

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
1 June 2006 (01.06.2006)

PCT

(10) International Publication Number
WO 2006/058043 A2

(51) International Patent Classification:
A61K 39/00 (2006.01)

(21) International Application Number:

PCT/US2005/042399

(22) International Filing Date:

23 November 2005 (23.11.2005)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/630,310	23 November 2004 (23.11.2004)	US
60/630,507	23 November 2004 (23.11.2004)	US
60/635,885	14 December 2004 (14.12.2004)	US
60/725,871	11 October 2005 (11.10.2005)	US

(71) Applicant (for all designated States except US): **UNIVERSITY OF VIRGINIA PATENT FOUNDATION** [US/US]; 250 West Main Street, Suite 300, Charlottesville, VA 22902 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **CHERTIHIN, Olga** [US/US]; 446 Peacock Drive, Charlottesville, VA 22903 (US). **HE, Wei** [CN/US]; 285 Peyton Court, Apt. 2, Charlottesville, VA 22903 (US). **HERR, John, C.** [US/US]; 895 Old Ballard Road, Charlottesville, VA 22901 (US).

(74) Agent: **SPARKS, Rodney, L.**; UNIVERSITY OF VIRGINIA PATENT FOUNDATION, 250 West Main Street, Suite 300, Charlottesville, Virginia 22902 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

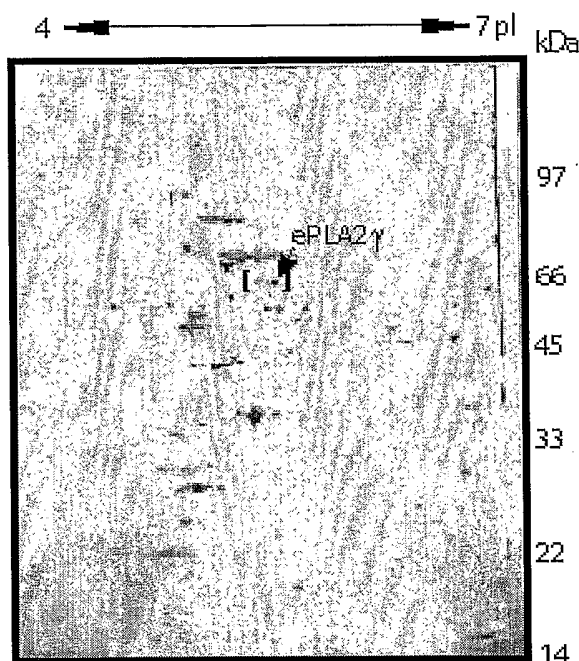
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (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.

(54) Title: CONTRACEPTIVE VACCINES FOR DOGS AND CATS BASED ON EGG MEMBRANE ANTIGENS



(57) Abstract: The present invention provides novel ePLA2γ, MOP, and ePAD nucleic acid and amino acid sequences and compounds and methods for their use in modulating fertility, inducing an immune response, and inhibiting or diminishing conception. In one aspect, the invention provides a contraceptive vaccine comprising at least one novel egg protein, or homolog, fragment, or derivative thereof.

WO 2006/058043 A2

CONTRACEPTIVE VACCINES FOR DOGS AND CATS BASED ON EGG MEMBRANE ANTIGENS

CROSS REFERENCE TO RELATED APPLICATIONS

5 This application is entitled to priority pursuant to 35 U.S.C. § 119(e) to U.S. Provisional Patent Application Serial Nos. 60/630,507, filed November 23, 2004, 60/630,310, filed November 23, 2004, 60/635,885, filed December 14, 2004, and 60/725,871, filed October 11, 2005, the entire disclosures of which are herein incorporated by reference.

10

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

 This invention was made with United States Government support under Grant No. HD38353, awarded by the National Institutes of Health. The United States Government
15 may have certain rights in the invention.

FIELD OF THE INVENTION

 The invention relates to novel egg genes and to novel egg nucleic acids, nucleic acid sequences, proteins, and protein sequences. The invention further relates to methods
20 of regulating the novel egg nucleic acids and proteins. The invention relates to novel ePAD, MOP, and ePLA2 γ genes and to novel nucleic acids, nucleic acid sequences, proteins, and protein sequences. The invention further relates to methods of inducing an immune response with one or more of these proteins. The invention also relates to contraceptive vaccines comprising one or more egg or ovary proteins.

25

BACKGROUND

 During growth, a pool of poorly characterized gene products required for oocyte maturation, fertilization, and preimplantation development, accumulate in the oocyte. Primarily, three different kinds of phospholipase A2's have been identified and cloned in
30 mammals: calcium-independent cytosolic phospholipase A2 (iPLA2), secretory phospholipase A2 (sPLA2), and cytosolic phospholipase A2 (cPLA2). The Ca²⁺

independent PLA2 (iPLA2) family contains two enzymes VIA (iPLA2 β) and VIB (iPLA2 γ) and may play a major role in phospholipid remodeling. The secretory PLA2 (sPLA2) family consists of low-molecular weight Ca²⁺-requiring secretory enzymes that have been implicated in a number of biological processes, such as regulation of eicosanoid production, inflammation, and host defense (Kudo et al., 2002, Prostaglandins Other Lipid Mediat. 68-69, 3-58). The cytosolic (cPLA2) family, known as Group IV according to the new nomenclature, consist of three isozymes, IVA (cPLA α), IVB (cPLA β) and IVC (cPLA γ), with molecular masses of 85, 110, and 60 KDa respectively (Kudo et al., 2002, Prostaglandins Other Lipid Mediat. 68-69, 3-58). The most studied member of this group is cPLA α , which has no detectable sequence homology with the secreted forms of PLA2 and contains two domains, a calcium-dependent lipid domain (CaLB) and a catalytic domain (Clark et al., 1991, Cell, 65:1043-1051). cPLA2 γ , the less studied member of this group, shows calcium-independent catalytic activity, and is constitutively associated with the lipid bilayer via a prenylation moiety located at the C-terminus. This prenylation site is critical for membrane localization and thus for cPLA2 γ function (Murakami et al., 2003, Biochem. J. 372:695-702). Phospholipase A2 enzymes catalyze the hydrolysis of the sn-2 position of membrane glycerophospholipids, leading to the production of free fatty acids and lysophospholipids. If arachidonic acid is the esterified fatty acid, then downstream metabolic enzymes, such as cyclooxygenases, convert this fatty acid to various bioactive lipophilic compounds called eicosanoids, which include prostaglandins (PGs) and leukotrienes (LTs).

Peptidylarginine deiminases (PADs) are a family of proteins that are widely distributed in mammals. The protein catalyses the conversion of arginine to citrulline. PAD expression is closely associated with different stages of development and cell differentiation. To date, five members of the human PAD family, PAD1–PAD4 and ePAD (PAD6), have been characterized by cDNA cloning. PAD catalyzed citrullinization of target proteins such as keratins and the expression of PAD appears related to processes such as hair shaft formation and skin cornification. The citrullinization appears to strengthen keratin cytoskeletal filaments prior to compaction in the hair filament or cornified layer of skin epithelium.

The PAD genes share 50–55% sequence identity, with greater similarity in the C-terminal region of the proteins. The ePAD protein is mainly expressed in egg cells and embryos but has also recently been identified in human leukocytes. The protein ePAD is one of the most abundant in eggs, and the sequence is highly conserved between human and mouse, though the exact function remains unknown at this time. By electron microscopy, it has been determined that ePAD localizes to the egg cytoskeleton sheets, and thus may have a role in cytoskeletal stabilization or reorganization in the egg and early embryo. One feature that characterizes eggs is the high cytoplasmic to nuclear ratio. It is hypothesized that ePAD plays a role in stabilizing the egg cytoskeleton contributing to the egg's considerable volume.

There is a long felt need in the art to identify new egg genes and proteins, to determine their role in biological processes, and to develop methods to modify the expression and function of these egg genes and their expression products, and to use this knowledge to prepare and use contraceptive vaccines. The present invention satisfies these needs.

SUMMARY OF THE INVENTION

The present invention provides compositions and methods useful as contraceptive vaccines, as methods of inducing an immune response, and as methods of modulating fertility. The compositions comprise at least one mammalian egg protein, or a homolog, derivative, or fragment thereof, or at least one isolated nucleic acid comprising a nucleic acid sequence encoding an egg protein. The compositions are particularly useful as contraceptives for dogs and cats.

In one aspect, the composition comprises a cocktail or mixture of two or more different novel egg proteins or two or more different isolated nucleic acids comprising nucleic acid sequences encoding different egg proteins, which can be used as a contraceptive vaccine or to induce an immune response directed against the egg proteins.

The novel proteins of the invention comprise mouse ePLA2 γ (SEQ ID NO:5), mouse MOP31 (SEQ ID NO:7), cat ePAD (SEQ ID NO:9), and dog ePAD (SEQ ID NO:11). The corresponding nucleic acid sequences encoding the these proteins are

mouse ePLA2 γ (SEQ ID NO:6), mouse MOP31 (SEQ ID NO:8), cat ePAD (SEQ ID NO:10), and dog ePAD (SEQ ID NO:12)

In one embodiment, the invention provides a pharmaceutical composition comprising a pharmaceutically-acceptable carrier and at least one novel egg protein, or a homolog, fragment or derivative thereof, wherein said protein is capable of inducing an immune response useful for inhibiting conception in a subject. In one aspect, the invention provides a pharmaceutical composition, wherein said novel egg protein comprises an amino acid sequence selected from the group consisting of SEQ ID NOs:5, 7, 9, 11, 15, and 17. In another aspect, the invention provides a pharmaceutical composition further comprising at least one known egg protein, wherein said known egg protein is capable of inducing an immune response useful for contraception. In one aspect, the known egg protein has the amino acid sequence of SEQ ID NO:13, or a homolog, fragment or derivative thereof.

In one embodiment, the invention provides a pharmaceutical composition, wherein said pharmaceutical composition comprises at least two different proteins, or homologs, fragments, or derivatives thereof, wherein said at least two different proteins comprise an amino acid sequence selected from the group consisting of SEQ ID NOs:5, 7, 9, 11, 15, and 17. In one embodiment, the pharmaceutical further comprises at least one known egg protein.

In one embodiment, the invention provides a contraceptive vaccine, said vaccine comprising a pharmaceutical composition of the invention.

In one embodiment, the invention provides a method for inhibiting conception in a subject, said method comprising administering to said subject a pharmaceutical composition comprising a pharmaceutically-acceptable carrier and at least one novel egg protein, or a homolog, fragment or derivative thereof, wherein said protein is capable of inducing an immune response useful for inhibiting conception in a subject. In one aspect, the novel egg protein comprises an amino acid sequence selected from the group consisting of SEQ ID NOs:5, 7, 9, 11, 15, and 17.

In one embodiment, the subject is selected from the group consisting of cattle, pigs, horses, sheep, goats, birds, cats, and dogs. In one aspect, the pharmaceutical composition is administered via a route selected from the group consisting of topical,

oral, intramuscular, and intravenous. In one aspect, the composition is administered as a controlled-release formulation.

In one embodiment, least two different novel egg proteins, or homologs, fragments, or derivatives thereof are administered. In one aspect, the at least two
5 proteins, or homologs, fragments, or derivatives thereof are administered in a mixture of approximately equimolar concentrations. In one aspect, the at least two proteins, or homologs, fragments, or derivatives thereof are administered at levels of between about 0.001 mg/kg body weight and about 100 mg/kg body weight.

In one embodiment, at least three different novel egg proteins, or homologs,
10 fragments, or derivatives thereof are administered.

In one embodiment, the invention provides isolated nucleic acids comprising nucleic acid sequences encoding proteins of the invention. In another embodiment, the invention provides a nucleic acid sequence selected from the group of sequences having SEQ ID NOs:6, 8, 10, 12, and 16.

