttttgaaaaa ataaatacgt gattagcctc aactaaacat 1740

tgctgactat aaagacagta tattcaccat gtcgctggca atatgtcatt gcgtaacacc 1800

aaataacccc ccagaagtag ccagaggcca gtttgaacat cacaattcta agtgttttag 1860

taactatttc tggcgtgagt caacagatca tgtagataga gtcaattatt gtttgtggag 1920

tttttcagct ataggggagg ggaactatta aaatccattt gtttctattc aataggtaat 1980

aaaaattagt tgtccctggg tttgggaaac ttaaatgccc attacagccc tggggaaggg 2040

ttttctgtct tatggagtga gtcttagcat ttaagttata cagttgctgc cttaaaatag 2100

tagcctgcta caatgacttc tttgggtagc cattttcata agaaataaaa tacaagatat 2160

gagtaaaaaa aaaaaaaaaa aaaaaa 2186

<210> 23

<211> 409

<212> DNA

<213> Homo sapiens

<400> 23
ggcctaattc gaaaccaaag cgcgggacgg atgaaagtac gggtcgcgag aggttgttcg 60

cgccttgaga gttaagcgaa gtgtggtggc ttccaaggaa tacaaacata aaggccttcg 120

accgttgcaa atagactaaa gtgaaaacaa atctgaatga agatgaagtt atttcagacc 180

atttgcaggc agctcaggag ttcaaagttt tctgtggaat cagctgccct tgtggctttc 240

tctacttcct cttactcatg tggccggaag aaaaaagtga acccatatga agaagtggac 300

caagaaaaat actctaattt agttcagtct gtcttgtcat ccagaggcgt cgcccagacc 360

ccgggatcgg tggaggaaga tgctttgctc tgtggacccg tgagcaagc 409

<210> 24

<211> 1159

<212> DNA

<213> Homo sapiens

<400> 24
acagcatgaa gagttcattc ttctgagtca aggagaggtg gaaaagctaa tcaagtgcga 60

cgaaattcag gtggattctg aagagccagt ctttgaggct gtcatcaact gggtgaagca 120

tgccaagaaa gagcgggaag aatccttgcc taacctgcta cagtatgtgc ggatgcccct 180

actaaccccc aggtatatca cagatgtaat agatgctgag cctttcatcc gctgtagttt 240

acaatgcagg gatctggttg atgaagcaaa gaagtttcat ctgaggcctg aacttcggag 300

tcagatgcag ggacccagga caagggctcg cctaggagcc aatgaagtgc ttttggtggt 360

tgggggcttt ggaagccagc agtctcccat tgatgtggta gagaaatatg accccaagac 420

tcaggagtgg agctttttgc caagcatcac tcgtaagaga cgttatgtgg cctcagtgtc 480

ccttcatgac cggatctacg tcattggtgg ctatgatggc cgttcccgcc ttagttcagt 540

ggaatgtcta gactacacag cagatgagga tggggtctgg tattctgtgg cccctatgaa 600

tgtccgacga ggtcttgctg gagccaccac cctgggagat atgatctatg tctctggagg 660

ctttgatgga agcaggcgtc acaccagtat ggagcgctat gatccaaaca ttgaccagtg 720

gagcatgctg ggagatatgc agacagcccg ggaaggtgcc ggactcgtag tggccagtgg 780

agtgatctac tgtctaggag gatatgacgg cttgaatatc ttaaattcag ttgagaaata 840

cgaccctcat acaggacatt ggactaatgt tacaccaatg gccaccaagc gttctgatat 900

gatggtaatt ccctgctaag tagcattgaa tgttatgacc ctatcatcga cagctgggaa 960

gtcgtgacat ccatgggaac ccagcgctgt gatgctggtg tttgtgctct ccgcgagaag 1020

tgaccattgt tggagcacca tccagagcta gtgaccagtc cagtggacag ttagtgggag 1080

aatcaaaaat cctttccaga atgtctgttt ctcactacgt gcaccgggtg attacaggca 1140

ccagtgcagt gatgattgt 1159

<210> 25

<211> 3561

<212> DNA

<213> Homo sapiens

<400> 25

cacacctggc cgtttattga gttttaagag ttgtttatat atgtatttat ttaatactag 60

tcctttgtca gataggtggt ttgcaaatat tttttaattt tgatgaagtc cagcttatca 120

atttttcttt ttatggctca tgcttttggg gtcaagtcta agaattcctt gcctagccct 180

agctagatcc tcaagatttt cttctggtca tcacgtcgcg tggccgcagg gagcagaccc 240

ggacagctcc agagcctccg ggccggggcg gcggcggcga cgcttcggct cctcctgagc 300

cacctgctgg acccgcaccc cactccatcc ccacaggctg gggacaggcc ctggcgcggc 360

tgtgtgggat cagaagcaga gttgcagaat ccaaggacct atttttgttc tttctccgca 420

ctgctttatg ggaggcatta tggcccccaa agacataatg acaaatactc atgctaaatc 480

aatcctcagt tcaatgaact ccgttgggaa gagcaatacc ttctgtgatg tgacattgag 540

tagagcagaa agactttcct gcccatgaga ttgtgctggc tgcctgtagt gattacttct 600

gtgccatgtt cactagtgac ctttatagaa ggggaaaccc tatgttgaca tccaaggttt 660

gactgcctct accatggaaa ttttattgga ctttgtgtac acggaaacag tacatgtgac 720

atggagaatg tacaggaact gcttcctgca gcctgtctgc ttcagttgaa aggtgtgaaa 780

caagcctgct gtgagttctt agaaagtcag ttggaccctt ctaattgcct gggtattagg 840

gattttgctg aaacccacaa ttgtgttgac ctgatgcaag cagctgaggt ttttagccag 900

aagcattttc ctgaagtggt acagcatgaa gagttcattc ttctgagtca aggagaggtg 960

gaaaagctaa tcaagtgcga cgaaattcag gtggattctg aagagccagt ctttgaggct 1020

gtcatcaact gggtgaagca tgccaagaaa gagcgggaag aatccttgcc taacctgcta 1080

cagtatgtgc ggatgcccct actaaccccc aggtatatca cagatgtaat agatgctgag 1140

cctttcatcc gctgtagttt acaatgcagg gatctggttg atgaagcaaa gaagtttcat 1200

ctgaggcctg aacttcggag tcagatgcag ggacccagga caagggctcg cctaggagcc 1260

aatgaagtgc ttttggtggt tgggggcttt ggaagccagc agtctcccat tgatgtggta 1320

gagaaatatg accccaagac tcaggagtgg agctttttgc caagcatcac tcgtaagaga 1380

cgttatgtgg cctcaatgtc ccttcatgac cggatctacg tcattggtgg ctatgatggc 1440

cgttcccgcc ttagttcagt ggaatgtcta gactacacag cagatgagga tggggtctgg 1500

tattctgtgg cccctatgaa tgtccgacga ggtcttgctg gagccaccac cctgggagat 1560

atgatctatg tctctggagg ctttgatgga agcaggcgtc acaccagtat ggagcgctat 1620

gatccaaaca ttgaccagtg gagcatgctg ggagatatgc agacagcccg ggaaggtgcc 1680

ggactcgtag tggccagtgg agtgatctac tgtctaggag gatatgacgg cttgaatatc 1740

ttaaattcag ttgagaaata cgaccctcat acaggacatt ggactaatgt tacaccaatg 1800

gccaccaagc gttctggtgc aggagtagcc ctgctgaatg accatattta tgtggtgggg 1860

ggatttgatg gtacagccca cctttcttcc gttgaagcat acaacattcg cactgattcc 1920

tggacaactg tcaccagtat gaccactcca cgatgctatg taggggccac agtgcttcgg 1980

gggagactct atgcaattgc aggatatgat ggtaattccc tgctaagtag cattgaatgt 2040

tatgacccta tcatcgacag ctgggaagtc gtgacatcca tgggaaccca gcgctgtgat 2100

gctggtgttt gtgttctccg cgagaagtga ccgttgttgg agcaccatcc agagctagtg 2160

accagtccag tggacagtta gtgggagaat caaaaatcct ttccagaatg tctgtttctc 2220

actatgtgca ccgggtgatt acaggcacca gtgcagtgat gattgtactt atttgacaca 2280

tactccccgt cgtcctggtt cttgttcctg agaagggtgg gtaacagata ttccaggaaa 2340

aagaatgcac attgaatgga tgtgagagac cacattgcct ctcccactgc tttggggagc 2400

actttcctgt catttctaac ttaccacatg cttggtgtac tatatgtatg ttgtgcctca 2460

tatgttgcaa agaactaagg tgagtatagc ctactagata tgggcaatat ccagcctaga 2520

tgattggaaa gataccagtt taagtaaact tggtaaaatc caagtctttt ttttttttt 2580

ccaggaacaa ctacattttc tcatatacag gtagctaggg gcaacacagt tccattctag 2640

agggaaacaa aagggagagc cccacaaaac tttggggaca agggagagag agactcatct 2700

gacacttctt ttggaggtca ggatttgtat atcagaattg aagttagaat taagtgaatt 2760

aaactgaatt tgattgtgag tgaacctaga acagcactga agtattacat aacctggaag 2820

actgagaagg gtatattatt tgaaggatct ttttatttcc ccgaggtctt tcgcactgga 2880

gacagcataa aagagtgaac aaatgttggg atgagagaag atgacatcaa tgtgggagtt 2940

cagtataact ggggataaac tagaagaacc tgtgatttta cagtcatctt attacctgcc 3000

agggctcatc tagccatggc aatgtttgcc ttgaatgggg gtgaaagcct ttctttgttg 3060

gaccaaatac tactacacta ttacacttcc acactattta tttggggatg ggctgggagt 3120

gacagtagcc tagtagttca gctacctgat tactgcccca ttcttttaga agcacatgtc 3180

tgccaaggag tggtttgtac tgctgtgttt ggtacatcta gtctttttc tgctataagt 3240

tttccttacc tgtcctttag tgtagatttt attcatcaca ggacagaata atcaaggaca 3300

accaaaatcc ttttgttagt ttcagtacct cagctatcaa catttctgag ctaccattca 3360

atgttcctct gtgtcatgga gtgaaattct tgttttgtgg gtattaggag tgtgggaatg 3420

tgataaccta aacaaccttt gctctgaaat tccatttttc cctctttccc tgagttgtat 3480

tgacctacag agttaatttc ctttgtattt ttttaagaaa atattaaaaa tcaacggtct 3540

caaaaaaaaa aaaaaaaaaa a 3561

<210> 26

<211> 1440

<212> DNA

<213> Homo sapiens

<400> 26

aggggccaga cccggacggc tccagagcct ccagagcctc cgggtctggg cggcgcttcg 60

gctcctcccg agccgcctgc tagccccgcg ccgcactcca tccccacagg ctggggacgg 120

gccaggtgcg gctgtgtggg ttcgggagcg gagttgcaga atccaaggac ccattttgtt 180

ctttctccgc actgctttat gggaggcatt atggccccca aagacataat gacaaatact 240

catgctaaat ccatcctcaa ttcaatgaac tccctcagga agagcaatac cctctgtgat 300

gtgacattga gagtagagca gaaagacttc cctgcccatc ggattgtgct ggctgcctgt 360

agtgattact tctgtgccat gttcactagt gagctctcag agaaggggaa accttatgtt 420

gacatccaag gtttgactgc ctctaccatg gaaattttat tggactttgt gtacacagaa 480

acagtacatg tgacagtgga gaatgtacaa gaactgcttc ctgcagcctg tctgcttcag 540

ttgaaaggtg tgaaacaagc ctgctgtgag ttcttagaaa gtcagttgga cccttctaat 600

tgcctgggta ttagggattt tgctgaaacc cacaattgtg ttgacctgat gcaagcagct 660

gaggttttta gccagaagca ttttcctgaa gtggtacagc atgaagagtt cattcttctg 720

agtcaaggag aggtggaaaa gctaatcaag tgcgacgaaa ttcaggtgga ttctgaagag 780

ccagtctttg aggctgtcat caactgggtg aagcatgcca agaaagagcg ggaagaatcc 840

ttgcctaacc tgctacagta tgtgcggatg cccctactaa cccccaggta tatcacagat 900

gtaatagatg ctgagccttt catccgctgt agtttacaat gcagggatct ggttgatgaa 960

gcaaagaagt ttcatctgag gcctgaactt cggagtcaga tgcagggacc caggacaagg 1020

gctcgcctag atatgatcta tgtctctgga ggctttgatg gaagcaggcg tcacaccagt 1080

atggagcgct atgatccaaa cattgaccag tggagcatgc tgggagatat gcagacagcc 1140

cgggaaggtg ccggactcgt agtggccagt ggagtgatct actgtctagg aggatatgac 1200

ggcttgaata tcttaaattc agttgagaaa tacgaccctc atacaggaca ttggactaat 1260

gttacaccaa tggccaccaa gcgttctggt gcaggagtag ccctgctgaa tgaccatatt 1320

tatgtggtgg ggggatttga tggtacagcc cacctttctt ccgttgaagc atacaacatt 1380

cgcactgatt cctggacaac tgtcaccagt atgaccactc cacgatgcta tgtaggggcc 1440

<210> 27

<211> 1126

<212> DNA

<213> Homo sapiens

<400> 27

acctagagcc cagctgaaca acaaggcttt gggtgtgaag ggactcccca gcctggagac 60

cctatttggc tgaaacagtt acaaaatatc aaatgtgttg tcagatattc ctccaattgt 120

tcacatagct gggatatttg ttgctcccct cacccccttgg attatgtagg gagccagtgc 180

acacagcctg tttgttttag tatccaagga agagaccaag gagccagctg gcgggaaggg 240

gtgggggtgt gcagtctgcc ctgtccttct gctcataacc tgacaaaatg ccaaactagt 300

aagcaggata gctgatacca cggctatgag ggagtaggct ctgagagggc acagacttgt 360

ggagctgggc gtctggatca aaactgcttt gggatggaac ctcgagccct agcagtgaag 420

aagactccat ttcttgtcca ggggatttaa aagagttttc tgctttgaga gagaaataga 480

gagtttagaa agcaattgct cttgggaaag ctatacacag ctctgttttg tcaatgacct 540

ttgttgtaag tctcccaacg tcccattagg agccacagca ggtgaggcat ttggtgcagc 600

aggaaacatg gggactgcct aggctcgaat ctgtggcacc ctgagcaatt acttaaattg 660

tggagcctag ttcctcatct gtaagatgga cttgagattc ctacctctca tgattactat 720

ggagattgaa taattggtaa aattctccta gctcagtgac tgccacagga tgggtctttc 780

agattttggt tctctttagc ttctggttct tgaaagaaat taatctgtat ataacataag 840

aaactttgaa agtcaaaaaa acaaaaaatt ttaattcctc gtagattaat tgatttgcta 900

tcttttagtt tttttttcta tgcatgtaga tgtattaaat atgtaaatgt ttttcaaagt 960

tgaggtaata ttgtatagaa ttttatagcc tacattttaa tttcttgata tcttaagcat 1020

ttcacttatt agatattgtt caaaaatgcc atttttaata tttgtataat accctatcat 1080

gtgcgtatac cttaactaag ccattcccat attcaacatt ttgtgt 1126

<210> 28

<211> 3215

<212> DNA

<213> Homo sapiens

<400> 28
gtgacttcct ttttctgccc actctggtaa cttattgctc tgctgggctc tttcccttag 60

ggtctctggc cctgttcttg ccccagcatg acttttatcg ggacgccgtt gtggaagcct 120

cacgcaggag ccctgccccc gtggagaaga tcccactggt gactccaacc ctaccaccat 180

gaatggggtc ctgatccccc atacgcccat cgcagtggac ttctggagcc tgcgccgggc 240

tggcaccgca cgtctcttct tcttgtctca catgcactcg gaccacaccg tgggcctgtc 300

tagcacctgg gcccggcccc tctactgctc cccaattaca gcccacctct tgcatcgtca 360

cctacaggtg attttcgata cacaccatcc atgctaaagg agccagccct gacactgggg 420

aaacagatcc atactttata cctagacaac accaattgca atccagccct ggttcttcct 480

tcccgacaag aagctgccca ccagattgtc cagctcattc gaaaacaccc acaacataac 540

ataaagattg gactctacag cctgggaaag gaatcactgc tggagcagct ggccctggag 600

tttcagacct gggtggtatt gagtcctcgg cgcctggagt tggtacagct actgggcctg 660

gcagatgtgt tcacagtgga ggagaaggct ggccgcatcc atgcagtaga ccatatggag 720

atctgccatt ccaacatgct gcgttggaac cagacccacc ctacgattgc tatccttccc 780

acaagccgaa aaatccacag ctcccaccct gatatccacg tcatccctta ctctgaccat 840

tcctcttact ccgagcttcg tgcctttgtc gcagcactga agccttgcca ggtggtgccc 900

attgtaagtc ggcggccctg tggaggcttt caggacagtc tgagccccag gatctccgtg 960