15 In one embodiment, the invention provides an antibody that specifically binds with a novel egg protein of the invention.

In one embodiment, the invention provides a method for inhibiting conception in a subject, comprising administering to the subject a pharmaceutical composition comprising at least one antibody directed against a protein of the invention. In one
20 aspect, the antibody inhibits the function of the novel egg protein. In one aspect, the antibody inhibits fertilization.

The invention further provides a kit for inhibiting conception in a subject.

ePLA2 γ

The present invention characterizes a novel maternal egg phospholipase A2 γ - like
25 (ePLA2 γ) protein (67.8 kDa, pI 5.7) that was originally cored from a Coomassie-stained two-dimensional (2D) gel of mouse egg proteins.

While there is considerable evidence in the literature suggesting that PLA2's are important for several stages of early embryonic development (Carnero et al., 1993; Riffo et al., 1997; Farber et al., 1999), the molecular identities of these phospholipases have not
30 been documented previously. The present invention provides ePLA2 γ , an egg and embryo abundant phospholipase A2 γ homologue, and discloses that this protein mainly

localizes to the inner leaflet of the plasma membrane and to the nucleoplasm in oocytes and preimplantation embryos. Without wishing to be bound by any particular theory, it is hypothesized that ePLA2 γ may be involved in lipid metabolism and membrane remodeling events during oogenesis and preimplantation development. The present disclosure supports the hypothesis that ePLA2 γ appears to concentrate at the nuclear envelope during germinal vesicle breakdown.

In one embodiment, ePLA2 γ contains a 75 bp 5' UTR, a 1791 bp open reading frame, and a 796 bp 3' UTR. Amino acid sequence analysis of ePLA2 γ revealed that the protein is ~56% identical to human cPLA2 γ . As with cPLA2 γ , ePLA2 γ contains a lipase consensus sequence and lacks a calcium binding domain. However, the ePLA2 γ protein is different from human cPLA2 γ in that it appears to be specifically expressed in the ovary, as demonstrated by northern blot analysis, and does not contain the myristoylation and prenylation lipid-anchoring motifs that are present in human cPLA2 γ .

ePLA2 γ appears to localize mainly to the inner leaflet of the plasma membrane and to the nucleoplasm in oocytes and preimplantation embryos. Given the abundance of this protein within oocytes and its continued expression until the blastocyst stage of development, it seems likely that ePLA2 γ may represent a previously uncharacterized maternal effect gene. Interestingly, in the oocyte, ePLA2 γ appears to aggregate in concentrated foci around the nuclear envelope during germinal vesicle breakdown suggesting a possible role for this putative egg-specific phospholipase in membrane remodeling.

The invention further provides MOP31 nucleic acid and amino acid sequences. MOP31 is a fragment of ePLA2 γ .

PAD Genes

The present application provides orthologous dog and cat ePAD genes. The invention provides the cat and dog homologues of the mouse ePAD gene, as well as fragments, modifications, derivatives and homologues thereof. In one embodiment, the invention provides a nucleic acid which comprises a nucleic acid sequence which encodes cat ePAD. The invention provides cat and dog ePAD proteins, as well as fragments, modifications, and derivatives thereof. In one aspect, the cat ePAD amino

acid sequence is provided. In one embodiment, the invention provides a nucleic acid which comprises a nucleic acid sequence which encodes dog ePAD. In one aspect, the dog ePAD amino acid sequence is provided. The invention also provides antibodies against the proteins and antisense oligonucleotides against the nucleic acids. Cloning the orthologous dog and cat ePAD genes, as well as identifying their function in the egg, are important steps in the development of a new contraceptive vaccine target for companion animals.

The present invention further provides nucleic acid and protein sequences for the novel egg protein MOP26.

BRIEF DESCRIPTION OF THE DRAWINGS

For the purpose of illustrating the invention, there are depicted in the drawings certain embodiments of the invention. However, the invention is not limited to the precise arrangements and instrumentalities of the embodiments depicted in the drawings.

Figure 1 represents an analysis of the location of ePLA2 γ (67.8 kDa, pI 5.7) on a Coomassie-stained two-dimensional electrophoretic gel of 2850 extracted mouse oocytes. An arrow indicates spot that was cored from the gel for tandem mass spectroscopic identification of ePLA2 γ peptides. Two other protein spots of higher acidity (to the left of the arrow and contained within brackets) were later shown to be reactive with ePLA2 γ immune sera, possibly indicating postranslationally modified isoforms.

Figure 2. Complementary DNA sequence and deduced amino acid sequence of ePLA2 γ . The untranslated flanking regions are shown in lower case letters. Peptide sequences originally obtained by mass spectrometry are underlined.

Figure 3. Amino acid sequence comparison of ePLA2 γ with human cPLA2 γ . The sequences were aligned using the ClustalW alignment program in BioEdit sequence alignment editor. Identical residues are shown in black boxes and conserved residues are shown in grey boxes. Dashes indicate gaps inserted to maximize sequence alignment. As opposed to human cPLA2 γ , ePLA2 γ lacks obvious N-terminal myristoylation (M-G-X-X-X(S/small uncharged)-X) and C-terminal prenylation (-CCLA) membrane anchoring motifs shown in white boxes

Figure 4. Tissue distribution of ePLA2 γ . Northern blot analysis was performed on various tissues using a P³²-labeled 900 bp ePLA2 γ DNA fragment. ePLA2 γ was detected as a single ~2.7 kilobase band in the ovarian RNA lane. A β -actin probe was used to control for RNA loading. Tissues included: Ovary, Heart, Brain, Spleen, Lung, Liver, Skeletal Muscle (SM), Kidney, and testis.

Figure 5. Western blot analysis demonstrating specificity of ePLA2 γ immune antisera. The preimmune sera was not reactive with either the recombinant ePLA2 γ protein or with egg proteins (50 mature zona-intact eggs per lane) while the anti- ePLA2 γ immune antisera was reactive with recombinant ePLA2 γ and with a ~67.8k Da egg protein.

Figure 6. Two dimensional western blot analysis demonstrating that ePLA2 γ immune antisera is reactive with the 67.8 kDa (pI 5.7) protein spot that was originally identified as ePLA2 γ . ePLA2 γ preimmune sera was not reactive with egg proteins. However, the ePLA2 γ immune antisera strongly reacts with the protein spot which contained ePLA2 γ peptides. An arrow indicates the spot which was originally cored for TMS analysis. Note that two other, slightly more acidic proteins are reactive with the immune sera (asterisks) possibly indicating that ePLA2 γ is post-translationally modified. Overlay image localizes reactive proteins with protein spots resolved using protogold.

Figure 7. Localization of ePLA2 γ on mouse ovary. Ovarian cross-sections from mature females were probed with ePLA2 γ antisera and visualized using immunoperoxidase staining. ePLA2 γ expression was observed in eggs (arrows) contained within, preantral (PA) and antral follicles (AF). 100X.

Figure 8. Developmental expression of ePLA2 γ in the oocyte and early embryo. Indirect immunofluorescence (IIF) was performed to determine if ePLA2 γ is expressed in immature and mature oocytes as well as the developing embryo. Immature (GV), Mature (MII), and pronuclear (PN) oocytes/embryos were collected from females, while 2 cell, 4 cell, 16 cell, morula (Mo.), and blastocyst (Bl.) were collected following culture from the pronuclear stage. Indirect immunofluorescence was then performed on the eggs/embryos using either ePLA2 γ immune antisera (A to G) or ePLA2 γ preimmune antisera (H, GV negative control). ePLA2 γ appears to be expressed at all stages shown and mainly

localizes to the egg cortex directly under the inner leaflet of the plasma membrane and nucleoplasm, except in MII where the expression is mainly limited to the egg cortex. Interestingly, in the morula, however, the cortical staining is limited to the external surface of the blastomeres. IIF – indirect immunofluorescence, Ph. – phase contrast,
 5 1000X.

Figure 9. Expression of ePLA2 γ in the oocyte during germinal vesicle breakdown (GVBD). Germinal vesicle stage oocytes were collected from females and either fixed immediately or fixed at different time points during spontaneous oocyte maturation and prepared for indirect immunofluorescence (IIF) in order to investigate the
 10 expression of ePLA2 γ during nuclear envelope dissolution. Dual staining was performed using the anti-ePLA2 γ immune sera (green) and anti-lamin B antibodies (red) to stain the nuclear envelope. Oocytes were divided according to the ePLA2 γ expression patterns: nuclear (A, at time 0), nuclear and aggregated around the nuclear envelope (B, C, D, E, between 1 and 3 h, white arrow) At the MI stage (F) when the nuclear envelope
 15 disappears entirely and ePLA2 γ expression again becomes dispersed. DAPI stain (blue) is shown in F to indicate the location of the chromatin. Black arrow indicates cytoplasmic ePLA2 γ aggregation IIF – indirect immunofluorescence, Ph. – phase contrast, 1000X.

Figure 10. Analysis of total RNA in cat's ovarian tissue

Figure 11. Analysis of total RNA in dog's ovarian tissue

Figure 12. Alignment of ePAD among human, mouse, cat and dog

Figure 13. Western blot of cat ePAD truncate protein. The right panel is stained with Coomassie blue and the left panel is the Western blot using anti-His antibody. (M: molecular standard, Con: bacterium without the insert plasmid, So: soluble part, Inso: insoluble part).
 25

Figure 14. The Prep Cell of the dog ePAD truncate protein and the Western blot. The three tubes (A, B and C) have a similar molecular weight and single band after Prep Cell with Coomassie Staining.

Figure 15. Western blot of the dog ePAD truncate protein after Prep Cell. The
 30 three tubes, (A, B and C), came from the Prep Cell of Figure 14. They have an obviously

single band in Western blot with anti-His antibody. The sample in Con. lane was from the bacteria without plasmid.

Figure 16. Western blot of cat and dog ePAD truncate protein after purification with anti-mouse ePAD antibody. Lanes- 1 and 4: molecular standard; 2 and 7: mouse testis tissue; 3: dog truncate protein; 6: cat truncate protein; 8: mouse egg.

Figure 17 represents an image of a photomicrograph depicting a mouse ovary stained with autoantibodies to a recombinant mouse ePAD vaccine. The entire egg cytoplasm in this antral follicle is recognized.

Figure 18 represents an image of a photomicrograph depicting a mouse ovary stained with mouse auto-antibodies to recombinant mouse ePAD demonstrating recognition of eggs in follicles at various stages of development and maturation.

Figure 19 represents an image of a gel demonstrating PCR products using mouse oocyte cDNA.

Figure 20, comprising left (A) and right (B) panels represents images of gels depicting MOP26 purification by Ni-ATN affinity column. FT- flow through, W1, W2 – washes.

Figure 21, demonstrates an image of a gel depicting the results of MOP26 purification on PrepCell and its quantification. (total protein stain).

Figure 22, comprising left (A) and right (B) panels, represents images of immunoblot analyses of recombinant MOP26 and total protein from 35 mouse oocytes probed with MOP26 antiserum raised in rat (A) and preimmune serum (B).

Figure 23, comprising four panels, represents photomicrographic images of an immunofluorescence analysis of mouse oocytes with anti-MOP26 (panel A; left upper photograph), its companion unstained image (right panel; C), preimmune (panel B; lower left panel), and the companion unstained preimmune image (lower right panel; D). Mouse oocytes were fixed with PFA and permeabilized with 0.5% Triton X-100.

Figure 24, comprising four panels, represents photomicrographic images of mouse ovary sections with anti-MOP26 (panel A; upper left panel), its companion unstained image (upper right panel; panel C), preimmune serum (panel B; lower left panel) and its unstained image (lower right panel; panel D).