cccctgattc cggactctgt acagcaatac atgagttctt cctctagaaa accaagcctt 1020

ctctggctgt tagaaaggag gctaaagagg ccgagaaccc aaggtgttgt gtttgaatcc 1080

cctgaggaaa gtgctgatca atctcaagct gacagagact caaagaaggc caagaaagag 1140

aaactttctc cctggcctgc ggaccttgaa aagcagcctt cccaccatcc tttgcggatc 1200

aagaagcagt tgttcccaga tctctatagc aaagaatgga acaaggcagt gcctttctgc 1260

agggtattct tccaggagat ttgaccagca agtggaaaaa taccataaac cctgctgaag 1320

acaggagagt acagaatgac aacattgagc ccacactgca gttttgaaga tagtaactga 1380

tggctggtgg gaaagagttt gtttttgggg cctacttttc tatctttaca agactcttat 1440

gggcccaccg tggagcagca cttcccaaaa cttgttcact ggggtcctcg tgcctatgga 1500

atccttcttt ttataactaa gtttaagaaa tacttttttt ataaaatctt tggagtatgc 1560

gtgagcaaat taaaagttct ttgaagtcct acagtaactt aatctgttta accttgttta 1620

acccagtatt tctcaaactt ttgtgaacat gcaatcatct tatgtgggta cagaaagagg 1680

taaagagtct gaatcaaaaa ggaccaggtt attgctgttg ctgttttgtg gtgtcatgag 1740

ccattctcca tgtccccttc tccctcttct cagatcaaaa tccctaggga gttctatttt 1800

taaaattatg aactatggcg ctgcatgctt caatcctgaa cgtcactgac ttgctgtgac 1860

catccaaata attttcctgt ctctgcctct gggagggaac aggaagcgat gaagaggtct 1920

tggaacagta gtgaaaattc tacctctatg tccttcatga ggatgtgcag tatcccagta 1980

tcactgggat ccatgtggaa cagagccagc tgggggggttg ggcagctctc tccaaggcag 2040

tacctagagc ccagctgaac aacaaggctt tgggtgtgaa gggactcccc agcctggaga 2100

ccctatttgg ctgaaacagt tacaaaatat caaatgtgtt gtcagatatt cctccaattg 2160

ttcacatagc tgggatattt gttgctcccc tcacccccttg gattatgtag ggagccagtg 2220

cacacagcct gtttgtttta gtatccaagg aagagaccaa ggagccagct ggcgggaagg 2280

ggtgggggtg tgcagtctgc cctgtccttc tgctcataac ctgacaaaat gccaaactag 2340

taagcaggat agctgatacc acggctatga gggagtaggc tctgagaggg cacagacttg 2400

tggagctggg cgtctggatc aaaactgctt tgggatggaa cctcgagccc tagcagtgaa 2460

gaagattcca tttcttgtcc aggggattta aaagagtttt ctgctttgag agagaaatag 2520

agagtttaga aagcaattgc tcttgggaaa gctatacaca gctctgtttt gtcaatgacc 2580

tttgttgtaa gtctcccaac gtcctattag gagccacagc aggtgaggca tttggtgcag 2640

caggaaacat ggggactgcc taggctcgaa tctgtggcac cctgagcaat tacttaaatt 2700

gtggagccta gttcctcatc tgtaagatgg acttgagatt cctacctctc atgattacta 2760

tggagattga ataattggta aaattctcct agctcagtga ctgccacagg atgggtcttt 2820

cagattttgg ttctctttag cttctggttc ttgaaagaaa ttaatctgta tataacataa 2880

gaaactttga aagtcaaaaa aacaaaaaat tttaattcct cgtagattaa ttgatttgct 2940

atcttttagt ttttttttct atgcatgtag atgtattaaa tatgtaaatg tttttcaaag 3000

ttgaggtaat attgtataga attttatagc ctacatttta atttcttgat atcttaagca 3060

tttcacttat tagatattgt tcaaaaatgc catttttaat atttgtataa taccctatca 3120

tgtgagtata ccttaactaa gccattccca tattcaacat tttgtgtact gtttttctaa 3180

ttacatatat tacaatgaaa aaaaaaaaaa aaaaa 3215

<210> 29

<211> 3183

<212> DNA

<213> Homo sapiens

<400> 29
attttggaac catcctctac acaggtgatt ttcgatacac accatccatg ctaaaggagc 60

cagccctgac actggggaaa cagatccata ctttatacct agacaacacc aattgcaatc 120

cagccctggt tcttccttcc cgacaagaag ctgcccacca gattgtccag ctcattcgaa 180

aacacccaca acataacata aagattggac tctacagcct gggaaaggaa tcactgctgg 240

agcagctggc cctggagttt cagacctggg tggtattgag tcctcggcgc ctggagttgg 300

tacagctact gggcctggca gatgtgttca cagtggagga gaaggctggc cgcatccatg 360

cagtagacca tatggagatc tgccattcca acatgctgcg ttggaaccag acccacccta 420

cgattgctat ccttcccaca agccgaaaaa tccacagctc ccaccctgat atccacgtca 480

tcccttactc tgaccattcc tcttactccg agcttcgtgc ctttgtcgca gcactgaagc 540

cttgccaggt ggtgcccatt gtaagtcggc ggccctgtgg aggctttcag gacagtctga 600

gccccaggat ctccgtgccc ctgattccgg actctgtaca gcaatacatg agttcttcct 660

ctagaaaacc aagccttctc tggctgttag aaaggaggct aaagaggccg agaacccaag 720

gtgttgtgtt tgaatcccct gaggaaagtg ctgatcaatc tcaagctgac agagactcaa 780

agaaggccaa gaaagagaaa ctttctccct ggcctgcgga ccttgaaaag cagccttccc 840

accatccttt gcggatcaag aagcagttgt tcccagatct ctatagcaaa gaatggaaca 900

aggcagtgcc tttctgtgag tctcaaaaga gggtgactat gttgacggcc ccactgggat 960

tttcagtgca cttaaggtct acagatgagg agtttatttc tcaaaaaacc agggaggaaa 1020

ttggtttagg gtccccctg gtacccatgg gagatgatga tggaggtcca gaagccacag 1080

ggaatcagag tgcctggatg ggccatggtt ctcccctgtc ccacagcagc aagggcaccc 1140

ctcttctagc tactgaattc aggggtctag cactcaaata tcttctgact ccagtgaact 1200

ttttccaggc agggtattct tccaggagat ttgaccagca agtggaaaaa taccataaac 1260

cctgctgaag acaggagagt acagaatgac aacattgagc ccacactgca gttttgaaga 1320

tagtaactga tggctggtgg gaaagagttt gtttttgggg cctacttttc tatctttaca 1380

agactcttat gggcccaccg tggagcagca cttcccaaaa cttgttcact ggggtcctcg 1440

tgcctatgga atccttcttt ttataactaa gtttaagaaa tacttttttt ataaaatctt 1500

tggagtatgc gtgagcaaat taaaagttct ttgaagtcct acagtaactt aatctgttta 1560

accttgttta acccagtatt tctcaaactt ttgtgaacat gcaatcatct tatgtgggta 1620

cagaaagagg taaagagtct gaatcaaaaa ggaccaggtt attgctgttg ctgttttgtg 1680

gtgtcatgag ccattctcca tgtccccttc tccctcttct cagatcaaaa tccctaggga 1740

gttctatttt taaaattatg aactatggcg ctgcatgctt caatcctgaa cgtcactgac 1800

ttgctgtgac catccaaata attttcctgt ctctgcctct gggagggaac aggaagcgat 1860

gaagaggtct tggaacagta gtgaaaattc tacctctatg tccttcatga ggatgtgcag 1920

tatcccagta tcactgggat ccatgtggaa cagagccagc tgggggggttg ggcagctctc 1980

tccaaggcag tacctagagc ccagctgaac aacaaggctt tgggtgtgaa gggactcccc 2040

agcctggaga ccctatttgg ctgaaacagt tacaaaatat caaatgtgtt gtcagatatt 2100

cctccaattg ttcacatagc tgggatattt gttgctcccc tcacccctg gattatgtag 2160

ggagccagtg cacacagcct gtttgtttta gtatccaagg aagagaccaa ggagccagct 2220

ggcgggaagg ggtgggggtg tgcagtctgc cctgtacttc tgctcataac ctgacaaaat 2280

gccaaactag taagcaggat agctgatacc acggctatga gggagtaggc tctgagaggg 2340

cacagacttg tggagctggg cgtctggatc aaaactgctt tgggatggaa cctcgagccc 2400

tagcagtgaa gaagattcca tttcttgtcc aggggattta aaagagtttt ctgctttgag 2460

agagaaatag agagtttaga aagcaattgc tcttgggaaa gctatacaca gctctgtttt 2520

gtcaatgacc tttgttgtaa gtctcccaac gtcctattag gagccacagc aggtgaggca 2580

tttggtgcag caggaaacat ggggactgcc taggctcgaa tctgtggcac cctgagcaat 2640

tacttaaatt gtggagccta gttcctcatc tgtaagatgg acttgagatt cctacctctc 2700

atgattacta tggagattga ataattggta aaattctcct agctcagtga ctgccacagg 2760

atgggtcttt cagattttgg ttctctttag cttctggttc ttgaaagaaa ttaatctgta 2820

tataacataa gaaactttga aagtcaaaaa aacaaaaaat tttaattcct cgtagattaa 2880

ttgatttgct atcttttagt tttttttct atgcatgtag atgtattaaa tatgtaaatg 2940

tttttcaaag ttgaggtaat attgtataga attttatagc ctacatttta atttcttgat 3000

atcttaagca tttcacttat tagatattgt tcaaaaatgc catttttaat atttgtataa 3060

taccctatca tgtgagtata ccttaactaa gccattccca tattcaacat tttgtgtact 3120

gtttttctaa ttacatatat tacaatgaac aaccttatgc aaaaaaaaaa aaaaaaaaaa 3180

aaa 3183

<210> 30

<211> 1157

<212> DNA

<213> Homo sapiens

<400> 30

ggcggttggg agtgtccagc gccctccgcg atttgggctc cagcgggcag ggtgacttcc 60

tttttctgcc cactctggta acttattgct ctgctgggct ctttccctta gggtctctgg 120

ccctgttctt gccccagcat gacttttatc gggacgccgt tgtggaagcc tcacgcagga 180

gccctgcccc cgtggagaag atcccactgg tgactccaac cctaccacca tgaatggggt 240

cctgatcccc catacgccca tcgcagtgga cttctggagc ctgcgccggg ctggcaccgc 300

acgtctcttc ttcttgtctc acatgcactc ggaccacacc gtgggcctgt ctagcacctg 360

ggcccggccc ctctactgct ccccaattac agcccacctc ttgcatcgtc acctacaggt 420

atctaagcaa tggatccaag ccctggaggt tggtgagagc catgtattac ccctagatga 480

aattggacaa gagaccatga ccgtaaccct cctcgatgcc aatcactgtc ctggttctgt 540

catgtttctc tttgaaggat attttggaac catcctctac acaggtgatt ttcgatacac 600

accatccatg ctaaaggagc cagccctgac actggggaaa cagatccata ctttatacct 660

agacaacacc aattgcaatc cagccctggt tcttccttcc cgacaagaag ctgcccacca 720

gattgtccag ctcattcgaa aacacccaca acataacata aagattggac tctacagcct 780

gggaaaggaa tcactgctgg agcagctggc cctggagttt cagacctggg tggtattgag 840

tcctcggcgc ctggagttgg tacagctact gggcctggca gatgtgttca cagtggagga 900

gaaggctggc cgcatccatg cagtagacca tatggagatc tgccattcca acatgctgcg 960

ttggaaccag acccacccta cgattgctat ccttcccaca agccgaaaaa tccacagctc 1020

ccaccctgat atccacgtca tcccttactc tgaccattcc tcttactccg agcttcgtgc 1080

ctttgtcgca gcactgaagc cttgccaggt ggtgcccatt gtaagtcggc ggccctggga 1140

ggctttcagg acagtct 1157

<210> 31

<211> 1014

<212> DNA

<213> Homo sapiens

<400> 31

ggaagatggc gacggccttg agcgaggagg agctggacaa tgaagactat tactcgttgc 60

tgaacgtgcg cagggaggcc tcttctgaag agctgaaagc tgcctaccgg aggctctgta 120

tgctctacca tccagacaag cacagagacc cagagctcaa gtcacaggcg gaacgactgt 180

ttaaccttgt tcaccaggct tatgaagtgc ttagtgaccc ccaaaccagg gccatctatg 240

atatatatgg gaagggagga ctggaaatgg aaggatggga ggttgtggaa aggaggagaa 300

cccctgctga aattcgagag gagtttgagc ggctgcagag agagagagaa gagaggagat 360

tgcagcagcg aaccaatccc aagggaacga tcagcgttgg agtaaatgcc accgaccttt 420

ttgatcgcta tgatgaggag tatgaagatg tgtccggcag tagctttccg cagattgaaa 480

ttaataaaat gcacatatcc cagtccattg aggcaccctt gacagcgaca gacacagcca 540

tcctctctgg aagcctctca acccagaatg gaaatggagg aggttccatt aactttgcgc 600

tcagacgagt aacttcggta aagggatggg gagagttgga atttggagct ggagacctac 660

aggggccttt gttcggtctc aagctgttcc gtaatctcac accaagatgc tttgtgacaa 720

caaactgtgc tctgcagttt tcatcccgtg gaatccgacc cggcctgacc actgtcctag 780

ctcggaacct agacaagaac accgtgggct acctgcagtg gcgatggggt atccagtcag 840

ccatgaacac tagcatcgtc cgagacacta aaaccagcca cttcactgtg gccctgcagc 900

tgggaatccc tcactccttt gcactgatca gctatcagca caaattccaa gatgacgatc 960

agactcgtgt gaagggatcc ctcaaagcag gcttctttgg gacggtggtg gagt 1014

<210> 32

<211> 3230

<212> DNA

<213> Homo sapiens

<400> 32

gggggtgaaa ggttgcgaag atggcgacgg ccttgagcga ggaggagctg gacaatgaag 60

actattactc gttgctgaac gtgcgcaggg aggcctcttc tgaagagctg aaagctgcct 120

accggaggct ctgtatgctc taccatccag acaagcacag agacccagag ctcaagtcac 180

aggcggaacg actgtttaac cttgttcacc aggcttatga agtgcttagt gacccccaaa 240

ccagggccat ctatgatata tatgggaaga gaggactgga aatggaagga tgggaggttg 300

tggaaaggag gagaacccct gctgaaattc gagaggagtt tgagcggctg cagagagaga 360

gagaagagag gagattgcag cagcgaacca atcccaaggg aacgatcagc gttggagtag 420

atgccaccga ccttttttgat cgctatgatg aggagtatga agatgtgtcc ggcagtagct 480

ttccgcagat tgaaattaat aaaatgcaca tatcccagtc cattgaggca cccttgacag 540

cgacagacac agccatcctc tctggaagcc tctcaaccca gaatggaaat ggaggaggtt 600

ccattaactt tgcgctcaga cgagtaactt cggcaaaggg atggggagag ttggaatttg 660

gagctggaga cctacagggg cctttgttcg gtctcaagct gttccgtaat ctcacaccaa 720

gatgctttgt gacaacaaac tgtgctctgc agttttcatc ccgtggaatc cgacccggcc 780

tgaccactgt cctagctcgg aacctagaca agaacaccgt gggctacctg cagtggcgat 840

ggggtatcca gtcagccatg aacactagca tcgtccgaga cactaaaacc agccacttca 900

ctgtggccct gcagctggga atccctcact cctttgcact gatcagctat cagcacaaat 960

tccaagatga cgatcagact cgtgtgaaag gatccctcaa agcaggcttc tttgggacgg 1020

tggtggagta cggagctgag aggaagatct ccaggcacag cgttttgggt gcagctgtca 1080

gcgttggagt tccacagggc gtttctctca aagtcaagct caacagggcc agtcagacat 1140

acttcttccc tattcacttg acggaccagc ttctgcccag cgccatgttc tatgccaccg 1200

tggggcctct agtggtctac tttgccatgc accgtctgat catcaaacca tacctcaggg 1260

ctcagaaaga gaaggaattg gagaagcaga gggaaagcgc cgccaccgat gtgctgcaga 1320

agaagcaaga ggcggagtcc gctgtccggc tgatgcagga atctgtccga aggataattg 1380

aggcagaaga gtccagaatg ggcctcatca tcgtcaatgc ctggtacggg aagtttgtca 1440

atgacaagag caggaagagc gagaaggtga aggtgattga cgtgactgtg cccctgcagt 1500

gcctggtgaa ggactcgaag ctcatcctca cggaggcctc caaggctggg ctgcctggct 1560