Figure 25, comprising left and right panels, represents photomicrographic images of an immunofluorescence analysis of live zona-intact and zona-free mouse oocytes with anti-MOP26 serum. In zona-intact oocytes, staining of the perivitelline space was observed.

5 **Figure 26**, comprises a map of proteins cored for sequencing (MOP26, 27, 28, 29, 30, 31, and 32).

10 **Figure 27** represents an image of an electrophoretic analysis of PCR products with mouse oocyte cDNA. For PCR, total RNA was extracted from 500 mouse eggs and reverse transcribed to make mouse egg cDNA. Specific PCR products were cloned into TOPO vector and sequenced, and subjected to electrophoretic analysis. Groups include MOP26, Control, and MOP31. The far right lane represents a Mass-Ladder.

15 **Figure 28**, comprising left and right panels, represents an electrophoretic analysis of MOP26 expression in HMS174 cells (left panel; panel A) and in BL21 cells (right panel; panel B). Ni-NTA blot.

20 **Figure 29** represents an image of a Western blot analysis of MOP26 expression in a variety of cells and tissues, including, oocyte, ovary, testes, thymus, liver, spleen kidney, intestine, stomach, lung, and heart.

25 **Figure 30**, comprising upper and lower panels, represents images of a multi-tissue RT-PCR analysis of MOP26 (upper panel; panel A) and β -actin (lower panel; panel B) in heart, lung, liver, pancreas, kidney, testis, ovary, spleen thymus, and oocyte.

30 **Figure 31**, comprising six panels, represents photomicrographic images of an immunofluorescent analysis of MOP26 expression through oocyte maturation and early embryonic development. The six states examined include gv (upper left), m2 (upper center), pn (upper right), 2-cell (lower left), 4-8 cell (lower middle), and blastocyst (lower right).

Other aspects and advantages of the present invention are described herein and in the following detailed description of the preferred embodiments thereof.

DETAILED DESCRIPTION OF THE INVENTION

35 **Definitions-**

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described herein. In describing and claiming the invention, the following terminology will be used in accordance with the definitions set forth below.

The articles “a” and “an” are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

A disease, disorder, or condition is “alleviated” if the severity of a symptom of the disease or disorder, the frequency with which such a symptom is experienced by a patient, or both, are reduced.

As used herein, “amino acids” are represented by the full name thereof, by the three letter code corresponding thereto, or by the one-letter code corresponding thereto, as indicated in the following table:

	<u>Full Name</u>	<u>Three-Letter Code</u>	<u>One-Letter Code</u>
	Aspartic Acid	Asp	D
	Glutamic Acid	Glu	E
20	Lysine	Lys	K
	Arginine	Arg	R
	Histidine	His	H
	Tyrosine	Tyr	Y
	Cysteine	Cys	C
25	Asparagine	Asn	N
	Glutamine	Gln	Q
	Serine	Ser	S
	Threonine	Thr	T
	Glycine	Gly	G
30	Alanine	Ala	A
	Valine	Val	V

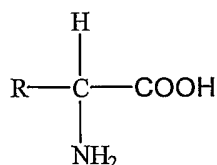
	Leucine	Leu	L
	Isoleucine	Ile	I
	Methionine	Met	M
	Proline	Pro	P
5	Phenylalanine	Phe	F
	Tryptophan	Trp	W

The expression "amino acid" as used herein is meant to include both natural and synthetic amino acids, and both D and L amino acids. "Standard amino acid" means any of the twenty standard L-amino acids commonly found in naturally occurring peptides. "Nonstandard amino acid residue" means any amino acid, other than the standard amino acids, regardless of whether it is prepared synthetically or derived from a natural source. As used herein, "synthetic amino acid" also encompasses chemically modified amino acids, including but not limited to salts, amino acid derivatives (such as amides), and substitutions. Amino acids contained within the peptides of the present invention, and particularly at the carboxy- or amino-terminus, can be modified by methylation, amidation, acetylation or substitution with other chemical groups which can change the peptide's circulating half-life without adversely affecting their activity. Additionally, a disulfide linkage may be present or absent in the peptides of the invention.

The term "amino acid" is used interchangeably with "amino acid residue," and may refer to a free amino acid and to an amino acid residue of a peptide. It will be apparent from the context in which the term is used whether it refers to a free amino acid or a residue of a peptide.

Amino acids have the following general structure:

25



Amino acids may be classified into seven groups on the basis of the side chain R:
 (1) aliphatic side chains, (2) side chains containing a hydroxylic (OH) group, (3) side chains containing sulfur atoms, (4) side chains containing an acidic or amide group, (5)

side chains containing a basic group, (6) side chains containing an aromatic ring, and (7) proline, an imino acid in which the side chain is fused to the amino group.

The nomenclature used to describe the peptide compounds of the present invention follows the conventional practice wherein the amino group is presented to the left and the carboxy group to the right of each amino acid residue. In the formulae representing selected specific embodiments of the present invention, the amino- and carboxy-terminal groups, although not specifically shown, will be understood to be in the form they would assume at physiologic pH values, unless otherwise specified.

As used herein, an "analog" of a chemical compound is a compound that, by way of example, resembles another in structure but is not necessarily an isomer (e.g., 5-fluorouracil is an analog of thymine).

The term "antibody," as used herein, refers to an immunoglobulin molecule which is able to specifically bind to a specific epitope on an antigen. Antibodies can be intact immunoglobulins derived from natural sources or from recombinant sources and can be immunoreactive portions of intact immunoglobulins. Antibodies are typically tetramers of immunoglobulin molecules. The antibodies in the present invention may exist in a variety of forms including, for example, polyclonal antibodies, monoclonal antibodies, Fv, Fab and F(ab)₂, as well as single chain antibodies and humanized antibodies (Harlow et al., 1999, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY; Harlow et al., 1989, Antibodies: A Laboratory Manual, Cold Spring Harbor, New York; Houston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; Bird et al., 1988, Science 242:423-426).

By the term "synthetic antibody" as used herein, is meant an antibody which is generated using recombinant DNA technology, such as, for example, an antibody expressed by a bacteriophage as described herein. The term should also be construed to mean an antibody which has been generated by the synthesis of a DNA molecule encoding the antibody and which DNA molecule expresses an antibody protein, or an amino acid sequence specifying the antibody, wherein the DNA or amino acid sequence has been obtained using synthetic DNA or amino acid sequence technology which is available and well known in the art.

As used herein, the term “antisense oligonucleotide” means a nucleic acid polymer, at least a portion of which is complementary to a nucleic acid which is present in a normal cell or in an affected cell. The antisense oligonucleotides of the invention include, but are not limited to, phosphorothioate oligonucleotides and other modifications of oligonucleotides. Methods for synthesizing oligonucleotides, phosphorothioate oligonucleotides, and otherwise modified oligonucleotides are well known in the art (U.S. Patent No: 5,034,506; Nielsen et al., 1991, Science 254: 1497). “Antisense” refers particularly to the nucleic acid sequence of the non-coding strand of a double stranded DNA molecule encoding a protein, or to a sequence which is substantially homologous to the non-coding strand. As defined herein, an antisense sequence is complementary to the sequence of a double stranded DNA molecule encoding a protein. It is not necessary that the antisense sequence be complementary solely to the coding portion of the coding strand of the DNA molecule. The antisense sequence may be complementary to regulatory sequences specified on the coding strand of a DNA molecule encoding a protein, which regulatory sequences control expression of the coding sequences.

The term “basic” or “positively charged” amino acid as used herein, refers to amino acids in which the R groups have a net positive charge at pH 7.0, and include, but are not limited to, the standard amino acids lysine, arginine, and histidine.

The term “biocompatible”, as used herein, refers to a material that does not elicit a substantial detrimental response in the host.

As used herein, the term “biologically active fragments” or “bioactive fragment” of the polypeptides encompasses natural or synthetic portions of the full-length protein that are capable of specific binding to their natural ligand or of performing the function of the protein.

The terms “cell,” “cell line,” and “cell culture” as used herein may be used interchangeably. All of these terms also include their progeny, which are any and all subsequent generations. It is understood that all progeny may not be identical due to deliberate or inadvertent mutations.

“Complementary” refers to the broad concept of sequence complementarity between regions of two nucleic acid strands or between two regions of the same nucleic acid strand. It is known that an adenine residue of a first nucleic acid region is capable of

forming specific hydrogen bonds ("base pairing") with a residue of a second nucleic acid region which is antiparallel to the first region if the residue is thymine or uracil. As used herein, the terms "complementary" or "complementarity" are used in reference to polynucleotides (i.e., a sequence of nucleotides) related by the base-pairing rules. For
5 example, for the sequence "A-G-T," is complementary to the sequence "T-C-A."

Similarly, it is known that a cytosine residue of a first nucleic acid strand is capable of base pairing with a residue of a second nucleic acid strand which is antiparallel to the first strand if the residue is guanine. A first region of a nucleic acid is complementary to a second region of the same or a different nucleic acid if, when the two
10 regions are arranged in an antiparallel fashion, at least one nucleotide residue of the first region is capable of base pairing with a residue of the second region. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, when the first and second portions are arranged in an antiparallel fashion, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about
15 95% of the nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion. More preferably, all nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion.

A "compound," as used herein, refers to a protein, polypeptide, an isolated nucleic acid, or other agent used in the method of the invention.

20 As used herein, the term "conservative amino acid substitution" is defined herein as an amino acid exchange within one of the following five groups:

I. Small aliphatic, nonpolar or slightly polar residues:

Ala, Ser, Thr, Pro, Gly;

II. Polar, negatively charged residues and their amides:

25 Asp, Asn, Glu, Gln;

III. Polar, positively charged residues:

His, Arg, Lys;

IV. Large, aliphatic, nonpolar residues:

Met Leu, Ile, Val, Cys

30 V. Large, aromatic residues:

Phe, Tyr, Trp

“Contraceptive”, as used herein, refers to an agent, compound, or method that diminishes the likelihood of or prevents conception.

5 A “control” cell, tissue, sample, or subject is a cell, tissue, sample, or subject of the same type as a test cell, tissue, sample, or subject. The control may, for example, be examined at precisely or nearly the same time the test cell, tissue, sample, or subject is examined. The control may also, for example, be examined at a time distant from the time at which the test cell, tissue, sample, or subject is examined, and the results of the examination of the control may be recorded so that the recorded results may be compared
10 with results obtained by examination of a test cell, tissue, sample, or subject. The control may also be obtained from another source or similar source other than the test group or a test subject, where the test sample is obtained from a subject suspected of having a disease or disorder for which the test is being performed.

A “test” cell, tissue, sample, or subject is one being examined or treated.

15 A “pathoindicative” cell, tissue, or sample is one which, when present, is an indication that the animal in which the cell, tissue, or sample is located (or from which the tissue was obtained) is afflicted with a disease or disorder. By way of example, the presence of one or more breast cells in a lung tissue of an animal is an indication that the animal is afflicted with metastatic breast cancer.

20 A tissue “normally comprises” a cell if one or more of the cell are present in the tissue in an animal not afflicted with a disease or disorder.

The use of the word “detect” and its grammatical variants is meant to refer to measurement of the species without quantification, whereas use of the word “determine” or “measure” with their grammatical variants are meant to refer to measurement of the
25 species with quantification. The terms “detect” and “identify” are used interchangeably herein.

As used herein, a “detectable marker” or a “reporter molecule” is an atom or a molecule that permits the specific detection of a compound comprising the marker in the presence of similar compounds without a marker. Detectable markers or reporter
30 molecules include, e.g., radioactive isotopes, antigenic determinants, enzymes, nucleic acids available for hybridization, chromophores, fluorophores, chemiluminescent

molecules, electrochemically detectable molecules, and molecules that provide for altered fluorescence-polarization or altered light-scattering.