tttatgaccc gtgtgtgggg gaagagaaga acctgaaagt gctctatcag ttccggggcg 1620

tcctgcatca ggtgatggtg ctggacagtg aggccctccg gataccaaag cagtcccaca 1680

ggatcgatac agatggataa actgccaaga accagatttt taaaaggccg caaaaaatct 1740

tttcctggga gtctacaaat ttggaaatga aaaaacccag acatcagatg tttttatttt 1800

atattattat tatagaaggt ggtaccatta tcaattatgt gaagggacat gcagacaccc 1860

cagcttttga gggtgctggg ggtaggactg aggcagcccc actgggaacc agactgcagc 1920

ctggcccatg gctgttttcc caaggatcag ttcctggagg gaagggctct ggccctgact 1980

ccgctgtgtc ccgagcacac gtgctgaccg cagcccgccg ccctgtagtt cttggctggg 2040

tctggaggtg tctgtggagc accctgccct caccacagga gcgtgagcca cttctgcagt 2100

ccacgctgaa catgggaaac aacctgaaaa gcaggcaggc ctcccggtca gggagcctct 2160

gctgtgctgg cttcccatga ccacctcctc ctgctgaaat attactgctt gaatctggag 2220

cagattgcgg gtttataaaa ctgcttttta tctgagaaca aacgggtttg gaaattagtc 2280

gtctttttc cccactccca gagctgctca aatcattcca ccggcccccc cggcttggga 2340

cagggtagtg taactcccga tcccagggcc tagccctgac acaggtggct tcccgtatcc 2400

cggtgggaaa acgccctgcc accagcgggc ttgagctggc ctgtgtccct ccactgcctg 2460

caccacccac ctccagagtg cagtgctggg caagggcagc tcaagaggac aggaccaggc 2520

gcttggcaag acatcagaca cacccaaccc aaaggcgtgg accccaggcc cggcccgtgg 2580

tacccagcag gtggcactgc agctccccgc tcctgcaggt ccagcgtcct cacaggaaca 2640

ccagggcctg tgctccggag ccttccttca gacccttcct ccacgtgccc acttgggatg 2700

cagaatgcag cggagctagg acccccctcca cggcctggac ctcggctgca gtaaagttac 2760

gtgaggcctg tctctcgggg cctggaagtg gcagccatca gttgctcttg ctgacccctc 2820

ggagcaagcg ccgcacaggt ggtggctgag acagctggcg tgggggggccc caagctgcgc 2880

cggcctccag cccacccaca gctgttgctg aagtcaggcc tccctcccca gcactggtat 2940

ctgagtaacg gctaagaacc tccttcctct ggttttgaaa agcagttcgg gttgtccaat 3000

tctgtaacat tcatctccat tttttaaaaa ggtttctctg acggccccac ggcccgagcc 3060

acggtgagcg tcgtgttgca tgagcctggg ccccgggctt cccgtgcgcc tctgccgcag 3120

gtgcttctgg gcacccatcc tctgcgtttc atttgcagtc gactgtacag aaggcactca 3180

ccacaataaa cctttcctga aagcagaaaa aaaaaaaaaa aaaaaaaaaa 3230

<210> 33

<211> 355

<212> DNA

<213> Homo sapiens

<400> 33
gagctgagat aatgcacggc acttcagctt gggcgacaga gcgagactcc gtctcaaaaa 60

agaaaaaaaa gaaagaaatt cataaggttg tggaaaggag gagaacccct gctgaaattc 120

gagaggagtt tgagcggctg cagagagaga gagaagagag gagattgcag cagcgaacca 180

atcccaaggg aacgatcagc gttggagtag atgccaccga ccttttttgat cgctatgatg 240

aggagtatga agatgtgtcc ggcagtagct ttccgcagat tgaaattaat aaaatgcaca 300

tatcccagtc cattgaggca cccttgacag cgacagacac agccatcctc tctgg 355

<210> 34

<211> 454

<212> DNA

<213> Homo sapiens

<400> 34
acaatatgaa gccttcattt aatctctgca gttcatctca tttcaaatgt ttatggaaga 60

agcacttcat tgaaagtagt gctgtaaata ttctgccata ggaatactgt ctacatgctt 120

tctcattcaa gaattcgtca tcacgcatca caggccgcgt ctttgacggt gggtgtccca 180

tttttatccg ctactcttta tttcatggag tcgtatcaac gctatgaacg caaggctgtg 240

atatggaacc agaaggctgt ctgaactttt gaaaccttgt gtgggattga tggtggtgcc 300

gaggcatgaa aggctagtat gagcgagaaa aggagagagc gcgtgcagag acttggtggt 360

gcataatgga tatttttaa cttggcgaga tgtgtctctc aatcctgtgg ctttggtgag 420

agagtgtgca gagagcaatg atagcaaata atgt 454

<210> 35

<211> 3661

<212> DNA

<213> Homo sapiens

<400> 35
cgtcgctgct tcggtgtccc tgtcgggctt cccagcagcg gcctagcggg aaaagtaaaa 60

gatgtctgaa tatattcggg taaccgaaga tgagaacgat gagcccattg aaataccatc 120

ggaagacgat gggacggtgc tgctctccac ggttacagcc cagtttccag gggcgtgtgg 180

gcttcgctac aggaatccag tgtctcagtg tatgagaggt gtccggctgg tagaaggaat 240

tctgcatgcc ccagatgctg gctggggaaa tctggtgtat gttgtcaact atccaaaaga 300

taacaaaaga aaaatggatg agacagatgc ttcatcagca gtgaaagtga aaagagcagt 360

ccagaaaaca tccgatttaa tagtgttggg tctcccatgg aaaacaaccg aacaggacct 420

gaaagagtat tttagtacct ttggagaagt tcttatggtg caggtcaaga aagatcttaa 480

gactggtcat tcaaaggggt ttggctttgt tcgttttacg gaatatgaaa cacaagtgaa 540

agtaatgtca cagcgacata tgatagatgg acgatggtgt gactgcaaac ttcctaattc 600

taagcaaagc caagatgagc ctttgagaag cagaaaagtg tttgtggggc gctgtacaga 660

ggacatgact gaggatgagc tgcgggagtt cttctctcag tacggggatg tgatggatgt 720

cttcatcccc aagccattca gggcctttgc ctttgttaca tttgcagatg atcagattgc 780

gcagtctctt tgtggagagg acttgatcat taaaggaatc agcgttcata tatccaatgc 840

cgaacctaag cacaatagca atagacagtt agaaagaagt ggaagatttg gtggtaatcc 900

aggtggcttt gggaatcagg gtggatttgg taatagcaga gggggtggag ctggtttggg 960

aaacaatcaa ggtagtaata tgggtggtgg gatgaacttt ggtgcgttca gcattaatcc 1020

agccatgatg gctgccgccc aggcagcact acagagcagt tggggtatga tgggcatgtt 1080

agccagccag cagaaccagt caggcccatc gggtaataac caaaaccaag gcaacatgca 1140

gagggagcca aaccaggcct tcggttctgg aaataactct tatagtggct ctaattctgg 1200

tgcagcaatt ggttggggat cagcatccaa tgcagggtcg ggcagtggtt ttaatggagg 1260

ctttggctca agcatggatt ctaagtcttc tggctgggga atgtagacag tggggttgtg 1320

gttggttggt atagaatggt gggaattcaa atttttctaa actcatggta agtatattgt 1380

aaaatacata tgtactaaga attttcaaaa ttggtttgtt cagtgtggag tatattcagc 1440

agtatttttg acatttttct ttagaaaaag gaagagctaa aggaatttta taagttttgt 1500

tacatgaaag gttgaaatat tgagtggttg aaagtgaact gctgtttgcc tgattggtaa 1560

accaacacac tacaattgat atcaaaaggt ttctcctgta atattttatc cctggacttg 1620

tcaagtgaat tctttgcatg ttcaaaacgg aaaccattga ttagaactac attctttacc 1680

ccttgtttta atttgaaccc caccatatgg atttttttcc ttaagaaaat ctccttttag 1740

gagatcatgg tgtcacagtg tttggttctt ttgttttgtt ttttaacact tgtctccccc 1800

catacacaaa agtacaatat gaagccttca tttaatctct gcagttcatc tcatttcaaa 1860

tgtttatgga agaagcactt cattgaaagt agtgctgtaa atattctgcc ataggaatac 1920
tgtctacatg ctttctcatt caagaattcg tcatcacgca tcacaggccg cgtctttgac 1980
ggtgggtgtc ccatttttat ccgctactct ttatttcatg gagtcgtatc aacgctatga 2040
acgcaaggct gtgatatgga accagaaggc tgtctgaact tttgaaacct tgtgtgggat 2100
tgatggtggt gccgaggcat gaaaggctag tatgagcgag aaaaggagag agcgcgtgca 2160
gagacttggt ggtgcataat ggatattttt taacttggcg agatgtgtct ctcaatcctg 2220
tggctttggt gagagagtgt gcagagagca atgatagcaa ataatgtacg aatgtttttt 2280
gcattcaaag gacatccaca tctgttggaa gacttttaag tgagtttttg ttcttagata 2340
acccacatta gatgaatgtg ttaagtgaaa tgatacttgt actcccccta ccccttttgtc 2400
aactgctgtg aatgctgtat ggtgtgtgtt ctcttctgtt actgatatgt aagtgtggca 2460
atgtgaactg aagctgatgg gctgagaaca tggactgagc ttgtggtgtg ctttgcagga 2520
ggacttgaag cagagttcac cagtgagctc aggtgtctca aagaagggtg gaagttctaa 2580
tgtctgttag ctacccataa gaatgctgtt tgctgcagtt ctgtgtcctg tgcttggatg 2640
ctttttataa gagttgtcat tgttggaaat tcttaaataa aactgattta aataatatgt 2700
gtctttgttt tgcagccctg aatgcaaaga attcatagca gttaattccc ctttttgac 2760
ccttttgaga tggaactttc ataaagtttc ttggcagtag tttattttgc ttcaaataaa 2820
cttatttgaa aagttgtctc aagtcaaatg gattcatcac ctgtcatgca ttgacacctg 2880
atacccagac ttaattggta tttgttcttg cattggccaa agtgaaaatt ttttttttt 2940
cttttgaaat ctagttttga ataagtctgg gtgaccgcac ctaaaatggt aagcagtacc 3000
ctccggcttt ttcttagtgc ctctgtgcat ttgggtgatg ttctatttac atggcctgtg 3060
taaatctcca ttgggaagtc atgccttcta aaaagattct tatttggggg agtgggcaaa 3120
atgttgatta ttttctaatg ctttgtagca aagcatatca attgaaaagg gaatatcagc 3180
accttcctag tttgggattt gaaaagtgga attaattgca gtagggataa agtagaagaa 3240
accacaaatt atcttgtgcc tgaaatccat taagaggcct gatagcttta agaattaggg 3300
tgggttgtct gtctggaagt gttaagtgga atgggctttg tcctccagga ggtgggggaa 3360
tgtggtaaca ttgaatacag ttgaataaaa tcgcttacaa aactcacact ctcacaatgc 3420
attgttaagt atgtaaaagc aataacattg attctctgtt gtacttttt gtaactaatt 3480
ctgtgagagt tgagctcatt ttctagttgg aagaatgtga tatttgttgt gttggtagtt 3540
tacctaatgc ccttacctaa ttagattatg ataaataggt ttgtcatttt gcaagttaca 3600
tgttttaaac atttatcaat gaagtcatcc tttagacttg taaaaaaaaa aaaaaaagaa 3660
a 3661

<210> 36

<211> 1095

<212> DNA

<213> Homo sapiens

<400> 36
tcgagcggcc gcccacgcgg gcacaccaag aatcagctga acctaaatac cttcctcata 60

aaacatgtaa cgaaattatt gtgcctaaag ccccctctca taaaacaatc caagaaacac 120

ctcattctga agactattca attgaaataa accaagaaac tcctgggtct gaaaaatatt 180

cacctgaaac gtatcaagaa atacctgggc ttgaagaata ttcacctgaa atataccaag 240

aaacatccca gcttgaagaa tattcacctg aaatatacca agaaacaccg gggcctgaag 300

acctctctac tgagacatat aaaaataagg atgtgcctaa agaatgcttt ccagaaccac 360

accaagaaac aggtgggccc caaggccagg atcctaaagc acaccaggaa gatgctaaag 420

atgcttatac ttttcctcaa gaaatgaaag aaaaacccaa agaagagcca ggaataccag 480

caattctgaa tgagagtcat ccagaaaatg atgtctatag ttatgttttg ttttaacaat 540

gctcaaccat aaagttgtgg tccaatggaa catacagctt aatagtttat gcgtgatttt 600

ctcaaaatat tgtaaaactt ttgacaatgc tcattaatat tatttttct atttgtagac 660

catatctgaa agaaataaca ttttttaagg ctctaccaca tagacaatat catgctagaa 720

tgtgtgtgtg tgtgtgtgtg tgtatgtatg tataggtcgg ggagaggata gtggtgggaa 780

cagacaaata aggaagcggg gaggactgga taattggttt tccccccctaa gaacatttat 840

ttacgtctta agagcagata agtgactaag actgaacaca tacattttgt ggagtatata 900

gttttcttgt aaatgctgtt caattattaa tgtaacagta gcatcaaaat tttattcagg 960

ctttagttga ctcttttggt cagttttaac aattctcctt aaaagatatt ttggagtgat 1020

gaatgtggtt tacttttgta tttgaatttt gattttctat ttttatttt taaatattgt 1080

atttgtgcac aatgt 1095

<210> 37

<211> 2103

<212> DNA

<213> Homo sapiens

<400> 37
ggaagtcaga ccaaaatagc aggaaggtat tgcagcaaga tggatttggg aaaggaccaa 60

tctcatttga agcaccatca gacacctgac cctcatcaag aagagaacca ttctccagaa 120
gtcattggaa cctggagttt gagaaacaga gaactactta gaaaaagaaa agctgaagtg 180
catgaaaagg aaacatcaca atggctattt ggagaacaga aaaaacgcaa gcagcagaga 240
acaggaaaag gaaatcgaag aggcagaaag agacaacaaa acacagaatt gaaggtggag 300
cctcagccac agatagaaaa ggaaatagtg gagaaagcac tggcacctat agagaaaaaa 360
actgagccac ctgggagcat aaccaaagta tttccttcag tagcctcccc gcaaaaagtt 420
gtgcctgagg aacacttttc tgaaatatgt caagaaagta acatatatca ggagaatttt 480
tctgagtacc aagaaatagc agtacaaaac cattcttctg aaacatgcca acatgtgtct 540
gaacctgaag acctctctcc taaaatgtac caagaaatat ctgtacttca agacaattct 600
tccaaaatat gccaagacat gaaggaacct gaagacaact ctcctaacac atgccaagta 660
atatctgtaa ttcaagacca tcctttcaaa atgtaccaag atatggctaa acgagaagat 720
ctggctccta aaatgtgcca agaagctgct gtacccaaaa tccttccttg tccaacatct 780
gaagacacag ctgatctggc aggatgctct cttcaagcat atccaaaacc agatgtgcct 840
aaaggctata ttcttgacac agaccaaaat ccagcagaac cagaggaata caatgaaaca 900
gatcaaggaa tagctgagac agaaggcctt tttcctaaaa tacaagaaat agctgagcct 960
aaagaccttt ctacaaaaac acaccaagaa tcagctgaac ctaaatacct tcctcataaa 1020
acatgtaacg aaattattgt gcctaaagcc ccctctcata aaacaatcca agaaacacct 1080
cattctgaag actattcaat tgaaataaac caagaaactc ctgggtctga aaaatattca 1140
cctgaaacgt atcaagaaat acctgggctt gaagaatatt cacctgaaat ataccaagaa 1200
acatcccagc ttgaagaata ttcacctgaa atataccaag aaacaccggg gcctgaagac 1260
ctctctactg agacatataa aaataaggat gtgcctaaag aatgctttcc agaaccacac 1320
caagaaacag gtgggcccca aggccaggat cctaaagcac accaggaaga tgctaaagat 1380
gcttatactt ttcctcaaga aatgaaagaa aaacccaaag aagagccagg aataccagca 1440
attctgaatg agagtcatcc agaaaatgat gtctatagtt atgttttgtt ttaacaatgc 1500
tcaaccataa agttgtggtc caatggaaca tacagcttaa tagtttatgc gtgattttct 1560
caaaatattg taaaactttt gacaatgctc attaatatta ttttttctat ttgtagacca 1620
tatctgaaag aaataacatt ttttaaggct ctaccacata gacaatatca tgctagaatg 1680
tgtgtgtgtg tgtgtgtgtg tgtgtgtgta tgtatgtata ggtcggggag aggatagtgg 1740
tgggaacaga caaataagga agcggggagg actggataat tggttttccc ccctaagaac 1800
atttatttac gtcttaagag cagataagtg actaagactg aacacataca ttttgtggag 1860
tatatagttt tcttgtaaat gctgttcaat tattaatgta acagtagcat caaaatttta 1920
ttcaggcttt agttgactct tttggtcagt tttaacaatt ctccttaaaa gatattttgg 1980

agtgatgaat gtagtttact tttgtatttg aattttgatt ttctattttt atttttaaa 2040

tattgtattt gtgcacaatg tacattaaat cattattaca tgcttaaaaa aaaaaaaaaa 2100

aaa 2103

<210> 38

<211> 1516

<212> DNA

<213> Homo sapiens

<400> 38

gcaagatgga tttgggaaag gaccaatctc atttgaagca ccatcagaca cctgaccctc 60

atcaagaaga gaaccattct ccagaagtca ttggaacctg gagtttgaga aacagagaac 120

tacttagaaa aagaaaagct gaagtgcatg aaaaggaaac atcacaatgg ctatttggag 180

aacagaaaaa acgcaagcag cagagaacag gaaaaggaaa tcgaagaggc agaaagagac 240

aacaaaacac agaattgaag gtggagcctc agccacagat agaaaaggaa atagtggaga 300

aagcactggc acctatagag aaaaaaactg agccacctgg gagcataacc aaagtatttc 360

cttcagtagc ctccccgcaa aaagttgtgc ctgaggaaca cttttctgaa atatgtcaag 420

aaagtaacat atatcaggag aatttttctg agtaccaaga aatagcagta caaaaccatt 480

cttctgaaac atgccaacat gtgtctgaac ctgaagacct ctctcctaaa atgtaccaag 540

aaatatctgt acttcaagac aattcttcca aaatatgcca agacatgaag gaacctgaag 600

acaactctcc taacacatgc caagtaatat ctgtaattca agaccatcct ttcaaaatgt 660

accaagatat ggctaaacga gaagatctgg ctcctaaaat gtgccaagaa gctgctgtac 720

ccaaaatcct tccttgtcca acatctgaag acacagctga tctggcagga tgctctcttc 780

aagcatatcc aaaaccagat gtgcctaaag gctatattct tgacacagac caaaatccag 840

cagaaccaga ggaatacaat gaaacagatc aaggaatagc tgagacagaa ggccttttc 900

ctaaaataca agaaatagct gagcctaaag acctttctac aaaaacacac caagaatcag 960

ctgaacctaa ataccttcct cataaaacat gtaacgaaat tattgtgcct aaagcccccc 1020

ctcataaaac aatccaagaa acacctcatt ctgaagacta ttcaattgaa ataaaccaag 1080

aaactcctgg gtctgaaaaa tattcacctg aaacgtatca agaaatacct gggcttgaag 1140

aatattcacc tgaaatatac caagaaacat cccagcttga agaatattca cctgaaatat 1200

accaagaaac accggggcct gaagacctct ctactgagac atataaaaat aaggatgtgc 1260

ctaaagaatg ctttccagaa ccacaccaag aaacaggtgg gccccaaggc caggatccta 1320

aagcacacca ggaagatgct aaagatgctt atacttttcc tcaagaaatg aaagaaaaac 1380

ccaaagaaga gccaggaata ccagcaattc tgaatgagag tcatccagaa aatgatgtct 1440

atagttatgt tttgttttaa caatgctcaa ccataaagtt gtggtccaat ggaacataaa 1500

aaaaaaaaaa aaaaaa 1516

<210> 39

<211> 426

<212> DNA

<213> Homo sapiens

<400> 39

acaactggca ttcaaatcta ggtcagtctg cccccagagc cactaccctt acccctcact 60

gaatctgcct ttatattgtt gagcccatga ccccaaactg ctctttccaa tttgaacttc 120

cagggatttt attgtgaact tacatagcaa cattaaaatg aagttgaatt gtttttaatg 180

gcaacgccgt ctgtctcctc tagcttaccg cttctcacct ttcaacccca tctgtggcct 240

ttgtccaggc ccacagctta gccatggctt ccctcctgca tccctgccgt gggttgctgg 300

cctcacactt gcagcagctg gacagtgatt ttagaaggcc accagtcccc atagctatgt 360

gacaatgaga agcaaacttt tttgtgacag attgtattgg cataggcatg atagatgggg 420

attggt 426

<210> 40

<211> 2864

<212> DNA

<213> Homo sapiens

<400> 40

ctcgcttctc gttctactgc cccaggagcc cggcgggtcc gggactcccg tccgtgccgg 60

tgcgggcgcc ggcatgtggc tgtgggagga ccagggcggc ctcctgggcc ctttctcctt 120

cctgctgcta gtgctgctgc tggtgacgcg gagcccggtc aatgcctgcc tcctcaccgg 180

cagcctcttc gttctactgc gcgtcttcag ctttgagccg gtgccctctt gcagggccct 240

gcaggtgctc aagccccggg accgcatttc tgccatcgcc caccgtggcg gcagccacga 300

cgcgcccgag aacacgctgg cggccattcg gcaggcagct aagaatggag caacaggcgt 360

ggagttggac attgagttta cttctgacgg gattcctgtc ttaatgcacg ataacacagt 420

agataggacg actgatggga ccgggcgatt gtgtgatttg acatttgaac aaattaggaa 480

gctgaatcct gcagcaaacc acagactcag gaatgatttc cctgatgaaa agatccctac 540

cctaagggaa gctgttgcag agtgcctaaa ccataacctc acaatcttct ttgatgtcaa 600
aggccatgca cacaaggcta ctgaggctct aaagaaaatg tatatggaat ttcctcaact 660
gtataataat agtgtggtct gttctttctt gccagaagtt atctacaaga tgagacaaac 720
agatcgggat gtaataacag cattaactca cagaccttgg agcctaagcc atacaggaga 780
tgggaaacca cgctatgata ctttctggaa acattttata tttgttatga tggacatttt 840
gctcgattgg agcatgcata atatcttgtg gtacctgtgt ggaatttcag ctttcctcat 900
gcaaaaggat tttgtatccc cggcctactt gaagaagtgg tcagctaaag gaatccaggt 960
tgttggttgg actgttaata cctttgatga aaagagttac tacgaatccc atcttggttc 1020
cagctatatc actgacagca tggtagaaga ctgcgaacct cacttctaga ctttcacggt 1080
gggacgaaac gggttcagaa actgccaggg gcctcataca gggatatcaa aatacccttt 1140
gtgctagccc aagccctggg gaatcaggtg actcacacaa atgcaatagt tggtcactgc 1200
atttttacct gaaccaaagc taaacccggt gttgccacca tgcaccatgg catgccagag 1260
ttcaacactg ttgctcttga aaatctgggt ctgaaaaaac gcacaagagc ccctgccctg 1320
ccctagctga ggcacacagg gagacccagt gaggataagc acagattgaa ttgtacaatt 1380
tgcagatgca gatgtaaatg catgggacat gcatgataac tcagagttga cattttaaaa 1440
cttgccacac ttatttcaaa tatttgtact cagctatgtt aacatgtact gtagacatca 1500
aacttgtggt catactaata aaattattaa aaggagcact aaaggaaaac tgtgtgccaa 1560
gcatcatatc ctaaggcata cggaatttgg ggaagccacc atgcaatcca gtgaggcttc 1620
agtgtacagc aaccaaaatg gtagggaggt cttgaagcca atgagggatt tatagcatct 1680
tgaatagaga gctgcaaacc accaggggggc agagttgcac ttttccaggc tttttaggaa 1740
gctctgcaac agatgtgatc tgatcatagg caattagaac tggaagaaac ttccaaaaat 1800
atctaggttt gtcctcattt tacaaatgag gaaactaaac tctgtggaag ggaaggggtt 1860
gcctcaaaag tcacagctta gctgggcaca gtggctcatg ccgataatcc cagcaattca 1920
gaaagctgag gcaggaggat tacttgaggc cagactgggc aatatagcaa gaccccatct 1980
ctaaaaaatt aggcatggtg gtgcatgcct gtattcccag ctactcagga ggttgaggtg 2040
ggaggatcac ttgagcccag aagttcaagg ttgcaatgag ccatgattac accacggcac 2100
tacaaccttg gtggcacagt gagaacctga ctcttaaaaa aaaaaaaaaa aaaaaaaaag 2160
ataactagaa cttctagaac atcttgttta cagttagcca gaaactatac aagtggttta 2220
acatgcatta tcttactcaa tccatacaaa agtcttatgg aggtgttagc actctttcta 2280
ctgatgaaga actgaggtac ttcataaaac cacttaccca aggtgtcttg agtctggtac 2340
aactggcatt caaatctagg tcagtctgcc cccagagcca ctacccttac ccctcactga 2400
atctgccttt atattgttga gcccatgacc ccaaactgct ctttccaatt tgaacttcca 2460

gggattttat tgtgaactta catagcaaca ttaaaatgaa gttgaattgt ttttaatggc 2520

aacgccgtct gtctcctcta gcttaccgct tctcaccttt caaccccatc tgtggccttt 2580

gtccaggccc acagcttagc catggcttcc ctcctgcatc cctgccgtgg gttgctggcc 2640

tcacacttgc agcagctgga cagtgatttt agaaggccac cagtccccat agctatgtga 2700

caatgagaag caaacttttt tgtgacagat tgtattggca taggcatgat agatggggat 2760

tggtacgttt tgaatcagca tttgcaaaaa aattgtcttg aattttaaaa taaacaacaa 2820

agatttgttc attgagtgca aaaaaaaaaa aaaaaaaaaa aaaa 2864

<210> 41

<211> 704

<212> DNA

<213> Homo sapiens

<400> 41
acttcacgac tggctgccac tactgggagg tgtatgtggg agacaagacc aaatggattc 60

ttggagtatg tagtgagtca gtgagcagga aggggaaggt tactgcctca cctgccaatg 120

gacactggct tctgcgacag agtcgtggga atgagtatga agctctcaca tccccgcaga 180

cctccttccg ccttaaagag cctccacggt gtgtggggat tttcctggac tatgaagcag 240

gagtcatctc tttctacaat gtgaccaaca agtcccacat ctttactttc acccacaatt 300

tctctggccc ccttcgccct ttctttgaac cttgccttca tgatggagga aaaaacacag 360

cacctctagt catttgttca gaactacaca aatcagagga atcaattgtc cccaggccag 420

aagggaaagg ccatgctaat ggagatgtgt ccctcaaggt gaactcttct ttactacccc 480

cgaaggcccc agagctgaag gatataatcc tgtccttgcc ccctgacctt ggcccagccc 540

ttcaggagct caaggctcct tctttttagg gatatgccac attacctgct cccatcacca 600

tccagcccag caccctggac ttcagtcgcc tggcccaacc ccatgattat ggaacgtctc 660

ttcaccttaa cccaaatcca gacccttttg tggtttctat ttgt 704

<210> 42

<211> 3381

<212> DNA

<213> Homo sapiens

<400> 42
ctcatatttc ataaaatatg aacttttccc ggcccacatc cctaggcctt cctgatgcgc 60

ttgcctgctc cctggtctct ctgcatgggg aaggagtgtt cccagcttgc aaactccagc 120

tttgcctgtg agaggaacaa gcgtccctga tccagaaggt gttcagatgg agatggcgag 180

ttctgctggc tcctggctct ctggctgcct catccctctc gtcttcctcc ggctgtctgt 240

gcatgtgtca ggccacgcag gggatgccgg caagttccac gtggccctac taggggcac 300

agccgagctg ctctgccctc tctccctctg gcccgggacg gtacccaagg aggtgaggtg 360

gctgcggtcc ccattcccgc agcgttccca ggctgttcac atattccggg atgggaagga 420

ccaggatgaa gatctgatgc cggaatataa ggggaggacg gtgctagtga gagatgccca 480

agagggaagt gtcactctgc agatccttga cgtgcgcctt gaggaccaag ggtcttaccg 540

atgtctgatc caagttggaa atctgagtaa agaggacacc gtgatcctgc aggttgcagc 600

cccatctgtg gggagtctct ccccctcagc agtggctctg gctgtgatcc tgcctgtcct 660

ggtacttctc atcatggtgt gcctttgcct tatctggaag caaagaagag caaaagaaaa 720

gcttctctat gaacatgtga cggaggtgga caatcttctt tcagaccatg ctaaagaaaa 780

aggaaaactc cataaagctg tcaagaaact ccggagtgaa ctgaagttga aaagagctgc 840

agcaaactca ggctggagaa gagcccggtt gcattttgtg gcagtgaccc tggacccaga 900

cacagcacat cccaaactca tcctttctga ggaccaaaga tgtgtaaggc ttggagacag 960

acggcagcct gtacctgaca acccccagag atttgatttc gttgtcagca tcctaggctc 1020

tgagtacttc acgactggct gccactactg ggaggtgtat gtgggagaca agaccaaatg 1080

gattcttgga gtatgtagtg agtcagtgag caggaagggg aaggttactg cctcacctgc 1140

caatggacac tggcttctgc gacagagtcg tgggaatgag tatgaagctc tcacatcccc 1200

gcagacctcc ttccgcctta aagagcctcc acggtgtgtg gggattttcc tggactatga 1260

agcaggagtc atctctttct acaatgtgac caacaagtcc cacatcttta ctttcaccca 1320

caatttctct ggcccccttc gcccttttctt tgaaccttgc cttcatgatg gaggaaaaaa 1380

cacagcacct ctagtcattt gttcagaact acacaaatca gaggaatcaa ttgtccccag 1440

gccagaaggg aaaggccatg ctaatggaga tgtgtccctc aaggtgaact cttctttact 1500

acccccgaag gccccagagc tgaaggatat aatcctgtcc ttgccccctg accttggccc 1560

agcccttcag gagctcaagg ctccttcttt ttagggatat gccacattac ctgctcccat 1620

caccatccag cccagcaccc tggacttcag tcgcctggcc caaccccatg attatggaac 1680

gtctcttcac cttaacccaa atccagaccc ttttgtggtt tctatttgta ccacttttct 1740

cccaggcctc agttctgaag cttacctttc ttctaaggaa ttgaagctcc cagtgacctg 1800

gagggaggat tcctggaaac caaacaatca gtttaggtgc aggtggagat gttgaatatg 1860

tgttaccaag atacagcaca ggttcaggga aaagagttcg ctactccagg ggttatttag 1920

aagacacttt ctctgcctca tcctgccctc aagctttagt caagaagtta tggcccccag 1980

tccctgactt cttacttatc ccattgagga ctgcctttct ctctctcagt tctggcctct 2040

gccccccaaa gtcagctctc taaaagcaag catgttttag accactcact ctttccctct 2100

ttcttcagga atgaattggg aaaggctgat gagtaaaaca taccatcctt ttctattttc 2160

ttgatgctgt ttacaacata gtttggtgat atccagagct aatgtacatg ctttcaaaag 2220

ctaatctgcc tgttgatgat aactaggtac agcgacttta aatacagttg ctataatcct 2280

gaaaagcccc aggagcacat cagggagctg ggaaacacag ttgcagagaa ctcagctgta 2340

tgttgccctc tgaccttggc ccagaccttc agaggctcag tctgttgagt ctgttgtgac 2400

tgtcttatga atcaatctgt tgtgccaacc ctttctgaga ttcagagagg tccagctaga 2460

aaaagtggca gttttaacca ccaacgtaga agctttttt tccccacaag accattgtac 2520

tgcaatactt gactatttgt gtagcaaaca gcctcttagg ttggagagta ttgatttcaa 2580

ataggaagtt ggtagactag gtgtgaggat gaaataatga ccttgatttt tgggtgtgta 2640

ttgcagaagc ctcgtcgctt tcaagtcaca tcatatatgc gatctggcct aaccatggag 2700

atattggtta tcagtttgca gatgaataca cagttctatc caacagaaac tgtatctttc 2760

tgtttgtcag aagttctccg ttagttcctg tattagtcaa ggttttccag agaaacagaa 2820

tcaatatata ttggggcagg ggggcggtgg ttattataag gaattggctc acgtgattat 2880

ggaggcggtt aaatcccaag atctgcagga taagtcagca agatggagtc ccatgagagc 2940

tgatggttta gttccagtct gatggcagca ggcttgacgc caaggaagag atgatgttta 3000

attcaagtcc gaaggcaagg aaaaagctga tggtcctgtc caaaggctat taggcaggaa 3060

gaattctctt agggcagagt tagctctttt gttctattca ggccttcaac tgattcagca 3120

aggcccgccc acatttggga ggacagtctg cattactcag tctactgatt tgaatgttaa 3180

tgtcattgag aaacaccctc acaggaacac tcagaataat gtttgaccaa atagctgggc 3240

atcttgtgac ccagttaagt ggatacataa gattaactat cacagtgact cagtggaatt 3300

ttttgtttgc ttttgtacag ttttaaaata aaagtttggt atttgtgttt taaaaaaaaa 3360

aaaaaaaaaa aaaaaaaaaa a 3381

<210> 43

<211> 184

<212> PRT

<213> Homo sapiens

<400> 43

Met Pro Phe Asn Gly Glu Lys Gln Cys Val Gly Glu Asp Gln Pro Ser
1 5 10 15

Asp Ser Asp Ser Ser Arg Phe Ser Glu Ser Met Ala Ser Leu Ser Asp
20 25 30

Tyr Glu Cys Ser Arg Gln Ser Phe Thr Ser Asp Ser Ser Ser Lys Ser
35 40 45

Ser Ser Pro Ala Ser Thr Ser Pro Pro Arg Val Val Thr Phe Asp Glu
50 55 60

Val Met Ala Thr Ala Arg Asn Leu Ser Asn Leu Thr Leu Ala His Glu
65 70 75 80

Ile Ala Val Asn Glu Asn Leu Gln Leu Lys Gln Glu Ala Leu Pro Glu
85 90 95

Lys Ser Leu Ala Gly Arg Val Lys His Ile Val His Gln Ala Phe Trp
100 105 110

Asp Val Leu Asp Ser Glu Leu Asn Ala Asp Pro Pro Glu Ile Glu His
115 120 125

Ala Ile Lys Leu Phe Glu Glu Ile Arg Glu Ile Leu Leu Ser Phe Leu
130 135 140

Thr Pro Gly Gly Asn Arg Leu Arg Asn Gln Ile Cys Glu Val Leu Asp
145 150 155 160

Thr Asp Leu Ile Arg Gln Gln Ala Glu His Ser Ala Val Asp Ile Gln
165 170 175

Gly Leu Ala Asn Tyr Val Ile Ser
180

<210> 44

<211> 258

<212> PRT

<213> Homo sapiens

<400> 44

Met Pro Phe Asn Gly Glu Lys Gln Cys Val Gly Glu Asp Gln Pro Ser
1 5 10 15

Asp Ser Asp Ser Ser Arg Phe Ser Glu Ser Met Ala Ser Leu Ser Asp
20 25 30

Tyr Glu Cys Ser Arg Gln Ser Phe Thr Ser Asp Ser Ser Ser Lys Ser
35 40 45

Ser Ser Pro Ala Ser Thr Ser Pro Pro Arg Val Val Thr Phe Asp Glu
50 55 60

Val Met Ala Thr Ala Arg Asn Leu Ser Asn Leu Thr Leu Ala His Glu
65 70 75 80

Ile Ala Val Asn Glu Asn Phe Gln Leu Lys Gln Glu Ala Leu Pro Glu
85 90 95

```
Lys Ser Leu Ala Gly Arg Val Lys His Ile Val His Gln Ala Phe Trp
            100                 105                 110

Asp Val Leu Asp Ser Glu Leu Asn Ala Asp Pro Pro Glu Phe Glu His
        115                 120                 125

Ala Ile Lys Leu Phe Glu Glu Ile Arg Glu Ile Leu Leu Ser Phe Leu
    130                 135                 140

Thr Pro Gly Gly Asn Arg Leu Arg Asn Gln Ile Cys Glu Val Leu Asp
145                 150                 155                 160

Thr Asp Leu Ile Arg Gln Gln Ala Glu His Ser Ala Val Asp Ile Gln
                165                 170                 175

Gly Leu Ala Asn Tyr Val Ile Ser Thr Met Gly Lys Leu Cys Ala Pro
            180                 185                 190

Val Arg Asp Asn Asp Ile Arg Glu Leu Lys Ala Thr Gly Asn Ile Val
        195                 200                 205

Glu Val Leu Arg Gln Ile Phe His Val Leu Asp Leu Met Gln Met Asp
    210                 215                 220

Met Ala Asn Phe Thr Ile Met Ser Leu Arg Pro His Leu Gln Arg Gln
225                 230                 235                 240

Leu Val Glu Tyr Glu Arg Thr Lys Phe Gln Glu Ile Leu Glu Glu Thr
                245                 250                 255

Pro Ser
```

<210> 45

<211> 17

<212> PRT

<213> Homo sapiens

<400> 45

```
Met Pro Phe Asn Gly Glu Lys Gln Cys Val Gly Glu Asp Gln Pro Ser
1               5                   10                  15

Asp
```