5 A “disease” is a state of health of an animal wherein the animal cannot maintain homeostasis, and wherein if the disease is not ameliorated then the animal's health continues to deteriorate.

In contrast, a “disorder” in an animal is a state of health in which the animal is able to maintain homeostasis, but in which the animal's state of health is less favorable than it would be in the absence of the disorder. Left untreated, a disorder does not necessarily cause a further decrease in the animal's state of health.

10 As used herein, the phrases “egg protein” or “egg-specific protein” refer to proteins which are expressed exclusively or predominately in eggs or ovaries. The proteins need not be expressed at all stages of egg or ovarian development.

“Encoding” refers to the inherent property of specific sequences of nucleotides in a polynucleotide, such as a gene, a cDNA, or an mRNA, to serve as templates for
15 synthesis of other polymers and macromolecules in biological processes having either a defined sequence of nucleotides (i.e., rRNA, tRNA and mRNA) or a defined sequence of amino acids and the biological properties resulting therefrom. Thus, a gene encodes a protein if transcription and translation of mRNA corresponding to that gene produces the protein in a cell or other biological system. Both the coding strand, the nucleotide
20 sequence of which is identical to the mRNA sequence and is usually provided in sequence listings, and the non-coding strand, used as the template for transcription of a gene or cDNA, can be referred to as encoding the protein or other product of that gene or cDNA.

Unless otherwise specified, a “nucleotide sequence encoding an amino acid
25 sequence” includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. Nucleotide sequences that encode proteins and RNA may include introns.

An “enhancer” is a DNA regulatory element that can increase the efficiency of transcription, regardless of the distance or orientation of the enhancer relative to the start
30 site of transcription.

As used herein, an “essentially pure” preparation of a particular protein or peptide is a preparation wherein at least about 95%, and preferably at least about 99%, by weight, of the protein or peptide in the preparation is the particular protein or peptide.

5 A “fragment” or “segment” is a portion of an amino acid sequence, comprising at least one amino acid, or a portion of a nucleic acid sequence comprising at least one nucleotide. The terms “fragment” and “segment” are used interchangeably herein.

As used herein, a “functional” biological molecule is a biological molecule in a form in which it exhibits a property or activity by which it is characterized. A functional enzyme, for example, is one which exhibits the characteristic catalytic activity by which
10 the enzyme is characterized.

“Homologous” as used herein, refers to the subunit sequence similarity between two polymeric molecules, e.g., between two nucleic acid molecules, e.g., two DNA molecules or two RNA molecules, or between two polypeptide molecules. When a subunit position in both of the two molecules is occupied by the same monomeric
15 subunit, e.g., if a position in each of two DNA molecules is occupied by adenine, then they are homologous at that position. The homology between two sequences is a direct function of the number of matching or homologous positions, e.g., if half (e.g., five positions in a polymer ten subunits in length) of the positions in two compound sequences are homologous then the two sequences are 50% homologous, if 90% of the
20 positions, e.g., 9 of 10, are matched or homologous, the two sequences share 90% homology. By way of example, the DNA sequences 3'ATTGCC5' and 3'TATGGC share 50% homology.

As used herein, “homology” is used synonymously with “identity.”

As used herein, the term “hybridization” is used in reference to the pairing of
25 complementary nucleic acids. Hybridization and the strength of hybridization (i.e., the strength of the association between the nucleic acids) is impacted by such factors as the degree of complementarity between the nucleic acids, stringency of the conditions involved, the length of the formed hybrid, and the G:C ratio within the nucleic acids.

The determination of percent identity between two nucleotide or amino acid
30 sequences can be accomplished using a mathematical algorithm. For example, a mathematical algorithm useful for comparing 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). This algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, et al. (1990, J. Mol. Biol. 215:403-410), and can be accessed, for example at the National Center for
5 Biotechnology Information (NCBI) world wide web site. BLAST nucleotide searches can be performed with the NBLAST program (designated "blastn" at the NCBI web site), using the following parameters: gap penalty = 5; gap extension penalty = 2; mismatch penalty = 3; match reward = 1; expectation value 10.0; and word size = 11 to obtain nucleotide sequences homologous to a nucleic acid described herein. BLAST protein
10 searches can be performed with the XBLAST program (designated "blastn" at the NCBI web site) or the NCBI "blastp" program, using the following parameters: expectation value 10.0, BLOSUM62 scoring matrix to obtain amino acid sequences homologous to a protein molecule described herein. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al. (1997, Nucleic
15 Acids Res. 25:3389-3402). Alternatively, PSI-Blast or PHI-Blast can be used to perform an iterated search which detects distant relationships between molecules (Id.) and relationships between molecules which share a common pattern. When utilizing BLAST, Gapped BLAST, PSI-Blast, and PHI-Blast programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used.

20 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.

By the term "immunizing a subject against an antigen" is meant, administering to the subject a composition, a protein complex, a DNA encoding a protein complex, an
25 antibody or a DNA encoding an antibody, which elicits an immune response in the subject, which immune response provides protection to the subject against the condition caused by the antigen or related to the presence of the antigen.

The term "inhibit," as used herein, refers to the ability of a compound of the invention to reduce or impede a described function. Preferably, inhibition is by at least
30 10%, more preferably by at least 25%, even more preferably by at least 50%, and most preferably, the function is inhibited by at least 75%.

The phrase “inhibit conception”, as used herein, refers to both direct and indirect inhibition of conception or impregnation, regardless of the mechanism. The phrase also includes reducing the rate of conception, and does not necessarily mean that conception is inhibited by 100%.

5 As used herein, an “instructional material” includes a publication, a recording, a diagram, or any other medium of expression which can be used to communicate the usefulness of the peptide of the invention in the kit for effecting alleviation of the various diseases or disorders recited herein. Optionally, or alternately, the instructional material may describe one or more methods of alleviating the diseases or disorders in a cell or a
10 tissue of a mammal. The instructional material of the kit of the invention may, for example, be affixed to a container which contains the identified compound invention or be shipped together with a container which contains the identified compound. Alternatively, the instructional material may be shipped separately from the container with the intention that the instructional material and the compound be used cooperatively
15 by the recipient.

 An “isolated nucleic acid” refers to a nucleic acid segment or fragment which has been separated from sequences which flank it in a naturally occurring state, e.g., a DNA fragment which has been removed from the sequences which are normally adjacent to the fragment, e.g., the sequences adjacent to the fragment in a genome in which it naturally
20 occurs. The term also applies to nucleic acids which have been substantially purified from other components which naturally accompany the nucleic acid, e.g., RNA or DNA or proteins, which naturally accompany it in the cell. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or
25 which exists as a separate molecule (e.g., as a cDNA or a genomic or cDNA fragment produced by PCR or restriction enzyme digestion) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence.

 Unless otherwise specified, a “nucleotide sequence encoding an amino acid
30 sequence” includes all nucleotide sequences that are degenerate versions of each other

and that encode the same amino acid sequence. Nucleotide sequences that encode proteins and RNA may include introns.

As used herein, a “ligand” is a compound that specifically binds to a target compound. A ligand (e.g., an antibody) “specifically binds to” or “is specifically immunoreactive with” a compound when the ligand functions in a binding reaction which is determinative of the presence of the compound in a sample of heterogeneous compounds. Thus, under designated assay (e.g., immunoassay) conditions, the ligand binds preferentially to a particular compound and does not bind to a significant extent to other compounds present in the sample. For example, an antibody specifically binds under immunoassay conditions to an antigen bearing an epitope against which the antibody was raised. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular antigen. For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with an antigen. See Harlow and Lane, 1988, Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity.

As used herein, the term “linkage” refers to a connection between two groups. The connection can be either covalent or non-covalent, including but not limited to ionic bonds, hydrogen bonding, and hydrophobic/hydrophilic interactions.

As used herein, the term “linker” refers to a molecule that joins two other molecules either covalently or noncovalently, e.g., through ionic or hydrogen bonds or van der Waals interactions.

By “nucleic acid” is meant any nucleic acid, whether composed of deoxyribonucleosides or ribonucleosides, and whether composed of phosphodiester linkages or modified linkages such as phosphotriester, phosphoramidate, siloxane, carbonate, carboxymethylester, acetamidate, carbamate, thioether, bridged phosphoramidate, bridged methylene phosphonate, bridged phosphoramidate, bridged phosphoramidate, bridged methylene phosphonate, phosphorothioate, methylphosphonate, phosphorodithioate, bridged phosphorothioate or sulfone linkages, and combinations of such linkages. The term nucleic acid also specifically includes nucleic acids composed of bases other than the five biologically occurring bases

(adenine, guanine, thymine, cytosine and uracil). Conventional notation is used herein to describe polynucleotide sequences: the left-hand end of a single-stranded polynucleotide sequence is the 5'-end; the left-hand direction of a double-stranded polynucleotide sequence is referred to as the 5'-direction. The direction of 5' to 3' addition of nucleotides to nascent RNA transcripts is referred to as the transcription direction. The DNA strand having the same sequence as an mRNA is referred to as the "coding strand"; sequences on the DNA strand which are located 5' to a reference point on the DNA are referred to as "upstream sequences"; sequences on the DNA strand which are 3' to a reference point on the DNA are referred to as "downstream sequences."

The term "oligonucleotide" typically refers to short polynucleotides, generally no greater than about 50 nucleotides. It will be understood that when a nucleotide sequence is represented by a DNA sequence (i.e., A, T, G, C), this also includes an RNA sequence (i.e., A, U, G, C) in which "U" replaces "T."

"Operably linked" refers to a juxtaposition wherein the components are configured so as to perform their usual function. Thus, control sequences or promoters operably linked to a coding sequence are capable of effecting the expression of the coding sequence. By describing two polynucleotides as "operably linked" is meant that a single-stranded or double-stranded nucleic acid moiety comprises the two polynucleotides arranged within the nucleic acid moiety in such a manner that at least one of the two polynucleotides is able to exert a physiological effect by which it is characterized upon the other. By way of example, a promoter operably linked to the coding region of a gene is able to promote transcription of the coding region.

As used herein, a "peptide" encompasses a sequence of 2 or more amino acid residues wherein the amino acids are naturally occurring or synthetic (non-naturally occurring) amino acids covalently linked by peptide bonds. No limitation is placed on the number of amino acid residues which can comprise a protein's or peptide's sequence. As used herein, the terms "peptide," "polypeptide," and "protein" are used interchangeably. Peptide mimetics include peptides having one or more of the following modifications:

1. peptides wherein one or more of the peptidyl --C(O)NR-- linkages (bonds) have been replaced by a non-peptidyl linkage such as a --CH₂--carbamate linkage

(--CH₂OC(O)NR--), a phosphonate linkage, a -CH₂-sulfonamide (-CH₂-S(O)₂NR--) linkage, a urea (--NHC(O)NH--) linkage, a --CH₂-secondary amine linkage, or with an alkylated peptidyl linkage (--C(O)NR--) wherein R is C₁-C₄ alkyl;

2. peptides wherein the N-terminus is derivatized to a --NRR₁ group, to a
 5 -- NRC(O)R group, to a --NRC(O)OR group, to a --NRS(O)₂R group, to a --NHC(O)NHR group where R and R₁ are hydrogen or C₁-C₄ alkyl with the proviso that R and R₁ are not both hydrogen;

3. peptides wherein the C terminus is derivatized to --C(O)R₂ where R₂ is
 10 selected from the group consisting of C₁-C₄ alkoxy, and --NR₃R₄ where R₃ and R₄ are independently selected from the group consisting of hydrogen and C₁-C₄ alkyl.