<210> 46

<211> 17

<212> PRT

<213> Homo sapiens

<400> 46

Met Pro Phe Asn Gly Glu Lys Gln Cys Val Gly Glu Asp Gln Pro Ser
1 5 10 15

Asp

<210> 47

<211> 122

<212> PRT

<213> Homo sapiens

<400> 47

Ala Arg Asp Tyr Leu Lys Thr Leu Thr Glu Arg Leu Ala Arg Leu Arg
1 5 10 15

Arg Ala Arg Arg Ala Leu Arg Arg Arg Asn Ser Ile Lys Lys Met Ala
20 25 30

Ala Leu Thr Pro Arg Lys Arg Lys Gln Asp Ser Leu Lys Cys Asp Ser
35 40 45

Leu Leu His Phe Thr Glu Asn Leu Phe Pro Ser Pro Asn Lys Lys His
50 55 60

Cys Phe Tyr Gln Asn Ser Asp Lys Asn Glu Glu Asn Leu His Cys Ser
65 70 75 80

Gln Gln Glu His Phe Val Leu Ser Ala Leu Lys Thr Thr Glu Ile Asn
85 90 95

Arg Leu Pro Ser Ala Asn Gln Gly Ser Pro Phe Lys Ser Ala Leu Ser
100 105 110

Thr Val Ser Phe Tyr Asn Gln Asn Lys Trp
115 120

<210> 48

<211> 297

<212> PRT

<213> Homo sapiens

<400> 48

Arg Arg Asn Ser Ile Lys Lys Met Ala Ala Leu Thr Pro Arg Lys Arg
1 5 10 15

Lys Gln Asp Ser Leu Lys Cys Asp Ser Leu Leu His Phe Thr Glu Asn
20 25 30

Leu Phe Pro Ser Pro Asn Lys Lys His Cys Phe Tyr Gln Asn Ser Asp
35 40 45

```
Lys Asn Glu Glu Asn Leu His Cys Ser Gln Gln Glu His Phe Val Leu
    50                  55                  60

Ser Ala Leu Lys Thr Thr Glu Ile Asn Arg Leu Pro Ser Ala Asn Gln
65                  70                  75                  80

Gly Ser Pro Phe Lys Ser Ala Leu Ser Thr Val Ser Phe Tyr Asn Gln
                85                  90                  95

Asn Lys Trp Tyr Leu Asn Pro Leu Glu Arg Lys Leu Ile Lys Glu Ser
            100                 105                 110

Arg Ser Thr Cys Leu Lys Thr Asn Asp Glu Asp Lys Ser Phe Pro Ile
        115                 120                 125

Val Thr Glu Lys Met Gln Gly Lys Pro Val Cys Ser Lys Lys Asn Asn
    130                 135                 140

Lys Lys Pro Gln Lys Ser Leu Thr Ala Lys Tyr Gln Pro Lys Tyr Arg
145                 150                 155                 160

His Ile Lys Pro Val Ser Arg Asn Ser Arg Asn Ser Lys Gln Asn Arg
                165                 170                 175

Val Ile Tyr Lys Pro Ile Val Glu Lys Glu Asn Asn Cys His Ser Ala
            180                 185                 190

Glu Asn Asn Ser Asn Ala Pro Arg Val Leu Ser Gln Lys Ile Lys Pro
        195                 200                 205

Gln Val Thr Leu Gln Gly Gly Ala Ala Phe Phe Val Arg Lys Lys Ser
    210                 215                 220

Ser Leu Arg Lys Ser Ser Leu Glu Asn Glu Pro Ser Leu Gly Arg Thr
225                 230                 235                 240

Gln Lys Ser Lys Ser Glu Val Ile Glu Asp Ser Asp Val Glu Thr Val
                245                 250                 255

Ser Glu Lys Lys Thr Phe Ala Thr Arg Gln Val Pro Lys Cys Leu Val
            260                 265                 270

Leu Glu Glu Lys Leu Lys Ile Gly Leu Leu Ser Ala Ser Ser Lys Asn
        275                 280                 285

Lys Glu Lys Leu Ile Lys Val Lys Leu
    290                 295

<210>  49

<211>  60

<212>  PRT

<213>  Homo sapiens


<400>  49

Ser Leu Leu His Phe Thr Glu Asn Leu Phe Pro Ser Pro Asn Lys Lys
1               5                   10                  15
```

His Cys Phe Tyr Gln Asn Ser Asp Lys Asn Glu Glu Asn Leu His Cys
20 25 30

Ser Gln Gln Glu His Phe Val Leu Ser Ala Leu Lys Thr Thr Glu Ile
35 40 45

Asn Arg Leu Pro Ser Ala Asn Gln Gly Ser Pro Phe
50 55 60

<210> 50

<211> 59

<212> PRT

<213> Homo sapiens

<400> 50

Gly Ala Arg Ile Leu Glu Thr Ala Thr Arg Val Gly Gly Ala Arg Ala
1 5 10 15

Gly Ala Pro Pro Ile Thr Glu Pro Ala Pro Gly Ser Ala Glu Pro Trp
20 25 30

Pro Ile Gly Thr Gln Arg Leu Pro Pro Ser Leu Ala Arg Asp Tyr Leu
35 40 45

Lys Thr Leu Thr Glu Arg Leu Ala Arg Leu Arg
50 55

<210> 51

<211> 23

<212> PRT

<213> Homo sapiens

<400> 51

Arg Asn Ser Ile Lys Lys Met Ala Ala Leu Thr Pro Arg Lys Arg Lys
1 5 10 15

Gln Asp Ser Leu Lys Cys Asp
20

<210> 52

<211> 156

<212> PRT

<213> Homo sapiens

<400> 52

```
Gly Ala Arg Ile Leu Glu Thr Ala Thr Arg Val Gly Gly Ala Arg Ala
1               5                  10                  15

Gly Ala Pro Pro Ile Thr Glu Pro Ala Pro Gly Ser Ala Glu Pro Trp
            20                  25                  30

Pro Ile Gly Thr Gln Arg Leu Pro Pro Ser Leu Ala Arg Asp Tyr Leu
        35                  40                  45

Lys Thr Leu Thr Glu Arg Leu Ala Arg Leu Arg Arg Ala Arg Arg Ala
    50                  55                  60

Leu Arg Arg Arg Asn Ser Ile Lys Lys Met Ala Ala Leu Thr Pro Arg
65                  70                  75                  80

Lys Arg Lys Gln Asp Ser Leu Lys Cys Asp Ser Leu Leu His Phe Thr
                85                  90                  95

Glu Asn Leu Phe Pro Ser Pro Asn Lys Lys His Cys Phe Tyr Gln Asn
            100                 105                 110

Ser Asp Lys Asn Glu Glu Asn Pro His Cys Ser Gln Gln Glu His Phe
        115                 120                 125

Val Leu Ser Ala Leu Lys Thr Thr Glu Ile Asn Arg Leu Pro Ser Ala
    130                 135                 140

Asn Gln Gly Ser Pro Phe Lys Ser Ala Leu Ser Thr
145                 150                 155
```

<210> 53

<211> 313

<212> PRT

<213> Homo sapiens

<400> 53

```
Leu Lys Thr Leu Thr Glu Arg Leu Ala Arg Leu Arg Arg Ala Arg Arg
1               5                  10                  15

Ala Leu Arg Arg Arg Asn Ser Ile Lys Lys Met Ala Ala Leu Thr Pro
            20                  25                  30

Arg Lys Arg Lys Gln Asp Ser Leu Lys Cys Asp Ser Leu Leu His Phe
        35                  40                  45

Thr Glu Asn Leu Phe Pro Ser Pro Asn Lys Lys His Cys Phe Tyr Gln
    50                  55                  60

Asn Ser Asp Lys Asn Glu Glu Asn Leu His Cys Ser Gln Gln Glu His
65                  70                  75                  80

Phe Val Leu Ser Ala Leu Lys Thr Thr Glu Ile Asn Arg Leu Pro Ser
                85                  90                  95

Ala Asn Gln Gly Ser Pro Phe Lys Ser Ala Leu Ser Thr Val Ser Phe
            100                 105                 110
```

Tyr Asn Gln Asn Lys Trp Tyr Leu Asn Pro Leu Glu Arg Lys Leu Ile
115 120 125

Lys Glu Ser Arg Ser Thr Cys Leu Lys Thr Asn Asp Glu Asp Lys Ser
130 135 140

Phe Pro Ile Val Thr Glu Lys Met Gln Gly Lys Pro Val Cys Ser Lys
145 150 155 160

Lys Asn Asn Lys Lys Pro Gln Lys Ser Leu Thr Ala Lys Tyr Gln Pro
165 170 175

Lys Tyr Arg His Ile Lys Pro Val Ser Arg Asn Ser Arg Asn Ser Lys
180 185 190

Gln Asn Arg Val Ile Tyr Lys Pro Ile Val Glu Lys Glu Asn Asn Cys
195 200 205

His Ser Ala Glu Asn Asn Ser Asn Ala Pro Arg Val Leu Ser Gln Lys
210 215 220

Ile Lys Pro Gln Val Thr Leu Gln Gly Gly Ala Ala Phe Phe Val Arg
225 230 235 240

Lys Lys Ser Ser Leu Arg Lys Ser Ser Leu Glu Asn Glu Pro Ser Leu
245 250 255

Gly Arg Thr Gln Lys Ser Lys Ser Glu Val Ile Glu Asp Ser Asp Val
260 265 270

Glu Thr Val Ser Glu Lys Lys Thr Phe Ala Thr Arg Gln Val Pro Lys
275 280 285

Cys Leu Val Leu Glu Glu Lys Leu Lys Ile Gly Leu Leu Ser Ala Ser
290 295 300

Ser Lys Asn Lys Glu Lys Leu Ile Lys
305 310

<210> 54

<211> 279

<212> PRT

<213> Homo sapiens

<400> 54

Arg Lys Lys Val Asn Pro Tyr Glu Glu Val Asp Gln Glu Lys Tyr Ser
1 5 10 15

Asn Leu Val Gln Ser Val Leu Ser Ser Arg Gly Val Ala Gln Thr Pro
20 25 30

Gly Ser Val Glu Glu Asp Ala Leu Leu Cys Gly Pro Val Ser Lys His
35 40 45

Lys Leu Pro Asn Gln Gly Glu Asp Arg Arg Val Pro Gln Asn Trp Phe
50 55 60

```
Pro Ile Phe Asn Pro Glu Arg Ser Asp Lys Pro Asn Ala Ser Asp Pro
65                  70                  75                  80

Ser Val Pro Leu Lys Ile Pro Leu Gln Arg Asn Val Ile Pro Ser Val
                85                  90                  95

Thr Arg Val Leu Gln Gln Thr Met Ala Lys Gln Gln Val Phe Leu Leu
            100                 105                 110

Glu Arg Trp Lys Gln Arg Met Ile Leu Glu Leu Gly Glu Asp Gly Phe
        115                 120                 125

Lys Glu Tyr Thr Ser Asn Val Phe Leu Gln Gly Lys Arg Phe His Glu
    130                 135                 140

Ala Leu Glu Ser Ile Leu Ser Pro Gln Glu Thr Leu Lys Glu Arg Asp
145                 150                 155                 160

Glu Asn Leu Leu Lys Ser Gly Tyr Ile Glu Ser Val Gln His Ile Leu
                165                 170                 175

Lys Asp Val Ser Gly Val Arg Ala Leu Glu Ser Ala Val Gln His Glu
            180                 185                 190

Thr Leu Asn Tyr Ile Gly Leu Leu Asp Cys Val Ala Glu Tyr Gln Gly
        195                 200                 205

Lys Leu Cys Val Ile Asp Trp Lys Thr Ser Glu Lys Pro Lys Pro Phe
    210                 215                 220

Ile Gln Ser Ile Phe Asp Asn Pro Leu Gln Val Val Ala Tyr Met Gly
225                 230                 235                 240

Ala Met Asn His Asp Thr Asn Tyr Ser Phe Gln Val Gln Cys Gly Leu
                245                 250                 255

Ile Val Val Ala Tyr Lys Asp Gly Ser Pro Ala His Pro His Phe Met
            260                 265                 270

Asp Ala Glu Leu Cys Ser Gln
        275

<210>  55

<211>  307

<212>  PRT

<213>  Homo sapiens


<400>  55

Arg Lys Lys Val Asn Pro Tyr Glu Glu Val Asp Gln Glu Lys Tyr Ser
1               5                   10                  15

Asn Leu Val Gln Ser Val Leu Ser Ser Arg Gly Val Ala Gln Thr Pro
            20                  25                  30

Gly Ser Val Glu Glu Asp Ala Leu Leu Cys Gly Pro Val Ser Lys His
        35                  40                  45
```

```
Lys Leu Pro Asn Gln Gly Glu Asp Arg Arg Val Pro Gln Asn Trp Phe
    50                  55                  60

Pro Ile Phe Asn Pro Glu Arg Ser Asp Lys Pro Asn Ala Ser Asp Pro
65                  70                  75                  80

Ser Val Pro Leu Lys Ile Pro Leu Gln Arg Asn Val Ile Pro Ser Val
                85                  90                  95

Thr Arg Val Leu Gln Gln Thr Met Thr Lys Gln Gln Val Phe Leu Leu
            100                 105                 110

Glu Arg Trp Lys Gln Arg Met Ile Leu Glu Leu Gly Glu Asp Gly Phe
        115                 120                 125

Lys Glu Tyr Thr Ser Asn Val Phe Leu Gln Gly Lys Arg Phe His Glu
    130                 135                 140

Ala Leu Glu Ser Ile Leu Ser Pro Gln Glu Thr Leu Lys Glu Arg Asp
145                 150                 155                 160

Glu Asn Leu Leu Lys Ser Gly Tyr Ile Glu Ser Val Gln His Ile Leu
                165                 170                 175

Lys Asp Val Ser Gly Val Arg Ala Leu Glu Ser Ala Val Gln His Glu
            180                 185                 190

Thr Leu Asn Tyr Ile Gly Leu Leu Asp Cys Val Ala Glu Tyr Gln Gly
        195                 200                 205

Lys Leu Cys Val Ile Asp Trp Lys Thr Ser Glu Lys Pro Lys Pro Phe
    210                 215                 220

Ile Gln Ser Thr Phe Asp Asn Pro Leu Gln Val Val Ala Tyr Met Gly
225                 230                 235                 240

Ala Met Asn His Asp Thr Asn Tyr Ser Phe Gln Val Gln Cys Gly Leu
                245                 250                 255

Ile Val Val Ala Tyr Lys Asp Gly Ser Pro Ala His Pro His Phe Met
            260                 265                 270

Asp Ala Glu Leu Cys Ser Gln Tyr Trp Thr Lys Trp Leu Leu Arg Leu
        275                 280                 285

Glu Glu Tyr Thr Glu Lys Lys Lys Asn Gln Asn Ile Gln Lys Pro Glu
    290                 295                 300

Tyr Ser Glu
305
```

<210> 56

<211> 82

<212> PRT

<213> Homo sapiens

<400> 56

```
Met Lys Leu Phe Gln Thr Ile Cys Arg Gln Leu Arg Ser Ser Lys Phe
1               5                  10                  15

Ser Val Glu Ser Ala Ala Leu Val Ala Phe Ser Thr Ser Ser Tyr Ser
            20                  25                  30

Cys Gly Arg Lys Lys Lys Val Asn Pro Tyr Glu Glu Val Asp Gln Glu
        35                  40                  45

Lys Tyr Ser Asn Leu Val Gln Ser Val Leu Ser Ser Arg Gly Val Ala
    50                  55                  60

Gln Thr Pro Gly Ser Val Glu Glu Asp Ala Leu Leu Cys Gly Pro Val
65                  70                  75                  80

Ser Lys
```

<210> 57

<211> 298

<212> PRT

<213> Homo sapiens

<400> 57

```
Gln His Glu Glu Phe Ile Leu Leu Ser Gln Gly Glu Val Glu Lys Leu
1               5                  10                  15

Ile Lys Cys Asp Glu Ile Gln Val Asp Ser Glu Glu Pro Val Phe Glu
            20                  25                  30

Ala Val Ile Asn Trp Val Lys His Ala Lys Lys Glu Arg Glu Glu Ser
        35                  40                  45

Leu Pro Asn Leu Leu Gln Tyr Val Arg Met Pro Leu Leu Thr Pro Arg
    50                  55                  60

Tyr Ile Thr Asp Val Ile Asp Ala Glu Pro Phe Ile Arg Cys Ser Leu
65                  70                  75                  80

Gln Cys Arg Asp Leu Val Asp Glu Ala Lys Lys Phe His Leu Arg Pro
                85                  90                  95

Glu Leu Arg Ser Gln Met Gln Gly Pro Arg Thr Arg Ala Arg Leu Gly
            100                 105                 110

Ala Asn Glu Val Leu Leu Val Val Gly Gly Phe Gly Ser Gln Gln Ser
        115                 120                 125

Pro Ile Asp Val Val Glu Lys Tyr Asp Pro Lys Thr Gln Glu Trp Ser
    130                 135                 140

Phe Leu Pro Ser Ile Thr Arg Lys Arg Arg Tyr Val Ala Ser Val Ser
145                 150                 155                 160

Leu His Asp Arg Ile Tyr Val Ile Gly Gly Tyr Asp Gly Arg Ser Arg
                165                 170                 175
```

Leu Ser Ser Val Glu Cys Leu Asp Tyr Thr Ala Asp Glu Asp Gly Val
180 185 190

Trp Tyr Ser Val Ala Pro Met Asn Val Arg Arg Gly Leu Ala Gly Ala
195 200 205

Thr Thr Leu Gly Asp Met Ile Tyr Val Ser Gly Gly Phe Asp Gly Ser
210 215 220

Arg Arg His Thr Ser Met Glu Arg Tyr Asp Pro Asn Ile Asp Gln Trp
225 230 235 240

Ser Met Leu Gly Asp Met Gln Thr Ala Arg Glu Gly Ala Gly Leu Val
245 250 255

Val Ala Ser Gly Val Ile Tyr Cys Leu Gly Gly Tyr Asp Gly Leu Asn
260 265 270

Ile Leu Asn Ser Val Glu Lys Tyr Asp Pro His Thr Gly His Trp Thr
275 280 285

Asn Val Thr Pro Met Ala Thr Lys Arg Ser
290 295

<210> 58

<211> 189

<212> PRT

<213> Homo sapiens

<400> 58

Val Asp Ser Glu Glu Pro Val Phe Glu Ala Val Ile Asn Trp Val Lys
1 5 10 15

His Ala Lys Lys Glu Arg Glu Glu Ser Leu Pro Asn Leu Leu Gln Tyr
20 25 30

Val Arg Met Pro Leu Leu Thr Pro Arg Tyr Ile Thr Asp Val Ile Asp
35 40 45

Ala Glu Pro Phe Ile Arg Cys Ser Leu Gln Cys Arg Asp Leu Val Asp
50 55 60

Glu Ala Lys Lys Phe His Leu Arg Pro Glu Leu Arg Ser Gln Met Gln
65 70 75 80

Gly Pro Arg Thr Arg Ala Arg Leu Gly Ala Asn Glu Val Leu Leu Val
85 90 95

Val Gly Gly Phe Gly Ser Gln Gln Ser Pro Ile Asp Val Val Glu Lys
100 105 110

Tyr Asp Pro Lys Thr Gln Glu Trp Ser Phe Leu Pro Ser Ile Thr Arg
115 120 125

Lys Arg Arg Tyr Val Ala Ser Val Ser Leu His Asp Arg Ile Tyr Val
130 135 140

Ile Gly Gly Tyr Asp Gly Arg Ser Arg Leu Ser Ser Val Glu Cys Leu
145 150 155 160

Asp Tyr Thr Ala Asp Glu Asp Gly Val Trp Tyr Ser Val Ala Pro Met
165 170 175

Asn Val Arg Arg Gly Leu Ala Gly Ala Thr Thr Leu Gly
180 185

<210> 59

<211> 448

<212> PRT

<213> Homo sapiens

<400> 59

Gln Leu Lys Gly Val Lys Gln Ala Cys Cys Glu Phe Leu Glu Ser Gln
1 5 10 15

Leu Asp Pro Ser Asn Cys Leu Gly Ile Arg Asp Phe Ala Glu Thr His
20 25 30

Asn Cys Val Asp Leu Met Gln Ala Ala Glu Val Phe Ser Gln Lys His
35 40 45

Phe Pro Glu Val Val Gln His Glu Glu Phe Ile Leu Leu Ser Gln Gly
50 55 60

Glu Val Glu Lys Leu Ile Lys Cys Asp Glu Ile Gln Val Asp Ser Glu
65 70 75 80

Glu Pro Val Phe Glu Ala Val Ile Asn Trp Val Lys His Ala Lys Lys
85 90 95

Glu Arg Glu Glu Ser Leu Pro Asn Leu Leu Gln Tyr Val Arg Met Pro
100 105 110

Leu Leu Thr Pro Arg Tyr Ile Thr Asp Val Ile Asp Ala Glu Pro Phe
115 120 125

Ile Arg Cys Ser Leu Gln Cys Arg Asp Leu Val Asp Glu Ala Lys Lys
130 135 140

Phe His Leu Arg Pro Glu Leu Arg Ser Gln Met Gln Gly Pro Arg Thr
145 150 155 160

Arg Ala Arg Leu Gly Ala Asn Glu Val Leu Leu Val Val Gly Gly Phe
165 170 175

Gly Ser Gln Gln Ser Pro Ile Asp Val Val Glu Lys Tyr Asp Pro Lys
180 185 190

Thr Gln Glu Trp Ser Phe Leu Pro Ser Ile Thr Arg Lys Arg Arg Tyr
195 200 205

Val Ala Ser Met Ser Leu His Asp Arg Ile Tyr Val Ile Gly Gly Tyr
210 215 220

```
Asp Gly Arg Ser Arg Leu Ser Ser Val Glu Cys Leu Asp Tyr Thr Ala
225                 230                 235                 240

Asp Glu Asp Gly Val Trp Tyr Ser Val Ala Pro Met Asn Val Arg Arg
                245                 250                 255

Gly Leu Ala Gly Ala Thr Thr Leu Gly Asp Met Ile Tyr Val Ser Gly
            260                 265                 270

Gly Phe Asp Gly Ser Arg Arg His Thr Ser Met Glu Arg Tyr Asp Pro
        275                 280                 285

Asn Ile Asp Gln Trp Ser Met Leu Gly Asp Met Gln Thr Ala Arg Glu
    290                 295                 300

Gly Ala Gly Leu Val Val Ala Ser Gly Val Ile Tyr Cys Leu Gly Gly
305                 310                 315                 320

Tyr Asp Gly Leu Asn Ile Leu Asn Ser Val Glu Lys Tyr Asp Pro His
                325                 330                 335

Thr Gly His Trp Thr Asn Val Thr Pro Met Ala Thr Lys Arg Ser Gly
            340                 345                 350

Ala Gly Val Ala Leu Leu Asn Asp His Ile Tyr Val Val Gly Gly Phe
        355                 360                 365

Asp Gly Thr Ala His Leu Ser Ser Val Glu Ala Tyr Asn Ile Arg Thr
    370                 375                 380

Asp Ser Trp Thr Thr Val Thr Ser Met Thr Thr Pro Arg Cys Tyr Val
385                 390                 395                 400

Gly Ala Thr Val Leu Arg Gly Arg Leu Tyr Ala Ile Ala Gly Tyr Asp
                405                 410                 415

Gly Asn Ser Leu Leu Ser Ser Ile Glu Cys Tyr Asp Pro Ile Ile Asp
            420                 425                 430

Ser Trp Glu Val Val Thr Ser Met Gly Thr Gln Arg Cys Asp Ala Gly
        435                 440                 445
```

<210> 60

<211> 189

<212> PRT

<213> Homo sapiens

<400> 60

```
Met Ile Tyr Val Ser Gly Gly Phe Asp Gly Ser Arg Arg His Thr Ser
1               5                   10                  15

Met Glu Arg Tyr Asp Pro Asn Ile Asp Gln Trp Ser Met Leu Gly Asp
            20                  25                  30

Met Gln Thr Ala Arg Glu Gly Ala Gly Leu Val Val Ala Ser Gly Val
        35                  40                  45
```

```
Ile Tyr Cys Leu Gly Gly Tyr Asp Gly Leu Asn Ile Leu Asn Ser Val
    50                  55                  60

Glu Lys Tyr Asp Pro His Thr Gly His Trp Thr Asn Val Thr Pro Met
65                  70                  75                  80

Ala Thr Lys Arg Ser Gly Ala Gly Val Ala Leu Leu Asn Asp His Ile
                85                  90                  95

Tyr Val Val Gly Gly Phe Asp Gly Thr Ala His Leu Ser Ser Val Glu
            100                 105                 110

Ala Tyr Asn Ile Arg Thr Asp Ser Trp Thr Thr Val Thr Ser Met Thr
        115                 120                 125

Thr Pro Arg Cys Tyr Val Gly Ala Thr Val Leu Arg Gly Arg Leu Tyr
    130                 135                 140

Ala Ile Ala Gly Tyr Asp Gly Asn Ser Leu Leu Ser Ser Ile Glu Cys
145                 150                 155                 160

Tyr Asp Pro Ile Ile Asp Ser Trp Glu Val Val Thr Ser Met Gly Thr
                165                 170                 175

Gln Arg Cys Asp Ala Gly Val Cys Val Leu Arg Glu Lys
            180                 185
```

<210> 61

<211> 189

<212> PRT

<213> Homo sapiens

<400> 61

```
Val Asp Ser Glu Glu Pro Val Phe Glu Ala Val Ile Asn Trp Val Lys
1               5                   10                  15

His Ala Lys Lys Glu Arg Glu Glu Ser Leu Pro Asn Leu Leu Gln Tyr
            20                  25                  30

Val Arg Met Pro Leu Leu Thr Pro Arg Tyr Ile Thr Asp Val Ile Asp
        35                  40                  45

Ala Glu Pro Phe Ile Arg Cys Ser Leu Gln Cys Arg Asp Leu Val Asp
    50                  55                  60

Glu Ala Lys Lys Phe His Leu Arg Pro Glu Leu Arg Ser Gln Met Gln
65                  70                  75                  80

Gly Pro Arg Thr Arg Ala Arg Leu Gly Ala Asn Glu Val Leu Leu Val
                85                  90                  95

Val Gly Gly Phe Gly Ser Gln Gln Ser Pro Ile Asp Val Val Glu Lys
            100                 105                 110

Tyr Asp Pro Lys Thr Gln Glu Trp Ser Phe Leu Pro Ser Ile Thr Arg
        115                 120                 125
```

Lys Arg Arg Tyr Val Ala Ser Met Ser Leu His Asp Arg Ile Tyr Val
130 135 140

Ile Gly Gly Tyr Asp Gly Arg Ser Arg Leu Ser Ser Val Glu Cys Leu
145 150 155 160

Asp Tyr Thr Ala Asp Glu Asp Gly Val Trp Tyr Ser Val Ala Pro Met
165 170 175

Asn Val Arg Arg Gly Leu Ala Gly Ala Thr Thr Leu Gly
180 185

<210> 62

<211> 414

<212> PRT

<213> Homo sapiens

<400> 62

Met Gly Gly Ile Met Ala Pro Lys Asp Ile Met Thr Asn Thr His Ala
1 5 10 15

Lys Ser Ile Leu Asn Ser Met Asn Ser Leu Arg Lys Ser Asn Thr Leu
20 25 30

Cys Asp Val Thr Leu Arg Val Glu Gln Lys Asp Phe Pro Ala His Arg
35 40 45

Ile Val Leu Ala Ala Cys Ser Asp Tyr Phe Cys Ala Met Phe Thr Ser
50 55 60

Glu Leu Ser Glu Lys Gly Lys Pro Tyr Val Asp Ile Gln Gly Leu Thr
65 70 75 80

Ala Ser Thr Met Glu Ile Leu Leu Asp Phe Val Tyr Thr Glu Thr Val
85 90 95

His Val Thr Val Glu Asn Val Gln Glu Leu Leu Pro Ala Ala Cys Leu
100 105 110

Leu Gln Leu Lys Gly Val Lys Gln Ala Cys Cys Glu Phe Leu Glu Ser
115 120 125

Gln Leu Asp Pro Ser Asn Cys Leu Gly Ile Arg Asp Phe Ala Glu Thr
130 135 140

His Asn Cys Val Asp Leu Met Gln Ala Ala Glu Val Phe Ser Gln Lys
145 150 155 160

His Phe Pro Glu Val Val Gln His Glu Glu Phe Ile Leu Leu Ser Gln
165 170 175

Gly Glu Val Glu Lys Leu Ile Lys Cys Asp Glu Ile Gln Val Asp Ser
180 185 190

Glu Glu Pro Val Phe Glu Ala Val Ile Asn Trp Val Lys His Ala Lys
195 200 205

Lys Glu Arg Glu Glu Ser Leu Pro Asn Leu Leu Gln Tyr Val Arg Met
210 215 220

Pro Leu Leu Thr Pro Arg Tyr Ile Thr Asp Val Ile Asp Ala Glu Pro
225 230 235 240

Phe Ile Arg Cys Ser Leu Gln Cys Arg Asp Leu Val Asp Glu Ala Lys
245 250 255

Lys Phe His Leu Arg Pro Glu Leu Arg Ser Gln Met Gln Gly Pro Arg
260 265 270

Thr Arg Ala Arg Leu Asp Met Ile Tyr Val Ser Gly Gly Phe Asp Gly
275 280 285

Ser Arg Arg His Thr Ser Met Glu Arg Tyr Asp Pro Asn Ile Asp Gln
290 295 300

Trp Ser Met Leu Gly Asp Met Gln Thr Ala Arg Glu Gly Ala Gly Leu
305 310 315 320

Val Val Ala Ser Gly Val Ile Tyr Cys Leu Gly Gly Tyr Asp Gly Leu
325 330 335

Asn Ile Leu Asn Ser Val Glu Lys Tyr Asp Pro His Thr Gly His Trp
340 345 350

Thr Asn Val Thr Pro Met Ala Thr Lys Arg Ser Gly Ala Gly Val Ala
355 360 365

Leu Leu Asn Asp His Ile Tyr Val Val Gly Gly Phe Asp Gly Thr Ala
370 375 380

His Leu Ser Ser Val Glu Ala Tyr Asn Ile Arg Thr Asp Ser Trp Thr
385 390 395 400

Thr Val Thr Ser Met Thr Thr Pro Arg Cys Tyr Val Gly Ala
405 410

<210> 63

<211> 164

<212> PRT

<213> Homo sapiens

<400> 63

Arg Lys Ser Asn Thr Leu Cys Asp Val Thr Leu Arg Val Glu Gln Lys
1 5 10 15

Asp Phe Pro Ala His Arg Ile Val Leu Ala Ala Cys Ser Asp Tyr Phe
20 25 30

Cys Ala Met Phe Thr Ser Glu Leu Ser Glu Lys Gly Lys Pro Tyr Val
35 40 45

Asp Ile Gln Gly Leu Thr Ala Ser Thr Met Glu Ile Leu Leu Asp Phe
50 55 60

```
Val Tyr Thr Glu Thr Val His Val Thr Val Glu Asn Val Gln Glu Leu
65                  70                  75                  80

Leu Pro Ala Ala Cys Leu Leu Gln Leu Lys Gly Val Lys Gln Ala Cys
                85                  90                  95

Cys Glu Phe Leu Glu Ser Gln Leu Asp Pro Ser Asn Cys Leu Gly Ile
            100                 105                 110

Arg Asp Phe Ala Glu Thr His Asn Cys Val Asp Leu Met Gln Ala Ala
        115                 120                 125

Glu Val Phe Ser Gln Lys His Phe Pro Glu Val Val Gln His Glu Glu
    130                 135                 140

Phe Ile Leu Leu Ser Gln Gly Glu Val Glu Lys Leu Ile Lys Cys Asp
145                 150                 155                 160

Glu Ile Gln Val
```

<210> 64

<211> 299

<212> PRT

<213> Homo sapiens

<400> 64

```
Thr Gly Asp Phe Arg Tyr Thr Pro Ser Met Leu Lys Glu Pro Ala Leu
1               5                   10                  15

Thr Leu Gly Lys Gln Ile His Thr Leu Tyr Leu Asp Asn Thr Asn Cys
            20                  25                  30

Asn Pro Ala Leu Val Leu Pro Ser Arg Gln Glu Ala Ala His Gln Ile
        35                  40                  45

Val Gln Leu Ile Arg Lys His Pro Gln His Asn Ile Lys Ile Gly Leu
    50                  55                  60

Tyr Ser Leu Gly Lys Glu Ser Leu Leu Glu Gln Leu Ala Leu Glu Phe
65                  70                  75                  80

Gln Thr Trp Val Val Leu Ser Pro Arg Arg Leu Glu Leu Val Gln Leu
                85                  90                  95

Leu Gly Leu Ala Asp Val Phe Thr Val Glu Glu Lys Ala Gly Arg Ile
            100                 105                 110

His Ala Val Asp His Met Glu Ile Cys His Ser Asn Met Leu Arg Trp
        115                 120                 125

Asn Gln Thr His Pro Thr Ile Ala Ile Leu Pro Thr Ser Arg Lys Ile
    130                 135                 140

His Ser Ser His Pro Asp Ile His Val Ile Pro Tyr Ser Asp His Ser
145                 150                 155                 160
```

```
Ser Tyr Ser Glu Leu Arg Ala Phe Val Ala Ala Leu Lys Pro Cys Gln
                165                 170                 175

Val Val Pro Ile Val Ser Arg Arg Pro Cys Gly Gly Phe Gln Asp Ser
            180                 185                 190

Leu Ser Pro Arg Ile Ser Val Pro Leu Ile Pro Asp Ser Val Gln Gln
        195                 200                 205

Tyr Met Ser Ser Ser Ser Arg Lys Pro Ser Leu Leu Trp Leu Leu Glu
    210                 215                 220

Arg Arg Leu Lys Arg Pro Arg Thr Gln Gly Val Val Phe Glu Ser Pro
225                 230                 235                 240

Glu Glu Ser Ala Asp Gln Ser Gln Ala Asp Arg Asp Ser Lys Lys Ala
                245                 250                 255

Lys Lys Glu Lys Leu Ser Pro Trp Pro Ala Asp Leu Glu Lys Gln Pro
            260                 265                 270

Ser His His Pro Leu Arg Ile Lys Lys Gln Leu Phe Pro Asp Leu Tyr
        275                 280                 285

Ser Lys Glu Trp Asn Lys Ala Val Pro Phe Cys
    290                 295
```

<210> 65

<211> 64

<212> PRT

<213> Homo sapiens

<400> 65

```
Met Asn Gly Val Leu Ile Pro His Thr Pro Ile Ala Val Asp Phe Trp
1               5                   10                  15

Ser Leu Arg Arg Ala Gly Thr Ala Arg Leu Phe Phe Leu Ser His Met
            20                  25                  30

His Ser Asp His Thr Val Gly Leu Ser Ser Thr Trp Ala Arg Pro Leu
        35                  40                  45

Tyr Cys Ser Pro Ile Thr Ala His Leu Leu His Arg His Leu Gln Val
    50                  55                  60
```