Synthetic or non-naturally occurring amino acids refer to amino acids which do not naturally occur in vivo but which, nevertheless, can be incorporated into the peptide structures described herein. The resulting "synthetic peptide" contains amino acids other than the 20 naturally occurring, genetically encoded amino acids at one, two, or more
 15 positions of the peptides. For instance, naphthylalanine can be substituted for tryptophan to facilitate synthesis. Other synthetic amino acids that can be substituted into peptides include L-hydroxypropyl, L-3,4-dihydroxyphenylalanyl, alpha-amino acids such as L-alpha-hydroxylysyl and D-alpha-methylalanyl, L-alpha-methylalanyl, beta-amino acids, and isoquinolyl. D amino acids and non-naturally occurring synthetic amino acids
 20 can also be incorporated into the peptides. Other derivatives include replacement of the naturally occurring side chains of the 20 genetically encoded amino acids (or any L or D amino acid) with other side chains.

As used herein, the term "pharmaceutically acceptable carrier" includes any of the standard pharmaceutical carriers, such as a phosphate buffered saline solution, water,
 25 emulsions such as an oil/water or water/oil emulsion, and various types of wetting agents. The term also encompasses any of the agents approved by a regulatory agency of the US Federal government or listed in the US Pharmacopeia for use in animals, including humans.

A “polylinker” is a nucleic acid sequence that comprises a series of three or more different restriction endonuclease recognitions sequences closely spaced to one another (i.e. less than 10 nucleotides between each site).

5 A “polynucleotide” means a single strand or parallel and anti-parallel strands of a nucleic acid. Thus, a polynucleotide may be either a single-stranded or a double-stranded nucleic acid.

“Polypeptide” refers to a polymer composed of amino acid residues, related naturally occurring structural variants, and synthetic non-naturally occurring analogs thereof linked via peptide bonds, related naturally occurring structural variants, and
10 synthetic non-naturally occurring analogs thereof. Synthetic polypeptides can be synthesized, for example, using an automated polypeptide synthesizer.

The term “protein” typically refers to large polypeptides.

The term “peptide” typically refers to short polypeptides.

As used herein, the term “promoter/regulatory sequence” means a nucleic acid
15 sequence which is required for expression of a gene product operably linked to the promoter/regulator sequence. In some instances, this sequence may be the core promoter sequence and in other instances, this sequence may also include an enhancer sequence and other regulatory elements which are required for expression of the gene product. The promoter/regulatory sequence may, for example, be one which expresses the gene
20 product in a tissue specific manner.

The term “non-native promoter” as used herein refers to any promoter that has been operably linked to a coding sequence wherein the coding sequence and the promoter are not naturally associated (i.e. a recombinant promoter/coding sequence construct).

25 A “tissue-specific” promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a living cell substantially only if the cell is a cell of the tissue type corresponding to the promoter.

Unless otherwise specified, a “nucleotide sequence encoding an amino acid sequence” includes all nucleotide sequences that are degenerate versions of each other
30 and that encode the same amino acid sequence. Nucleotide sequences that encode proteins and RNA may include introns.

“Plurality” means at least two.

As used herein, “protecting group” with respect to a terminal amino group refers to a terminal amino group of a peptide, which terminal amino group is coupled with any of various amino-terminal protecting groups traditionally employed in peptide synthesis. Such protecting groups include, for example, acyl protecting groups such as formyl, acetyl, benzoyl, trifluoroacetyl, succinyl, and methoxysuccinyl; aromatic urethane protecting groups such as benzyloxycarbonyl; and aliphatic urethane protecting groups, for example, tert-butoxycarbonyl or adamantyloxycarbonyl. See Gross and Mienhofer, eds., *The Peptides*, vol. 3, pp. 3-88 (Academic Press, New York, 1981) for suitable protecting groups.

As used herein, “protecting group” with respect to a terminal carboxy group refers to a terminal carboxyl group of a peptide, which terminal carboxyl group is coupled with any of various carboxyl-terminal protecting groups. Such protecting groups include, for example, tert-butyl, benzyl or other acceptable groups linked to the terminal carboxyl group through an ester or ether bond.

As used herein, the term “purified” and like terms relate to an enrichment of a molecule or compound relative to other components normally associated with the molecule or compound in a native environment. The term “purified” does not necessarily indicate that complete purity of the particular molecule has been achieved during the process. A “highly purified” compound as used herein refers to a compound that is greater than 90% pure. In particular, purified sperm cell DNA refers to DNA that does not produce significant detectable levels of non-sperm cell DNA upon PCR amplification of the purified sperm cell DNA and subsequent analysis of that amplified DNA.

“Recombinant polynucleotide” refers to a polynucleotide having sequences that are not naturally joined together. An amplified or assembled recombinant polynucleotide may be included in a suitable vector, and the vector can be used to transform a suitable host cell. A recombinant polynucleotide may serve a non-coding function (e.g., promoter, origin of replication, ribosome-binding site, etc.) as well.

A host cell that comprises a recombinant polynucleotide is referred to as a “recombinant host cell.” A gene which is expressed in a recombinant host cell wherein the gene comprises a recombinant polynucleotide, produces a “recombinant polypeptide.”

A “recombinant polypeptide” is one which is produced upon expression of a recombinant polynucleotide.

5 A “reversibly implantable” device is one which may be inserted (e.g. surgically or by insertion into a natural orifice of the animal) into the body of an animal and thereafter removed without great harm to the health of the animal.

A “sample,” as used herein, refers preferably to a biological sample from a subject, including, but not limited to, normal tissue samples, diseased tissue samples, biopsies, blood, saliva, feces, semen, tears, and urine. A sample can also be any other source of material obtained from a subject which contains cells, tissues, or fluid of
10 interest. A sample can also be obtained from cell or tissue culture.

As used herein, the term “secondary antibody” refers to an antibody that binds to the constant region of another antibody (the primary antibody).

By the term “signal sequence” is meant a polynucleotide sequence which encodes a peptide that directs the path a polypeptide takes within a cell, i.e., it directs the cellular
15 processing of a polypeptide in a cell, including, but not limited to, eventual secretion of a polypeptide from a cell. A signal sequence is a sequence of amino acids which are typically, but not exclusively, found at the amino terminus of a polypeptide which targets the synthesis of the polypeptide to the endoplasmic reticulum. In some instances, the signal peptide is proteolytically removed from the polypeptide and is thus absent from the
20 mature protein.

As used herein, the term “solid support” relates to a solvent insoluble substrate that is capable of forming linkages (preferably covalent bonds) with various compounds. The support can be either biological in nature, such as, without limitation, a cell or bacteriophage particle, or synthetic, such as, without limitation, an acrylamide derivative,
25 agarose, cellulose, nylon, silica, or magnetized particles.

By the term “specifically binds,” as used herein, is meant an antibody or compound which recognizes and binds a molecule of interest (e.g., an antibody directed against a polypeptide of the invention), but does not substantially recognize or bind other molecules in a sample.

30 The term “standard,” as used herein, refers to something used for comparison. For example, a standard can be a known standard agent or compound which is

administered or added to a control sample and used for comparing results when measuring said compound in a test sample. Standard can also refer to an "internal standard," such as an agent or compound which is added at known amounts to a sample and is useful in determining such things as purification or recovery rates when a sample is processed or subjected to purification or extraction procedures before a marker of interest is measured.

A "subject" of analysis, diagnosis, or treatment is an animal. Such animals include mammals.

As used herein, a "substantially homologous amino acid sequences" includes those amino acid sequences which have at least about 95% homology, preferably at least about 96% homology, more preferably at least about 97% homology, even more preferably at least about 98% homology, and most preferably at least about 99% or more homology to an amino acid sequence of a reference antibody chain. Amino acid sequence similarity or identity can be computed by using the BLASTP and TBLASTN programs which employ the BLAST (basic local alignment search tool) 2.0.14 algorithm. The default settings used for these programs are suitable for identifying substantially similar amino acid sequences for purposes of the present invention.

"Substantially homologous nucleic acid sequence" means a nucleic acid sequence corresponding to a reference nucleic acid sequence wherein the corresponding sequence encodes a peptide having substantially the same structure and function as the peptide encoded by the reference nucleic acid sequence; e.g., where only changes in amino acids not significantly affecting the peptide function occur. Preferably, the substantially identical nucleic acid sequence encodes the peptide encoded by the reference nucleic acid sequence. The percentage of identity between the substantially similar nucleic acid sequence and the reference nucleic acid sequence is at least about 50%, 65%, 75%, 85%, 95%, 99% or more. Substantial identity of nucleic acid sequences can be determined by comparing the sequence identity of two sequences, for example by physical/chemical methods (i.e., hybridization) or by sequence alignment via computer algorithm. Suitable nucleic acid hybridization conditions to determine if a nucleotide sequence is substantially similar to a reference nucleotide sequence are: 7% sodium dodecyl sulfate SDS, 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 2X standard saline citrate

(SSC), 0.1% SDS at 50°C; preferably in 7% (SDS), 0.5 M NaPO₄, 1 mM EDTA at 50°C. with washing in 1X SSC, 0.1% SDS at 50°C; preferably 7% SDS, 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.5X SSC, 0.1% SDS at 50°C; and more preferably in 7% SDS, 0.5 M NaPO₄, 1 mM EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at
5 65°C. Suitable computer algorithms to determine substantial similarity between two nucleic acid sequences include, GCS program package (Devereux et al., 1984 Nucl. Acids Res. 12:387), and the BLASTN or FASTA programs (Altschul et al., 1990 Proc. Natl. Acad. Sci. USA. 1990 87:14:5509-13; Altschul et al., J. Mol. Biol. 1990 215:3:403-10; Altschul et al., 1997 Nucleic Acids Res. 25:3389-3402). The default settings
10 provided with these programs are suitable for determining substantial similarity of nucleic acid sequences for purposes of the present invention.

The term “substantially pure” describes a compound, e.g., a protein or polypeptide which has been separated from components which naturally accompany it. Typically, a compound is substantially pure when at least 10%, more preferably at least
15 20%, more preferably at least 50%, more preferably at least 60%, more preferably at least 75%, more preferably at least 90%, and most preferably at least 99% of the total material (by volume, by wet or dry weight, or by mole percent or mole fraction) in a sample is the compound of interest. Purity can be measured by any appropriate method, e.g., in the case of polypeptides by column chromatography, gel electrophoresis, or HPLC analysis.
20 A compound, e.g., a protein, is also substantially purified when it is essentially free of naturally associated components or when it is separated from the native contaminants which accompany it in its natural state.

A “therapeutic” treatment is a treatment administered to a subject who exhibits signs of pathology for the purpose of diminishing or eliminating those signs.

25 A “therapeutically effective amount” of a compound is that amount of compound which is sufficient to provide a beneficial effect to the subject to which the compound is administered.

As used herein, the term “transgene” means an exogenous nucleic acid sequence comprising a nucleic acid which encodes a promoter/regulatory sequence operably linked
30 to nucleic acid which encodes an amino acid sequence, which exogenous nucleic acid is encoded by a transgenic mammal.

As used herein, the term “transgenic mammal” means a mammal, the germ cells of which comprise an exogenous nucleic acid.

As used herein, a “transgenic cell” is any cell that comprises a nucleic acid sequence that has been introduced into the cell in a manner that allows expression of a
5 gene encoded by the introduced nucleic acid sequence.

As used herein, the term “treating” includes prophylaxis of the specific disorder or condition, or alleviation of the symptoms associated with a specific disorder or condition and/or preventing or eliminating said symptoms. A “prophylactic” treatment is a treatment administered to a subject who does not exhibit signs of a disease or exhibits
10 only early signs of the disease for the purpose of decreasing the risk of developing pathology associated with the disease. As used herein, the term “treating” includes alleviating the symptoms associated with a specific disease, disorder or condition and/or preventing or eliminating said symptoms.

By the term “vaccine,” as used herein, is meant a composition which when
15 inoculated into a subject has the effect of stimulating an immune response in the subject, which serves to fully or partially protect the subject against a condition, disease or its symptoms. In one aspect, the condition is conception. The term vaccine encompasses prophylactic as well as therapeutic vaccines. A combination vaccine is one which combines two or more vaccines, or two or more compounds or agents.

20 A “vector” is a composition of matter which comprises an isolated nucleic acid and which can be used to deliver the isolated nucleic acid to the interior of a cell. Numerous vectors are known in the art including, but not limited to, linear polynucleotides, polynucleotides associated with ionic or amphiphilic compounds, plasmids, and viruses. Thus, the term “vector” includes an autonomously replicating
25 plasmid or a virus. The term should also be construed to include non-plasmid and non-viral compounds which facilitate transfer of nucleic acid into cells, such as, for example, polylysine compounds, liposomes, and the like. Examples of viral vectors include, but are not limited to, adenoviral vectors, adeno-associated virus vectors, retroviral vectors, plasmids, cosmids, lambda phage vectors, and the like.

30 “Expression vector” refers to a vector comprising a recombinant polynucleotide comprising expression control sequences operatively linked to a nucleotide sequence to

be expressed. An expression vector comprises sufficient cis-acting elements for expression; other elements for expression can be supplied by the host cell or in an in vitro expression system. Expression vectors include all those known in the art, such as cosmids, plasmids (e.g., naked or contained in liposomes) and viruses that incorporate the recombinant polynucleotide.

The present invention provides novel egg proteins, and homologs, derivatives, and fragments thereof, useful for contraceptive vaccines in animals. Such proteins elicit an immune response against said proteins. The present invention further provides combination vaccines of at least one novel egg protein and at least one known egg protein. The present invention further provides isolated nucleic acids comprising nucleic acid sequences encoding the egg proteins of the invention.

In one aspect, the novel egg proteins of the invention are selected from the group of proteins having the amino acid sequences SEQ ID NOs:5, 7, 9, 11, 15 and 17 (i.e., murine ePLA2 γ , murine MOP31, cat ePAD, dog ePAD, and murine MOP26, and a fragment of EPLA not present in MOP31). In one aspect, the known egg proteins is SEQ ID NO:13.

The present invention further provides antibodies against the egg proteins of the invention.

It will be appreciated, of course, that the proteins or peptides of the invention may incorporate amino acid residues which are modified without affecting activity. For example, the termini may be derivatized to include blocking groups, i.e. chemical substituents suitable to protect and/or stabilize the N- and C-termini from "undesirable degradation", a term meant to encompass any type of enzymatic, chemical or biochemical breakdown of the compound at its termini which is likely to affect the function of the compound, i.e. sequential degradation of the compound at a terminal end thereof.

Blocking groups include protecting groups conventionally used in the art of peptide chemistry which will not adversely affect the *in vivo* activities of the peptide. For example, suitable N-terminal blocking groups can be introduced by alkylation or acylation of the N-terminus. Examples of suitable N-terminal blocking groups include C₁-C₅ branched or unbranched alkyl groups, acyl groups such as formyl and acetyl groups, as well as substituted forms thereof, such as the acetamidomethyl (Acm) group.

Desamino analogs of amino acids are also useful N-terminal blocking groups, and can either be coupled to the N-terminus of the peptide or used in place of the N-terminal residue. Suitable C-terminal blocking groups, in which the carboxyl group of the C-terminus is either incorporated or not, include esters, ketones or amides. Ester or ketone-forming alkyl groups, particularly lower alkyl groups such as methyl, ethyl and propyl, and amide-forming amino groups such as primary amines ($-NH_2$), and mono- and di-alkylamino groups such as methylamino, ethylamino, dimethylamino, diethylamino, methylethylamino and the like are examples of C-terminal blocking groups.

Descarboxylated amino acid analogues such as agmatine are also useful C-terminal blocking groups and can be either coupled to the peptide's C-terminal residue or used in place of it. Further, it will be appreciated that the free amino and carboxyl groups at the termini can be removed altogether from the peptide to yield desamino and descarboxylated forms thereof without affect on peptide activity.

Other modifications can also be incorporated without adversely affecting the activity and these include, but are not limited to, substitution of one or more of the amino acids in the natural L-isomeric form with amino acids in the D-isomeric form. Thus, the peptide may include one or more D-amino acid residues, or may comprise amino acids which are all in the D-form. Retro-inverso forms of peptides in accordance with the present invention are also contemplated, for example, inverted peptides in which all amino acids are substituted with D-amino acid forms.

Acid addition salts of the present invention are also contemplated as functional equivalents. Thus, a peptide in accordance with the present invention treated with an inorganic acid such as hydrochloric, hydrobromic, sulfuric, nitric, phosphoric, and the like, or an organic acid such as an acetic, propionic, glycolic, pyruvic, oxalic, malic, malonic, succinic, maleic, fumaric, tataric, citric, benzoic, cinnamic, mandelic, methanesulfonic, ethanesulfonic, p-toluenesulfonic, salicylic and the like, to provide a water soluble salt of the peptide is suitable for use in the invention.

Modifications (which do not normally alter primary sequence) include in vivo, or in vitro chemical derivatization of polypeptides, e.g., acetylation, or carboxylation. Also included are modifications of glycosylation, e.g., those made by modifying the glycosylation patterns of a polypeptide during its synthesis and processing or in further

processing steps; e.g., by exposing the polypeptide to enzymes which affect glycosylation, e.g., mammalian glycosylating or deglycosylating enzymes. Also embraced are sequences which have phosphorylated amino acid residues, e.g., phosphotyrosine, phosphoserine, or phosphothreonine.

5 Also included are polypeptides which have been modified using ordinary molecular biological techniques so as to improve their resistance to proteolytic degradation or to optimize solubility properties or to render them more suitable as a therapeutic agent. Analogs of such polypeptides include those containing residues other than naturally occurring L-amino acids, e.g., D-amino acids or non-naturally occurring
10 synthetic amino acids. The peptides of the invention are not limited to products of any of the specific exemplary processes listed herein.

 Nucleic acids useful in the present invention include, by way of example and not limitation, oligonucleotides and polynucleotides such as antisense DNAs and/or RNAs; ribozymes; DNA for gene therapy; viral fragments including viral DNA and/or RNA;
15 DNA and/or RNA chimeras; mRNA; plasmids; cosmids; genomic DNA; cDNA; gene fragments; various structural forms of DNA including single-stranded DNA, double-stranded DNA, supercoiled DNA and/or triple-helical DNA; Z-DNA; and the like. The nucleic acids may be prepared by any conventional means typically used to prepare nucleic acids in large quantity. For example, DNAs and RNAs may be chemically
20 synthesized using commercially available reagents and synthesizers by methods that are well-known in the art (see, e.g., Gait, 1985, OLIGONUCLEOTIDE SYNTHESIS: A PRACTICAL APPROACH (IRL Press, Oxford, England)). RNAs may be produced in high yield via in vitro transcription using plasmids such as SP65 (Promega Corporation, Madison, WI).

25 The peptides of the present invention may be readily prepared by standard, well-established techniques, such as solid-phase peptide synthesis (SPPS) as described by Stewart et al. in Solid Phase Peptide Synthesis, 2nd Edition, 1984, Pierce Chemical Company, Rockford, Illinois; and as described by Bodanszky and Bodanszky in The Practice of Peptide Synthesis, 1984, Springer-Verlag, New York. At the outset, a
30 suitably protected amino acid residue is attached through its carboxyl group to a derivatized, insoluble polymeric support, such as cross-linked polystyrene or polyamide

resin. "Suitably protected" refers to the presence of protecting groups on both the α -amino group of the amino acid, and on any side chain functional groups. Side chain protecting groups are generally stable to the solvents, reagents and reaction conditions used throughout the synthesis, and are removable under conditions which will not affect the final peptide product. Stepwise synthesis of the oligopeptide is carried out by the removal of the N-protecting group from the initial amino acid, and couple thereto of the carboxyl end of the next amino acid in the sequence of the desired peptide. This amino acid is also suitably protected. The carboxyl of the incoming amino acid can be activated to react with the N-terminus of the support-bound amino acid by formation into a reactive group such as formation into a carbodiimide, a symmetric acid anhydride or an "active ester" group such as hydroxybenzotriazole or pentafluorophenyl esters.

Examples of solid phase peptide synthesis methods include the BOC method which utilized tert-butoxycarbonyl as the α -amino protecting group, and the Fmoc method which utilizes 9-fluorenylmethyloxycarbonyl to protect the α -amino of the amino acid residues, both methods of which are well-known by those of skill in the art.

Incorporation of N- and/or C- blocking groups can also be achieved using protocols conventional to solid phase peptide synthesis methods. For incorporation of C-terminal blocking groups, for example, synthesis of the desired peptide is typically performed using, as solid phase, a supporting resin that has been chemically modified so that cleavage from the resin results in a peptide having the desired C-terminal blocking group. To provide peptides in which the C-terminus bears a primary amino blocking group, for instance, synthesis is performed using a p-methylbenzhydrylamine (MBHA) resin so that, when peptide synthesis is completed, treatment with hydrofluoric acid releases the desired C-terminally amidated peptide. Similarly, incorporation of an N-methylamine blocking group at the C-terminus is achieved using N-methylaminoethyl-derivatized DVB, resin, which upon HF treatment releases a peptide bearing an N-methylamidated C-terminus. Blockage of the C-terminus by esterification can also be achieved using conventional procedures. This entails use of resin/blocking group combination that permits release of side-chain peptide from the resin, to allow for subsequent reaction with the desired alcohol, to form the ester function. Fmoc protecting group, in combination with DVB resin derivatized with methoxyalkoxybenzyl

alcohol or equivalent linker, can be used for this purpose, with cleavage from the support being effected by TFA in dichloromethane. Esterification of the suitably activated carboxyl function e.g. with DCC, can then proceed by addition of the desired alcohol, followed by deprotection and isolation of the esterified peptide product.

5 Incorporation of N-terminal blocking groups can be achieved while the synthesized peptide is still attached to the resin, for instance by treatment with a suitable anhydride and nitrile. To incorporate an acetyl blocking group at the N-terminus, for instance, the resin-coupled peptide can be treated with 20% acetic anhydride in acetonitrile. The N-blocked peptide product can then be cleaved from the resin,
10 deprotected and subsequently isolated.

To ensure that the proteins or peptides obtained from either chemical or biological synthetic techniques is the desired peptide, analysis of the peptide composition should be conducted. Such amino acid composition analysis may be conducted using high resolution mass spectrometry to determine the molecular weight of the peptide.

15 Alternatively, or additionally, the amino acid content of the peptide can be confirmed by hydrolyzing the peptide in aqueous acid, and separating, identifying and quantifying the components of the mixture using HPLC, or an amino acid analyzer. Protein sequenators, which sequentially degrade the peptide and identify the amino acids in order, may also be used to determine definitely the sequence of the peptide.

20 Prior to its use, the peptide can be purified to remove contaminants. In this regard, it will be appreciated that the peptide will be purified to meet the standards set out by the appropriate regulatory agencies. Any one of a number of a conventional purification procedures may be used to attain the required level of purity including, for example, reversed-phase high-pressure liquid chromatography (HPLC) using an alkylated
25 silica column such as C₄-, C₈- or C₁₈- silica. A gradient mobile phase of increasing organic content is generally used to achieve purification, for example, acetonitrile in an aqueous buffer, usually containing a small amount of trifluoroacetic acid. Ion-exchange chromatography can be also used to separate peptides based on their charge.

30 Substantially pure peptide obtained as described herein may be purified by following known procedures for protein purification, wherein an immunological, enzymatic or other assay is used to monitor purification at each stage in the procedure.

Protein purification methods are well known in the art, and are described, for example in Deutscher et al. (ed., 1990, Guide to Protein Purification, Harcourt Brace Jovanovich, San Diego).