<210> 66

<211> 19

<212> PRT

<213> Homo sapiens

<400> 66

Ala Gly Tyr Ser Ser Arg Arg Phe Asp Gln Gln Val Glu Lys Tyr His
1 5 10 15

Lys Pro Cys

<210> 67

<211> 421

<212> PRT

<213> Homo sapiens

<400> 67

Phe Gly Thr Ile Leu Tyr Thr Gly Asp Phe Arg Tyr Thr Pro Ser Met
1 5 10 15

Leu Lys Glu Pro Ala Leu Thr Leu Gly Lys Gln Ile His Thr Leu Tyr
20 25 30

Leu Asp Asn Thr Asn Cys Asn Pro Ala Leu Val Leu Pro Ser Arg Gln
35 40 45

Glu Ala Ala His Gln Ile Val Gln Leu Ile Arg Lys His Pro Gln His
50 55 60

Asn Ile Lys Ile Gly Leu Tyr Ser Leu Gly Lys Glu Ser Leu Leu Glu
65 70 75 80

Gln Leu Ala Leu Glu Phe Gln Thr Trp Val Val Leu Ser Pro Arg Arg
85 90 95

Leu Glu Leu Val Gln Leu Leu Gly Leu Ala Asp Val Phe Thr Val Glu
100 105 110

Glu Lys Ala Gly Arg Ile His Ala Val Asp His Met Glu Ile Cys His
115 120 125

Ser Asn Met Leu Arg Trp Asn Gln Thr His Pro Thr Ile Ala Ile Leu
130 135 140

Pro Thr Ser Arg Lys Ile His Ser Ser His Pro Asp Ile His Val Ile
145 150 155 160

Pro Tyr Ser Asp His Ser Ser Tyr Ser Glu Leu Arg Ala Phe Val Ala
165 170 175

Ala Leu Lys Pro Cys Gln Val Val Pro Ile Val Ser Arg Arg Pro Cys
180 185 190

Gly Gly Phe Gln Asp Ser Leu Ser Pro Arg Ile Ser Val Pro Leu Ile
195 200 205

Pro Asp Ser Val Gln Gln Tyr Met Ser Ser Ser Ser Arg Lys Pro Ser
210 215 220

Leu Leu Trp Leu Leu Glu Arg Arg Leu Lys Arg Pro Arg Thr Gln Gly
225 230 235 240

Val Val Phe Glu Ser Pro Glu Glu Ser Ala Asp Gln Ser Gln Ala Asp
245 250 255

Arg Asp Ser Lys Lys Ala Lys Lys Glu Lys Leu Ser Pro Trp Pro Ala
260 265 270

Asp Leu Glu Lys Gln Pro Ser His His Pro Leu Arg Ile Lys Lys Gln
275 280 285

Leu Phe Pro Asp Leu Tyr Ser Lys Glu Trp Asn Lys Ala Val Pro Phe
290 295 300

Cys Glu Ser Gln Lys Arg Val Thr Met Leu Thr Ala Pro Leu Gly Phe
305 310 315 320

Ser Val His Leu Arg Ser Thr Asp Glu Glu Phe Ile Ser Gln Lys Thr
325 330 335

Arg Glu Glu Ile Gly Leu Gly Ser Pro Leu Val Pro Met Gly Asp Asp
340 345 350

Asp Gly Gly Pro Glu Ala Thr Gly Asn Gln Ser Ala Trp Met Gly His
355 360 365

Gly Ser Pro Leu Ser His Ser Ser Lys Gly Thr Pro Leu Leu Ala Thr
370 375 380

Glu Phe Arg Gly Leu Ala Leu Lys Tyr Leu Leu Thr Pro Val Asn Phe
385 390 395 400

Phe Gln Ala Gly Tyr Ser Ser Arg Arg Phe Asp Gln Gln Val Glu Lys
405 410 415

Tyr His Lys Pro Cys
420

<210> 68

<211> 307

<212> PRT

<213> Homo sapiens

<400> 68

Met Asn Gly Val Leu Ile Pro His Thr Pro Ile Ala Val Asp Phe Trp
1 5 10 15

Ser Leu Arg Arg Ala Gly Thr Ala Arg Leu Phe Phe Leu Ser His Met
20 25 30

His Ser Asp His Thr Val Gly Leu Ser Ser Thr Trp Ala Arg Pro Leu
35 40 45

Tyr Cys Ser Pro Ile Thr Ala His Leu Leu His Arg His Leu Gln Val
50 55 60

Ser Lys Gln Trp Ile Gln Ala Leu Glu Val Gly Glu Ser His Val Leu
65 70 75 80

```
Pro Leu Asp Glu Ile Gly Gln Glu Thr Met Thr Val Thr Leu Leu Asp
                85                  90                  95

Ala Asn His Cys Pro Gly Ser Val Met Phe Leu Phe Glu Gly Tyr Phe
            100                 105                 110

Gly Thr Ile Leu Tyr Thr Gly Asp Phe Arg Tyr Thr Pro Ser Met Leu
        115                 120                 125

Lys Glu Pro Ala Leu Thr Leu Gly Lys Gln Ile His Thr Leu Tyr Leu
    130                 135                 140

Asp Asn Thr Asn Cys Asn Pro Ala Leu Val Leu Pro Ser Arg Gln Glu
145                 150                 155                 160

Ala Ala His Gln Ile Val Gln Leu Ile Arg Lys His Pro Gln His Asn
                165                 170                 175

Ile Lys Ile Gly Leu Tyr Ser Leu Gly Lys Glu Ser Leu Leu Glu Gln
            180                 185                 190

Leu Ala Leu Glu Phe Gln Thr Trp Val Val Leu Ser Pro Arg Arg Leu
        195                 200                 205

Glu Leu Val Gln Leu Leu Gly Leu Ala Asp Val Phe Thr Val Glu Glu
    210                 215                 220

Lys Ala Gly Arg Ile His Ala Val Asp His Met Glu Ile Cys His Ser
225                 230                 235                 240

Asn Met Leu Arg Trp Asn Gln Thr His Pro Thr Ile Ala Ile Leu Pro
                245                 250                 255

Thr Ser Arg Lys Ile His Ser Ser His Pro Asp Ile His Val Ile Pro
            260                 265                 270

Tyr Ser Asp His Ser Ser Tyr Ser Glu Leu Arg Ala Phe Val Ala Ala
        275                 280                 285

Leu Lys Pro Cys Gln Val Val Pro Ile Val Ser Arg Arg Pro Trp Glu
    290                 295                 300

Ala Phe Arg
305
```

<210> 69

<211> 336

<212> PRT

<213> Homo sapiens

<400> 69

```
Met Ala Thr Ala Leu Ser Glu Glu Glu Leu Asp Asn Glu Asp Tyr Tyr
1               5                   10                  15

Ser Leu Leu Asn Val Arg Arg Glu Ala Ser Ser Glu Glu Leu Lys Ala
            20                  25                  30
```

```
Ala Tyr Arg Arg Leu Cys Met Leu Tyr His Pro Asp Lys His Arg Asp
        35                  40                  45

Pro Glu Leu Lys Ser Gln Ala Glu Arg Leu Phe Asn Leu Val His Gln
    50                  55                  60

Ala Tyr Glu Val Leu Ser Asp Pro Gln Thr Arg Ala Ile Tyr Asp Ile
65                  70                  75                  80

Tyr Gly Lys Gly Gly Leu Glu Met Glu Gly Trp Glu Val Val Glu Arg
                85                  90                  95

Arg Arg Thr Pro Ala Glu Ile Arg Glu Glu Phe Glu Arg Leu Gln Arg
            100                 105                 110

Glu Arg Glu Glu Arg Arg Leu Gln Gln Arg Thr Asn Pro Lys Gly Thr
        115                 120                 125

Ile Ser Val Gly Val Asn Ala Thr Asp Leu Phe Asp Arg Tyr Asp Glu
    130                 135                 140

Glu Tyr Glu Asp Val Ser Gly Ser Ser Phe Pro Gln Ile Glu Ile Asn
145                 150                 155                 160

Lys Met His Ile Ser Gln Ser Ile Glu Ala Pro Leu Thr Ala Thr Asp
                165                 170                 175

Thr Ala Ile Leu Ser Gly Ser Leu Ser Thr Gln Asn Gly Asn Gly Gly
            180                 185                 190

Gly Ser Ile Asn Phe Ala Leu Arg Arg Val Thr Ser Val Lys Gly Trp
        195                 200                 205

Gly Glu Leu Glu Phe Gly Ala Gly Asp Leu Gln Gly Pro Leu Phe Gly
    210                 215                 220

Leu Lys Leu Phe Arg Asn Leu Thr Pro Arg Cys Phe Val Thr Thr Asn
225                 230                 235                 240

Cys Ala Leu Gln Phe Ser Ser Arg Gly Ile Arg Pro Gly Leu Thr Thr
                245                 250                 255

Val Leu Ala Arg Asn Leu Asp Lys Asn Thr Val Gly Tyr Leu Gln Trp
            260                 265                 270

Arg Trp Gly Ile Gln Ser Ala Met Asn Thr Ser Ile Val Arg Asp Thr
        275                 280                 285

Lys Thr Ser His Phe Thr Val Ala Leu Gln Leu Gly Ile Pro His Ser
    290                 295                 300

Phe Ala Leu Ile Ser Tyr Gln His Lys Phe Gln Asp Asp Asp Gln Thr
305                 310                 315                 320

Arg Val Lys Gly Ser Leu Lys Ala Gly Phe Phe Gly Thr Val Val Glu
                325                 330                 335
```

<210> 70

<211> 559

<212> PRT

<213> Homo sapiens

<400> 70

```
Met Ala Thr Ala Leu Ser Glu Glu Glu Leu Asp Asn Glu Asp Tyr Tyr
1               5                  10                  15

Ser Leu Leu Asn Val Arg Arg Glu Ala Ser Ser Glu Glu Leu Lys Ala
            20                  25                  30

Ala Tyr Arg Arg Leu Cys Met Leu Tyr His Pro Asp Lys His Arg Asp
        35                  40                  45

Pro Glu Leu Lys Ser Gln Ala Glu Arg Leu Phe Asn Leu Val His Gln
    50                  55                  60

Ala Tyr Glu Val Leu Ser Asp Pro Gln Thr Arg Ala Ile Tyr Asp Ile
65                  70                  75                  80

Tyr Gly Lys Arg Gly Leu Glu Met Glu Gly Trp Glu Val Val Glu Arg
                85                  90                  95

Arg Arg Thr Pro Ala Glu Ile Arg Glu Glu Phe Glu Arg Leu Gln Arg
            100                 105                 110

Glu Arg Glu Glu Arg Arg Leu Gln Gln Arg Thr Asn Pro Lys Gly Thr
        115                 120                 125

Ile Ser Val Gly Val Asp Ala Thr Asp Leu Phe Asp Arg Tyr Asp Glu
    130                 135                 140

Glu Tyr Glu Asp Val Ser Gly Ser Ser Phe Pro Gln Ile Glu Ile Asn
145                 150                 155                 160

Lys Met His Ile Ser Gln Ser Ile Glu Ala Pro Leu Thr Ala Thr Asp
                165                 170                 175

Thr Ala Ile Leu Ser Gly Ser Leu Ser Thr Gln Asn Gly Asn Gly Gly
            180                 185                 190

Gly Ser Ile Asn Phe Ala Leu Arg Arg Val Thr Ser Ala Lys Gly Trp
        195                 200                 205

Gly Glu Leu Glu Phe Gly Ala Gly Asp Leu Gln Gly Pro Leu Phe Gly
    210                 215                 220

Leu Lys Leu Phe Arg Asn Leu Thr Pro Arg Cys Phe Val Thr Thr Asn
225                 230                 235                 240

Cys Ala Leu Gln Phe Ser Ser Arg Gly Ile Arg Pro Gly Leu Thr Thr
                245                 250                 255

Val Leu Ala Arg Asn Leu Asp Lys Asn Thr Val Gly Tyr Leu Gln Trp
            260                 265                 270

Arg Trp Gly Ile Gln Ser Ala Met Asn Thr Ser Ile Val Arg Asp Thr
        275                 280                 285

Lys Thr Ser His Phe Thr Val Ala Leu Gln Leu Gly Ile Pro His Ser
    290                 295                 300
```

```
Phe Ala Leu Ile Ser Tyr Gln His Lys Phe Gln Asp Asp Asp Gln Thr
305                 310                 315                 320

Arg Val Lys Gly Ser Leu Lys Ala Gly Phe Phe Gly Thr Val Val Glu
                325                 330                 335

Tyr Gly Ala Glu Arg Lys Ile Ser Arg His Ser Val Leu Gly Ala Ala
            340                 345                 350

Val Ser Val Gly Val Pro Gln Gly Val Ser Leu Lys Val Lys Leu Asn
        355                 360                 365

Arg Ala Ser Gln Thr Tyr Phe Phe Pro Ile His Leu Thr Asp Gln Leu
    370                 375                 380

Leu Pro Ser Ala Met Phe Tyr Ala Thr Val Gly Pro Leu Val Val Tyr
385                 390                 395                 400

Phe Ala Met His Arg Leu Ile Ile Lys Pro Tyr Leu Arg Ala Gln Lys
                405                 410                 415

Glu Lys Glu Leu Glu Lys Gln Arg Glu Ser Ala Ala Thr Asp Val Leu
            420                 425                 430

Gln Lys Lys Gln Glu Ala Glu Ser Ala Val Arg Leu Met Gln Glu Ser
        435                 440                 445

Val Arg Arg Ile Ile Glu Ala Glu Glu Ser Arg Met Gly Leu Ile Ile
    450                 455                 460

Val Asn Ala Trp Tyr Gly Lys Phe Val Asn Asp Lys Ser Arg Lys Ser
465                 470                 475                 480

Glu Lys Val Lys Val Ile Asp Val Thr Val Pro Leu Gln Cys Leu Val
                485                 490                 495

Lys Asp Ser Lys Leu Ile Leu Thr Glu Ala Ser Lys Ala Gly Leu Pro
            500                 505                 510

Gly Phe Tyr Asp Pro Cys Val Gly Glu Glu Lys Asn Leu Lys Val Leu
        515                 520                 525

Tyr Gln Phe Arg Gly Val Leu His Gln Val Met Val Leu Asp Ser Glu
    530                 535                 540

Ala Leu Arg Ile Pro Lys Gln Ser His Arg Ile Asp Thr Asp Gly
545                 550                 555

<210>  71

<211>  103

<212>  PRT

<213>  Homo sapiens


<400>  71

Asp Ser Val Ser Lys Lys Lys Lys Lys Lys Glu Ile His Lys Val Val
1               5                   10                  15
```

Glu Arg Arg Arg Thr Pro Ala Glu Ile Arg Glu Glu Phe Glu Arg Leu
20 25 30

Gln Arg Glu Arg Glu Glu Arg Arg Leu Gln Gln Arg Thr Asn Pro Lys
35 40 45

Gly Thr Ile Ser Val Gly Val Asp Ala Thr Asp Leu Phe Asp Arg Tyr
50 55 60

Asp Glu Glu Tyr Glu Asp Val Ser Gly Ser Ser Phe Pro Gln Ile Glu
65 70 75 80

Ile Asn Lys Met His Ile Ser Gln Ser Ile Glu Ala Pro Leu Thr Ala
85 90 95

Thr Asp Thr Ala Ile Leu Ser
100

<210> 72

<211> 414

<212> PRT

<213> Homo sapiens

<400> 72

Met Ser Glu Tyr Ile Arg Val Thr Glu Asp Glu Asn Asp Glu Pro Ile
1 5 10 15

Glu Ile Pro Ser Glu Asp Asp Gly Thr Val Leu Leu Ser Thr Val Thr
20 25 30

Ala Gln Phe Pro Gly Ala Cys Gly Leu Arg Tyr Arg Asn Pro Val Ser
35 40 45

Gln Cys Met Arg Gly Val Arg Leu Val Glu Gly Ile Leu His Ala Pro
50 55 60

Asp Ala Gly Trp Gly Asn Leu Val Tyr Val Val Asn Tyr Pro Lys Asp
65 70 75 80

Asn Lys Arg Lys Met Asp Glu Thr Asp Ala Ser Ser Ala Val Lys Val
85 90 95

Lys Arg Ala Val Gln Lys Thr Ser Asp Leu Ile Val Leu Gly Leu Pro
100 105 110

Trp Lys Thr Thr Glu Gln Asp Leu Lys Glu Tyr Phe Ser Thr Phe Gly
115 120 125

Glu Val Leu Met Val Gln Val Lys Lys Asp Leu Lys Thr Gly His Ser
130 135 140

Lys Gly Phe Gly Phe Val Arg Phe Thr Glu Tyr Glu Thr Gln Val Lys
145 150 155 160

Val Met Ser Gln Arg His Met Ile Asp Gly Arg Trp Cys Asp Cys Lys
165 170 175

Leu Pro Asn Ser Lys Gln Ser Gln Asp Glu Pro Leu Arg Ser Arg Lys
180 185 190

Val Phe Val Gly Arg Cys Thr Glu Asp Met Thr Glu Asp Glu Leu Arg
195 200 205

Glu Phe Phe Ser Gln Tyr Gly Asp Val Met Asp Val Phe Ile Pro Lys
210 215 220

Pro Phe Arg Ala Phe Ala Phe Val Thr Phe Ala Asp Asp Gln Ile Ala
225 230 235 240

Gln Ser Leu Cys Gly Glu Asp Leu Ile Ile Lys Gly Ile Ser Val His
245 250 255

Ile Ser Asn Ala Glu Pro Lys His Asn Ser Asn Arg Gln Leu Glu Arg
260 265 270

Ser Gly Arg Phe Gly Gly Asn Pro Gly Gly Phe Gly Asn Gln Gly Gly
275 280 285

Phe Gly Asn Ser Arg Gly Gly Gly Ala Gly Leu Gly Asn Asn Gln Gly
290 295 300

Ser Asn Met Gly Gly Gly Met Asn Phe Gly Ala Phe Ser Ile Asn Pro
305 310 315 320

Ala Met Met Ala Ala Ala Gln Ala Ala Leu Gln Ser Ser Trp Gly Met
325 330 335

Met Gly Met Leu Ala Ser Gln Gln Asn Gln Ser Gly Pro Ser Gly Asn
340 345 350

Asn Gln Asn Gln Gly Asn Met Gln Arg Glu Pro Asn Gln Ala Phe Gly
355 360 365

Ser Gly Asn Asn Ser Tyr Ser Gly Ser Asn Ser Gly Ala Ala Ile Gly
370 375 380

Trp Gly Ser Ala Ser Asn Ala Gly Ser Gly Ser Gly Phe Asn Gly Gly
385 390 395 400

Phe Gly Ser Ser Met Asp Ser Lys Ser Ser Gly Trp Gly Met
405 410

<210> 73

<211> 173

<212> PRT

<213> Homo sapiens

<400> 73

Thr Arg Ala His Gln Glu Ser Ala Glu Pro Lys Tyr Leu Pro His Lys
1 5 10 15

Thr Cys Asn Glu Ile Ile Val Pro Lys Ala Pro Ser His Lys Thr Ile
20 25 30

```
Gln Glu Thr Pro His Ser Glu Asp Tyr Ser Ile Glu Ile Asn Gln Glu
        35                  40                  45

Thr Pro Gly Ser Glu Lys Tyr Ser Pro Glu Thr Tyr Gln Glu Ile Pro
    50                  55                  60

Gly Leu Glu Glu Tyr Ser Pro Glu Ile Tyr Gln Glu Thr Ser Gln Leu
65                  70                  75                  80

Glu Glu Tyr Ser Pro Glu Ile Tyr Gln Glu Thr Pro Gly Pro Glu Asp
                85                  90                  95

Leu Ser Thr Glu Thr Tyr Lys Asn Lys Asp Val Pro Lys Glu Cys Phe
            100                 105                 110

Pro Glu Pro His Gln Glu Thr Gly Gly Pro Gln Gly Gln Asp Pro Lys
        115                 120                 125

Ala His Gln Glu Asp Ala Lys Asp Ala Tyr Thr Phe Pro Gln Glu Met
    130                 135                 140

Lys Glu Lys Pro Lys Glu Glu Pro Gly Ile Pro Ala Ile Leu Asn Glu
145                 150                 155                 160

Ser His Pro Glu Asn Asp Val Tyr Ser Tyr Val Leu Phe
                165                 170
```