5 The present invention is also directed to pharmaceutical compositions comprising the compounds of the present invention. More particularly, such compounds can be formulated as pharmaceutical compositions using standard pharmaceutically acceptable carriers, fillers, solubilizing agents and stabilizers known to those skilled in the art.

10 The invention is also directed to methods of administering the compounds of the invention to a subject. In one embodiment, the invention provides a method of treating a subject by administering compounds identified using the methods of the invention description. Pharmaceutical compositions comprising the present compounds are administered to a subject in need thereof by any number of routes including, but not limited to, topical, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, 15 enteral, topical, sublingual, or rectal means.

In accordance with one embodiment, a method of treating a subject in need of such treatment is provided. The method comprises administering a pharmaceutical composition comprising at least one compound of the present invention to a subject in need thereof. Compounds identified by the methods of the invention can be administered 20 with known compounds or other medications as well.

The invention also encompasses the use of pharmaceutical compositions of an appropriate compound, and homologs, fragments, analogs, or derivatives thereof to practice the methods of the invention, the composition comprising at least one appropriate compound, and homolog, fragment, analog, or derivative thereof and a 25 pharmaceutically-acceptable carrier.

The pharmaceutical compositions useful for practicing the invention may be administered to deliver a dose of between 1 ng/kg/day and 100 mg/kg/day.

30 The invention encompasses the preparation and use of pharmaceutical compositions comprising a compound useful for treatment of the diseases disclosed herein as an active ingredient. Such a pharmaceutical composition may consist of the active ingredient alone, in a form suitable for administration to a subject, or the

pharmaceutical composition may comprise the active ingredient and one or more pharmaceutically acceptable carriers, one or more additional ingredients, or some combination of these. The active ingredient may be present in the pharmaceutical composition in the form of a physiologically acceptable ester or salt, such as in
5 combination with a physiologically acceptable cation or anion, as is well known in the art.

As used herein, the term "physiologically acceptable" ester or salt means an ester or salt form of the active ingredient which is compatible with any other ingredients of the pharmaceutical composition, which is not deleterious to the subject to which the
10 composition is to be administered.

The formulations of the pharmaceutical compositions described herein may be prepared by any method known or hereafter developed in the art of pharmacology. In general, such preparatory methods include the step of bringing the active ingredient into association with a carrier or one or more other accessory ingredients, and then, if
15 necessary or desirable, shaping or packaging the product into a desired single- or multi-dose unit.

It will be understood by the skilled artisan that such pharmaceutical compositions are generally suitable for administration to animals of all sorts. Subjects to which administration of the pharmaceutical compositions of the invention is contemplated
20 include, but are not limited to, humans and other primates, mammals including commercially relevant mammals such as cattle, pigs, horses, sheep, cats, and dogs, birds including commercially relevant birds such as chickens, ducks, geese, and turkeys. The invention is also contemplated for use in contraception for nuisance animals such as rodents.

25 A pharmaceutical composition of the invention may be prepared, packaged, or sold in bulk, as a single unit dose, or as a plurality of single unit doses. As used herein, a "unit dose" is discrete amount of the pharmaceutical composition comprising a predetermined amount of the active ingredient. The amount of the active ingredient is generally equal to the dosage of the active ingredient which would be administered to a
30 subject or a convenient fraction of such a dosage such as, for example, one-half or one-third of such a dosage.

The relative amounts of the active ingredient, the pharmaceutically acceptable carrier, and any additional ingredients in a pharmaceutical composition of the invention will vary, depending upon the identity, size, and condition of the subject treated and further depending upon the route by which the composition is to be administered. By way of example, the composition may comprise between 0.1% and 100% (w/w) active ingredient.

In addition to the active ingredient, a pharmaceutical composition of the invention may further comprise one or more additional pharmaceutically active agents. Particularly contemplated additional agents include anti-emetics and scavengers such as cyanide and cyanate scavengers.

Controlled- or sustained-release formulations of a pharmaceutical composition of the invention may be made using conventional technology.

As used herein, "additional ingredients" include, but are not limited to, one or more of the following: excipients; surface active agents; dispersing agents; inert diluents; granulating and disintegrating agents; binding agents; lubricating agents; sweetening agents; flavoring agents; coloring agents; preservatives; physiologically degradable compositions such as gelatin; aqueous vehicles and solvents; oily vehicles and solvents; suspending agents; dispersing or wetting agents; emulsifying agents, demulcents; buffers; salts; thickening agents; fillers; emulsifying agents; antioxidants; antibiotics; antifungal agents; stabilizing agents; and pharmaceutically acceptable polymeric or hydrophobic materials. Other "additional ingredients" which may be included in the pharmaceutical compositions of the invention are known in the art and described, for example in Genaro, ed., 1985, Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, PA, which is incorporated herein by reference.

Typically, dosages of the compound of the invention which may be administered to an animal, preferably a human, range in amount from 1 μ g to about 100 g per kilogram of body weight of the animal. While the precise dosage administered will vary depending upon any number of factors, including but not limited to, the type of animal and type of disease state being treated, the age of the animal and the route of administration. Preferably, the dosage of the compound will vary from about 1 mg to

about 10 g per kilogram of body weight of the animal. More preferably, the dosage will vary from about 10 mg to about 1 g per kilogram of body weight of the animal.

The compound may be administered to an animal as frequently as several times daily, or it may be administered less frequently, such as once a day, once a week, once every two weeks, once a month, or even less frequently, such as once every several months or even once a year or less. The frequency of the dose will be readily apparent to the skilled artisan and will depend upon any number of factors, such as, but not limited to, the type and severity of the disease being treated, the type and age of the animal, etc.

The invention also includes a kit comprising the composition of the invention and an instructional material which describes adventitiously administering the composition to a cell or a tissue of a mammal. In another embodiment, this kit comprises a (preferably sterile) solvent suitable for dissolving or suspending the composition of the invention prior to administering the compound to the mammal.

As used herein, an "instructional material" includes a publication, a recording, a diagram, or any other medium of expression which can be used to communicate the usefulness of the peptide of the invention in the kit for effecting alleviation of the various diseases or disorders recited herein. Optionally, or alternately, the instructional material may describe one or more methods of alleviation the diseases or disorders in a cell or a tissue of a mammal. The instructional material of the kit of the invention may, for example, be affixed to a container which contains the peptide of the invention or be shipped together with a container which contains the peptide. Alternatively, the instructional material may be shipped separately from the container with the intention that the instructional material and the compound be used cooperatively by the recipient.

The invention is now described with reference to the following Examples and Embodiments. Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the present invention and practice the claimed methods. The following working examples therefore, are provided for the purpose of illustration only and specifically point out the preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure. Therefore, the examples should

be construed to encompass any and all variations which become evident as a result of the teaching provided herein.

Examples

5 **Example 1- ePLA2 γ**

Materials and Methods.

Identification of ePLA2 γ peptide sequences by tandem mass spectroscopy.

Mouse oocyte proteins (2850) were collected and separated on 16 cm 2D electrophoretic gels, as previously described (Wright, et al., 2003, Dev. Biol. 256, 73-88).
10 A protein spot of approximately 67.8 kDa (pI 5.7) was cored from a Coomassie-stained gel, digested with trypsin, and microsequenced by tandem mass spectrometry as previously described (Wright, et al., 2003, Dev. Biol. 256, 73-88). The obtained peptide sequence information was then compared against database sequence using the NCBI Blast Website (<http://www.ncbi.nlm.nih.gov/BLAST/>).

15 ***RACE-PCR Cloning of ePLA2 γ cDNA.***

The ePLA2 γ open reading frame was amplified from a mouse oocyte adapter-ligated cDNA library by RT-PCR using DNA polymerase (Amplitaq Gold, Perkin-Elmer, Norwalk, CT) as previously described (Wright et al., 2003). Cycling parameters were: 94°C, 10 minutes; 94°C, 15 seconds; 60°C, 30 seconds; 72°C, 2 minutes; and 72°C, 10
20 minutes, for 40 cycles. PCR of plasmid templates was performed under the same conditions except that taq polymerase (Promega Advantage 2) was used. The primers for cloning the full length ORF of ePLA2 γ into the TOPO cloning vector (Invitrogen) were 5'-ATGGA ACTAAGCTCTGGGGTC - 3' (SEQ ID NO:1) and 5'-
AGGGTGTGTATAGATTCCTA-3' (SEQ ID NO:2), and the AP1 primer was supplied
25 with in the Marathon ready cDNA kit (Clontech, Palo Alto, CA).

Northern blot analysis.

A randomly primed probe was generated corresponding to the 900 bp N-terminal region of ePLA2 γ using the Prime-a-Gene Labeling System (Promega, Madison, WI). Northern blot analysis was performed as described previously (Wright et al., 2003).
30 Signals were detected by exposure to X-ray film for 10 days.

Expression of Recombinant Protein and Antibody Production

Primers were designed to generate a PCR product encoding amino acids 1-110 of the ePLA2 γ open reading frame. The sense gene-specific primer sequence was 5'-ATGGAAGTAAGCTCTGGGGTC - 3' (SEQ ID NO:1) with the three prime end containing an engineered BamHI restriction site. The antisense gene specific primer
5 sequence was 5'-TCCAGTCTAAAGAAGCAAGTG-3' (SEQ ID NO:3) and contained an engineered XhoI restriction site. The primers were used to amplify the cDNA fragment, and the PCR product was cloned into the BamHI-XhoI restriction sites of the pET22b expression vector (Novagen, Madison, WI).

The recombinant protein was then expressed and purified as described previously
10 (Wright et al., 2003, Dev. Biol. 256:73-88). Preimmune serum was collected from adult male guinea pigs, and each animal was then injected with 100 μ g of purified recombinant ePLA2 γ in an emulsion with Freund's complete adjuvant. The animals received one booster immunization with 50 μ g of purified recombinant ePLA2 γ in Freund's incomplete adjuvant at a 3-week interval. The polyclonal sera was then collected and
15 prepared as previously described (Wright et al., 2003, Dev. Biol. 256:73-88).

Western blot analysis of ePLA2 γ .

BL21 cells expressing recombinant ePLA2 γ were sonicated and 20 μ g of bacterial protein as well as fifty oocytes, were solubilized in Laemmli buffer (Laemmli, 1970) and heat denatured at 95°C. The samples were then loaded onto a 12.5% linear gel, and
20 separated at 100 V for three hours. Proteins were then electro-transferred onto 0.2 μ m nitrocellulose membrane (Bio-Rad Laboratories, Hercules, CA) for 40 min at 100 V for western blotting.

For two dimensional western blot analysis, proteins from ~ 800 zona-intact ovulated metaphase II arrested eggs were extracted, separated by two-dimensional gel
25 electrophoresis (16 X 16 cm PAGE), and immunoblotted as previously described. A 1:2000 dilution of ePLA2 γ preimmune and immune antisera were used for these experiments.

Indirect Immunofluorescence of Oocytes and Embryos.

All oocytes and embryos were obtained from ICR 25-30 g females. Germinal
30 vesicle oocytes and metaphase II eggs were obtained as described previously (Coonrod, et al., 2001 Genesis. 30, 198-200; Coonrod et al., 1999, Dev. Biol. 207, 334-349).