<210> 74

<211> 484

<212> PRT

<213> Homo sapiens

<400> 74

```
Met Asp Leu Gly Lys Asp Gln Ser His Leu Lys His His Gln Thr Pro
1               5                   10                  15

Asp Pro His Gln Glu Glu Asn His Ser Pro Glu Val Ile Gly Thr Trp
            20                  25                  30

Ser Leu Arg Asn Arg Glu Leu Leu Arg Lys Arg Lys Ala Glu Val His
        35                  40                  45

Glu Lys Glu Thr Ser Gln Trp Leu Phe Gly Glu Gln Lys Lys Arg Lys
    50                  55                  60

Gln Gln Arg Thr Gly Lys Gly Asn Arg Arg Gly Arg Lys Arg Gln Gln
65                  70                  75                  80

Asn Thr Glu Leu Lys Val Glu Pro Gln Pro Gln Ile Glu Lys Glu Ile
                85                  90                  95

Val Glu Lys Ala Leu Ala Pro Ile Glu Lys Lys Thr Glu Pro Pro Gly
            100                 105                 110

Ser Ile Thr Lys Val Phe Pro Ser Val Ala Ser Pro Gln Lys Val Val
        115                 120                 125
```

```
Pro Glu Glu His Phe Ser Glu Ile Cys Gln Glu Ser Asn Ile Tyr Gln
    130                 135                 140

Glu Asn Phe Ser Glu Tyr Gln Glu Ile Ala Val Gln Asn His Ser Ser
145                 150                 155                 160

Glu Thr Cys Gln His Val Ser Glu Pro Glu Asp Leu Ser Pro Lys Met
                165                 170                 175

Tyr Gln Glu Ile Ser Val Leu Gln Asp Asn Ser Ser Lys Ile Cys Gln
            180                 185                 190

Asp Met Lys Glu Pro Glu Asp Asn Ser Pro Asn Thr Cys Gln Val Ile
        195                 200                 205

Ser Val Ile Gln Asp His Pro Phe Lys Met Tyr Gln Asp Met Ala Lys
    210                 215                 220

Arg Glu Asp Leu Ala Pro Lys Met Cys Gln Glu Ala Ala Val Pro Lys
225                 230                 235                 240

Ile Leu Pro Cys Pro Thr Ser Glu Asp Thr Ala Asp Leu Ala Gly Cys
                245                 250                 255

Ser Leu Gln Ala Tyr Pro Lys Pro Asp Val Pro Lys Gly Tyr Ile Leu
            260                 265                 270

Asp Thr Asp Gln Asn Pro Ala Glu Pro Glu Glu Tyr Asn Glu Thr Asp
        275                 280                 285

Gln Gly Ile Ala Glu Thr Glu Gly Leu Phe Pro Lys Ile Gln Glu Ile
    290                 295                 300

Ala Glu Pro Lys Asp Leu Ser Thr Lys Thr His Gln Glu Ser Ala Glu
305                 310                 315                 320

Pro Lys Tyr Leu Pro His Lys Thr Cys Asn Glu Ile Ile Val Pro Lys
                325                 330                 335

Ala Pro Ser His Lys Thr Ile Gln Glu Thr Pro His Ser Glu Asp Tyr
            340                 345                 350

Ser Ile Glu Ile Asn Gln Glu Thr Pro Gly Ser Glu Lys Tyr Ser Pro
        355                 360                 365

Glu Thr Tyr Gln Glu Ile Pro Gly Leu Glu Glu Tyr Ser Pro Glu Ile
    370                 375                 380

Tyr Gln Glu Thr Ser Gln Leu Glu Glu Tyr Ser Pro Glu Ile Tyr Gln
385                 390                 395                 400

Glu Thr Pro Gly Pro Glu Asp Leu Ser Thr Glu Thr Tyr Lys Asn Lys
                405                 410                 415

Asp Val Pro Lys Glu Cys Phe Pro Glu Pro His Gln Glu Thr Gly Gly
            420                 425                 430

Pro Gln Gly Gln Asp Pro Lys Ala His Gln Glu Asp Ala Lys Asp Ala
        435                 440                 445

Tyr Thr Phe Pro Gln Glu Met Lys Glu Lys Pro Lys Glu Glu Pro Gly
    450                 455                 460
```

Ile Pro Ala Ile Leu Asn Glu Ser His Pro Glu Asn Asp Val Tyr Ser
465 470 475 480

Tyr Val Leu Phe

<210> 75

<211> 484

<212> PRT

<213> Homo sapiens

<400> 75

Met Asp Leu Gly Lys Asp Gln Ser His Leu Lys His His Gln Thr Pro
1 5 10 15

Asp Pro His Gln Glu Glu Asn His Ser Pro Glu Val Ile Gly Thr Trp
20 25 30

Ser Leu Arg Asn Arg Glu Leu Leu Arg Lys Arg Lys Ala Glu Val His
35 40 45

Glu Lys Glu Thr Ser Gln Trp Leu Phe Gly Glu Gln Lys Lys Arg Lys
50 55 60

Gln Gln Arg Thr Gly Lys Gly Asn Arg Arg Gly Arg Lys Arg Gln Gln
65 70 75 80

Asn Thr Glu Leu Lys Val Glu Pro Gln Pro Gln Ile Glu Lys Glu Ile
85 90 95

Val Glu Lys Ala Leu Ala Pro Ile Glu Lys Lys Thr Glu Pro Pro Gly
100 105 110

Ser Ile Thr Lys Val Phe Pro Ser Val Ala Ser Pro Gln Lys Val Val
115 120 125

Pro Glu Glu His Phe Ser Glu Ile Cys Gln Glu Ser Asn Ile Tyr Gln
130 135 140

Glu Asn Phe Ser Glu Tyr Gln Glu Ile Ala Val Gln Asn His Ser Ser
145 150 155 160

Glu Thr Cys Gln His Val Ser Glu Pro Glu Asp Leu Ser Pro Lys Met
165 170 175

Tyr Gln Glu Ile Ser Val Leu Gln Asp Asn Ser Ser Lys Ile Cys Gln
180 185 190

Asp Met Lys Glu Pro Glu Asp Asn Ser Pro Asn Thr Cys Gln Val Ile
195 200 205

Ser Val Ile Gln Asp His Pro Phe Lys Met Tyr Gln Asp Met Ala Lys
210 215 220

Arg Glu Asp Leu Ala Pro Lys Met Cys Gln Glu Ala Ala Val Pro Lys
225 230 235 240

```
Ile Leu Pro Cys Pro Thr Ser Glu Asp Thr Ala Asp Leu Ala Gly Cys
                245                 250                 255

Ser Leu Gln Ala Tyr Pro Lys Pro Asp Val Pro Lys Gly Tyr Ile Leu
            260                 265                 270

Asp Thr Asp Gln Asn Pro Ala Glu Pro Glu Glu Tyr Asn Glu Thr Asp
        275                 280                 285

Gln Gly Ile Ala Glu Thr Glu Gly Leu Phe Pro Lys Ile Gln Glu Ile
    290                 295                 300

Ala Glu Pro Lys Asp Leu Ser Thr Lys Thr His Gln Glu Ser Ala Glu
305                 310                 315                 320

Pro Lys Tyr Leu Pro His Lys Thr Cys Asn Glu Ile Ile Val Pro Lys
                325                 330                 335

Ala Pro Ser His Lys Thr Ile Gln Glu Thr Pro His Ser Glu Asp Tyr
            340                 345                 350

Ser Ile Glu Ile Asn Gln Glu Thr Pro Gly Ser Glu Lys Tyr Ser Pro
        355                 360                 365

Glu Thr Tyr Gln Glu Ile Pro Gly Leu Glu Glu Tyr Ser Pro Glu Ile
    370                 375                 380

Tyr Gln Glu Thr Ser Gln Leu Glu Glu Tyr Ser Pro Glu Ile Tyr Gln
385                 390                 395                 400

Glu Thr Pro Gly Pro Glu Asp Leu Ser Thr Glu Thr Tyr Lys Asn Lys
                405                 410                 415

Asp Val Pro Lys Glu Cys Phe Pro Glu Pro His Gln Glu Thr Gly Gly
            420                 425                 430

Pro Gln Gly Gln Asp Pro Lys Ala His Gln Glu Asp Ala Lys Asp Ala
        435                 440                 445

Tyr Thr Phe Pro Gln Glu Met Lys Glu Lys Pro Lys Glu Glu Pro Gly
    450                 455                 460

Ile Pro Ala Ile Leu Asn Glu Ser His Pro Glu Asn Asp Val Tyr Ser
465                 470                 475                 480

Tyr Val Leu Phe
```

<210> 76

<211> 331

<212> PRT

<213> Homo sapiens

<400> 76

```
Met Trp Leu Trp Glu Asp Gln Gly Gly Leu Leu Gly Pro Phe Ser Phe
1               5                   10                  15
```

```
Leu Leu Leu Val Leu Leu Leu Val Thr Arg Ser Pro Val Asn Ala Cys
            20                  25                  30

Leu Leu Thr Gly Ser Leu Phe Val Leu Leu Arg Val Phe Ser Phe Glu
        35                  40                  45

Pro Val Pro Ser Cys Arg Ala Leu Gln Val Leu Lys Pro Arg Asp Arg
    50                  55                  60

Ile Ser Ala Ile Ala His Arg Gly Gly Ser His Asp Ala Pro Glu Asn
65                  70                  75                  80

Thr Leu Ala Ala Ile Arg Gln Ala Ala Lys Asn Gly Ala Thr Gly Val
                85                  90                  95

Glu Leu Asp Ile Glu Phe Thr Ser Asp Gly Ile Pro Val Leu Met His
            100                 105                 110

Asp Asn Thr Val Asp Arg Thr Thr Asp Gly Thr Gly Arg Leu Cys Asp
        115                 120                 125

Leu Thr Phe Glu Gln Ile Arg Lys Leu Asn Pro Ala Ala Asn His Arg
    130                 135                 140

Leu Arg Asn Asp Phe Pro Asp Glu Lys Ile Pro Thr Leu Arg Glu Ala
145                 150                 155                 160

Val Ala Glu Cys Leu Asn His Asn Leu Thr Ile Phe Phe Asp Val Lys
                165                 170                 175

Gly His Ala His Lys Ala Thr Glu Ala Leu Lys Lys Met Tyr Met Glu
            180                 185                 190

Phe Pro Gln Leu Tyr Asn Asn Ser Val Val Cys Ser Phe Leu Pro Glu
        195                 200                 205

Val Ile Tyr Lys Met Arg Gln Thr Asp Arg Asp Val Ile Thr Ala Leu
    210                 215                 220

Thr His Arg Pro Trp Ser Leu Ser His Thr Gly Asp Gly Lys Pro Arg
225                 230                 235                 240

Tyr Asp Thr Phe Trp Lys His Phe Ile Phe Val Met Met Asp Ile Leu
                245                 250                 255

Leu Asp Trp Ser Met His Asn Ile Leu Trp Tyr Leu Cys Gly Ile Ser
            260                 265                 270

Ala Phe Leu Met Gln Lys Asp Phe Val Ser Pro Ala Tyr Leu Lys Lys
        275                 280                 285

Trp Ser Ala Lys Gly Ile Gln Val Val Gly Trp Thr Val Asn Thr Phe
    290                 295                 300

Asp Glu Lys Ser Tyr Tyr Glu Ser His Leu Gly Ser Ser Tyr Ile Thr
305                 310                 315                 320

Asp Ser Met Val Glu Asp Cys Glu Pro His Phe
                325                 330
```

<210> 77

<211> 188

<212> PRT

<213> Homo sapiens

<400> 77

```
Phe Thr Thr Gly Cys His Tyr Trp Glu Val Tyr Val Gly Asp Lys Thr
1               5                  10                  15

Lys Trp Ile Leu Gly Val Cys Ser Glu Ser Val Ser Arg Lys Gly Lys
            20                  25                  30

Val Thr Ala Ser Pro Ala Asn Gly His Trp Leu Leu Arg Gln Ser Arg
        35                  40                  45

Gly Asn Glu Tyr Glu Ala Leu Thr Ser Pro Gln Thr Ser Phe Arg Leu
    50                  55                  60

Lys Glu Pro Pro Arg Cys Val Gly Ile Phe Leu Asp Tyr Glu Ala Gly
65                  70                  75                  80

Val Ile Ser Phe Tyr Asn Val Thr Asn Lys Ser His Ile Phe Thr Phe
                85                  90                  95

Thr His Asn Phe Ser Gly Pro Leu Arg Pro Phe Phe Glu Pro Cys Leu
            100                 105                 110

His Asp Gly Gly Lys Asn Thr Ala Pro Leu Val Ile Cys Ser Glu Leu
        115                 120                 125

His Lys Ser Glu Glu Ser Ile Val Pro Arg Pro Glu Gly Lys Gly His
    130                 135                 140

Ala Asn Gly Asp Val Ser Leu Lys Val Asn Ser Ser Leu Leu Pro Pro
145                 150                 155                 160

Lys Ala Pro Glu Leu Lys Asp Ile Ile Leu Ser Leu Pro Pro Asp Leu
                165                 170                 175

Gly Pro Ala Leu Gln Glu Leu Lys Ala Pro Ser Phe
            180                 185
```

<210> 78

<211> 475

<212> PRT

<213> Homo sapiens

<400> 78

```
Met Glu Met Ala Ser Ser Ala Gly Ser Trp Leu Ser Gly Cys Leu Ile
1               5                  10                  15

Pro Leu Val Phe Leu Arg Leu Ser Val His Val Ser Gly His Ala Gly
            20                  25                  30

Asp Ala Gly Lys Phe His Val Ala Leu Leu Gly Gly Thr Ala Glu Leu
```

```
        35                  40                  45

Leu Cys Pro Leu Ser Leu Trp Pro Gly Thr Val Pro Lys Glu Val Arg
    50                  55                  60

Trp Leu Arg Ser Pro Phe Pro Gln Arg Ser Gln Ala Val His Ile Phe
65                  70                  75                  80

Arg Asp Gly Lys Asp Gln Asp Glu Asp Leu Met Pro Glu Tyr Lys Gly
                85                  90                  95

Arg Thr Val Leu Val Arg Asp Ala Gln Glu Gly Ser Val Thr Leu Gln
            100                 105                 110

Ile Leu Asp Val Arg Leu Glu Asp Gln Gly Ser Tyr Arg Cys Leu Ile
        115                 120                 125

Gln Val Gly Asn Leu Ser Lys Glu Asp Thr Val Ile Leu Gln Val Ala
    130                 135                 140

Ala Pro Ser Val Gly Ser Leu Ser Pro Ser Ala Val Ala Leu Ala Val
145                 150                 155                 160

Ile Leu Pro Val Leu Val Leu Leu Ile Met Val Cys Leu Cys Leu Ile
                165                 170                 175

Trp Lys Gln Arg Arg Ala Lys Glu Lys Leu Leu Tyr Glu His Val Thr
            180                 185                 190

Glu Val Asp Asn Leu Leu Ser Asp His Ala Lys Glu Lys Gly Lys Leu
        195                 200                 205

His Lys Ala Val Lys Lys Leu Arg Ser Glu Leu Lys Leu Lys Arg Ala
    210                 215                 220

Ala Ala Asn Ser Gly Trp Arg Arg Ala Arg Leu His Phe Val Ala Val
225                 230                 235                 240

Thr Leu Asp Pro Asp Thr Ala His Pro Lys Leu Ile Leu Ser Glu Asp
                245                 250                 255

Gln Arg Cys Val Arg Leu Gly Asp Arg Arg Gln Pro Val Pro Asp Asn
            260                 265                 270

Pro Gln Arg Phe Asp Phe Val Val Ser Ile Leu Gly Ser Glu Tyr Phe
        275                 280                 285

Thr Thr Gly Cys His Tyr Trp Glu Val Tyr Val Gly Asp Lys Thr Lys
    290                 295                 300

Trp Ile Leu Gly Val Cys Ser Glu Ser Val Ser Arg Lys Gly Lys Val
305                 310                 315                 320

Thr Ala Ser Pro Ala Asn Gly His Trp Leu Leu Arg Gln Ser Arg Gly
                325                 330                 335

Asn Glu Tyr Glu Ala Leu Thr Ser Pro Gln Thr Ser Phe Arg Leu Lys
            340                 345                 350

Glu Pro Pro Arg Cys Val Gly Ile Phe Leu Asp Tyr Glu Ala Gly Val
        355                 360                 365

Ile Ser Phe Tyr Asn Val Thr Asn Lys Ser His Ile Phe Thr Phe Thr
```

```
    370                 375                 380

His Asn Phe Ser Gly Pro Leu Arg Pro Phe Phe Glu Pro Cys Leu His
385                 390                 395                 400

Asp Gly Gly Lys Asn Thr Ala Pro Leu Val Ile Cys Ser Glu Leu His
                405                 410                 415

Lys Ser Glu Glu Ser Ile Val Pro Arg Pro Glu Gly Lys Gly His Ala
            420                 425                 430

Asn Gly Asp Val Ser Leu Lys Val Asn Ser Ser Leu Leu Pro Pro Lys
        435                 440                 445

Ala Pro Glu Leu Lys Asp Ile Ile Leu Ser Leu Pro Pro Asp Leu Gly
    450                 455                 460

Pro Ala Leu Gln Glu Leu Lys Ala Pro Ser Phe
465                 470                 475
```