Pronuclear staged embryos were isolated from the oviducts of superovulated female mice in Whitten's media with HEPES (Specialty media, Phillipsburg, NJ) containing 0.05% hyaluronidase. Embryos were washed thoroughly in immunofluorescence (IF) media (PBS + 1% BSA + 0.5 NGS) and either removed for fixation or allowed to develop in vitro at 37°C and 5% CO₂ in maturation buffer (TYH buffer) as previously described. Oocytes and embryos were fixed in 4% paraformaldehyde in PBS for 30 min at room temperature. Following fixation, oocytes and embryos were washed 5 times in IF media and then permeabilized with 0.5% Triton-X 100 in PBS for 30 min. Oocytes and embryos were then washed 5 times and where incubated with ePLA2 γ guinea pig preimmune and immune antisera 1:500 in IF media for 1 h at room temperature, with a subset of these being incubated with a 1:1000 dilution of lamin B IgG (kindly donated by David Spector) overnight. Oocytes and embryos were washed 5 times and incubated for 3 h at room temperature with goat anti-guinea pig FITC-labeled secondary antibody (Jackson Immunoresearch) alone or together with goat anti-rabbit Texas Red-labeled secondary antibody (Jackson Immunoresearch) to recognize lamin B. Oocytes/embryos were mounted on slides and visualized at 1000X under a Zeiss Axiovert-200 deconvolution fluorescence microscope and imaged.

Immunohistochemistry of Ovarian Sections.

Five μ m sections of adult female mouse ovaries were fixed in 4% paraformaldehyde, embedded in paraffin wax, and processed for immunohistochemistry. In brief, the paraffin wax was removed using Neo Clear ClarificationTM (EM Science, NJ, USA) and the section was rehydrated through a graded series of ethanol solutions (100-70%) in H₂O. To localize ePLA2 γ in the histological sections, endogenous peroxidase activity was first removed by treatment with 3% H₂O₂ for 20 minutes at room temperature. After washing with PBS-BSA (0.15%), slides were blocked for 20 minutes with BSA (2%), then incubated overnight with the ePLA2 γ guinea pig preimmune or immune antisera (1:500 in PBS-BSA 0.15%). Next, the slides were washed and incubated with biotinylated secondary antibody for 30 minutes, washed again and treated with VECTASTAIN elite ABC reagent (Vector Laboratories, Burlingame, CA) for 30 minutes. Finally, the sections were incubated with a peroxidase substrate solution until the desired staining levels were obtained, rinsed in tap water, counterstained with

hematoxylin, cleared with Neo Clear ClarificationTM, mounted on slides using mounting media (Permout, Fisher Chemicals) and visualized at 100X.

RESULTS-

Tandem Mass Spectroscopic Identification of ePLA2 γ , a Highly Abundant Novel Egg Protein.

Proteins from extracts of 2850 zona-free mouse eggs were separated on a two dimensional electrophoretic gel and stained with Coomassie (Fig. 1). One protein spot, later designated ePLA2 γ (Fig. 1, arrow), having a molecular weight of 67.8 kDa and isoelectric point of 5.7, was cored from the gel, digested with trypsin, and microsequenced by CAD mass spectrometry. Twenty-one unique peptides were identified in that spot and, when searched against non-redundant database sequence, did not match previously characterized protein sequence. We therefore concluded that this protein was novel. However, one peptide with the sequence FIEGPVTYSEAPR (SEQ ID NO:4) did match an EST clone, J0258C02, from a fertilized egg cDNA library.

Cloning and Characterization of ePLA2 γ .

Oligonucleotide primers (designed from either the FIEGPVTYSEAPR (SEQ ID NO:4) peptide sequence or from the matching EST sequences) were used for repeated RACE-PCR reactions to amplify a 2701 bp cDNA from the ovarian cDNA library which contained the ePLA2 γ open reading frame. The DNA was inserted into a cloning vector and subsequent sequence analysis revealed a 75 bp 5' untranslated region, a contiguous open reading frame of 1791 bp (encoding a deduced protein of 597 amino acids), and a 796 bp 3' untranslated region (Fig. 2). The untranslated flanking regions are shown in lower case letters. The GenBank accession number for this clone is **AY694793**. The computed mass and isoelectric point (ExPASy compute pi/MW algorithm) of the deduced amino acid sequence are 66 and 5.36, respectively, which closely matches the mass and pI of the protein spot that was cored from the 2D gel for TMS analysis. Thirteen of the 21 peptides originally identified by TMS were found in the deduced amino acid sequence and are underlined. The deduced amino acid sequence was compared against database sequence using the NCBI conserved domain function located at the NCBI website) and significant similarities to a cytoplasmic phospholipase A2 (PLA2) domain (E value = 2e-31) and a PLA2 B lysophospholipase catalytic domain (E

value = $3e-51$) were observed. The first domain is found in enzymes that hydrolyze arachidonyl phospholipids and overlaps the ~ 300 amino acid N-terminus of ePLA2 γ . The latter, observed between residues 45 and 503, is found in enzymes that catalyze the release of fatty acids from lysophospholipids.

5 The mouse ePLA2 γ amino acid (SEQ ID NO:5) and nucleic acid (SEQ ID NO:6) sequences of the invention are as follows:

MELSSGVCPATRLQEAEKAAVHKRSPKVLEALRKLNIQADQAPVIAVLGSGGGGL
 RAHIACLGVLSELKELGLLDAVTYLAGVSGSTWALSSLYTKNGNMEGIEEELKH
 10 RYEKNEWDFHESLEKAIQASKRENYSLTDFWAYLIVSRQIRELQDSNLSSLKKQV
 EEGVLPYPIFAAIDEDLLADWRERKTQNSWFEFTPHHAGYPALGAYVPITEFGSR
 FENGKLVKSEPERDLTFLRGLWGSADFADIKNIKNYILNYFRNPFGLKLFIEGPVTY
 SEAPRMNVDAMLLDLVMAFYFTDMNDPSIKDKLCALQQALGTETDEFGIEMAEII
 QNWNETSAEKKEQFLDHLDRFKKTQEDTTTYSLMNWNTGLVWDRCVFNVT
 15 RKCYSKWQWGTVYNFLYKHGKIADETMCSRELLHLVDAGFAINTPYPLVLPPV
 RETHLILSFDIFSAGDPLETIRATADYCRHEIPFEVSEDQLKEWAKAPASCYVLR
 GETGPVVMHFTLFNKDNCGDDIETWRKKYGTVKLSDSYTPDLVRDLLRVSKEN
 VKKNKINILSEMRKVAGNPGNIPRVNKEACLGDRVKDPQGSQTVEFKKSHNISK
 D (SEQ ID NO:5)

20

1 atggcttctg cggttgctcct gctccctgac atgcttgagt tccagtcctg acgtccttgg
 61 tgataaacag cagcatggaa ctaagctctg gggctctgccc tgccactaga ctccaggaag
 121 cagaaaaggc agctgtgcac aaacgcagtc ccaagggtttt ggaggctcta cgaaagctca
 181 acatccaagc tgaccaggct ccagtcattg ctgtcttggg ctctggcggg gggctgcggg
 241 cccacatcgc ttgtcttggg gtgctgagtg agctgaaaga acttggcctg ttggatgctg
 25 301 tcacatacct cgcaggggtc tctgggtcca cttgggcact gtcttcactc tacaccaaga
 361 atggaaatat ggaagggata gaagaggagc tgaaacatcg gtatgagaag aatgagtggg
 421 actttcatga gagcctggag aaagccatcc aggcatacaa gagggagaat tactccctga
 481 ctgacttttg ggcctattta attgtttcca ggcaaatacag agaacttcag gattcgaatt
 541 tgtccagtct aaagaagcaa gtggaagaag gagtgtctgcc ctatcccatc tttgcagcca
 601 ttgatgagga ctttctggct gattggaggg agaggaaaac tcagaattcc tggtttgaat
 661 tcaactcctca tcatgctggc taccctgcac ttggggctta tgtcccatc acagagtttg
 721 gaagcagatt tgagaatggg aaactgggta aatctgagcc tgagagagat ttgactttcc
 781 tgagagggtt atggggaagt gcttttctg atattaaaga aattaagaat tatattttga

```

841 actacttcag gaaccctttt ggaaaattga agtttataga aggaccagtg acatactcag
901 aagcaccaag gatgaatgtg gatgcaatgc tcttgattt agtgatggct tatttcacag
961 atatgaatga cccagcatc aaggataagc tctgtgccct tcagcaggct ctgggtactg
1021 agacagatga atttggcata gagatggctg agatcatcca gaattggaat gagacctccg
5 1081 cagagaagaa ggagcagttt ctcgaccatc tgttgatcg cttcaagaag acacaagaag
1141 acaccaccac atacagtttg atgaactgga acacgggtct agtttgggac cgttgcgttt
1201 ttgtgaatga aactcgcaag tgtgtctcca aatggcagtg gggaaactgtt tacaacttcc
1261 tctataaaca cggtaaaatt gcagatgaga ccatgtgcag ccgagagctt ctccatctgg
1321 tggatgctgg ttttgccatc aacactccct atccacttgt tctgcctcct gtgcgtgaaa
10 1381 ctcacctcat cctctcatth gacttcagt gctggggacc actagagacc atcagggcca
1441 cagcagacta ctgccaacgc catgaaatcc cttttcctga ggtgagcgag gatcagttga
1501 aggaatgggc caaagcccca gcaagctgct atgtcctcag aggggaaaca ggacctgttg
1561 tcatgcactt tactctgttc aacaaagaca actgtggaga tgatattgag acatggagaa
1621 aaaaatatgg gacagtaaaa ctatctgact catacacacc agacctgggtg agagatttgc
15 1681 tgagggtatc caaggagaat gtgaagaaaa acaaaattaa taccctcagt gagatgagga
1741 aagtggctgg gaatcctggg aacatcccaa gagtgaacaa ggaggcctgc ttgggagaca
1801 gagtaaagga tcccaaggc tctcagactg tggagttaa gaaatccac aacatatcta
1861 aggattaagg atcctttgca ctacctggag agttggcatc tattagtgga ctcaccacct
1921 gtgctctctc caaagtccac cgtaagctgt ctctttgtga agaagaactg cagttttcag
20 1981 ggtactgtgg gcctactgct ttaccagggc ccaagtgtca acctggcctg atgttctctg
2041 ccactcataa aattgcattt gccttagcta ttaccattga gcaagtctcc agagacctag
2101 atgggtgtctt tgatagcatt atttagggca ctgggtcacc agagctaaaa gaggcaaagt
2161 atgttcttca gatactcag gggacctaa agctagagca caactccttg atcatcaaag
2221 ccacctgtc caacatccat gtggttaaaa tgaatgtagt gattgagacc aaagcctcat
25 2281 aaatTTTTTct ctattcttn atntgctcat tttatacacc cagtcttant aaatTTTTTt
2341 tttntataa aaggatagac cttagctcaa atcatgacc atcacttcag agcaattaa
2401 aaaaccagn caactntgag aagacatgtt ctgtatatgc ttngtatct gctgccttag
2461 tgctctctc cngcttgct acttctntc ctttccatcc cttccacctg taactttnt
2521 ttgatttgga ttccacaatn ttntcattgt aatatcacag gcaagttcct tcacatgaat
30 2581 tcctttaatc atntatggaa taaatgtgca ttgtggttgt gccctggtnt ttccaaaaaa
2641 aaaaaaaaaa aaaaaaaaaa aaa (SEQ ID NO:6).

```

Comparison of the sequence against the NCBI non-redundant database using the BLASTp algorithm found that the sequence was 100% identical to a deduced protein sequence (XP_149881.23) named phospholipase A2 gamma, group IVC (cytosolic calcium-independent). The sequence was ~ 80% identical to rat PLA2 γ (a deduced sequence), ~ 56% identical to previously characterized human cPLA2 γ (a protein which

has been characterized) and ~ 27% identical to chicken cPLA2 α . Based on this homology, we propose naming the identified molecule ePLA2 γ , for egg and embryo-abundant phospholipase A2 γ -like protein. A further comparison of the ePLA2 γ sequence with human PLA2 γ shows that both molecules lack the cPLA2 α calcium-dependent lipid-binding domain yet retain the cPLA2 α catalytic triad (Fig. 3, arrows). However, as opposed to human cPLA2 γ , ePLA2 γ lacks obvious N-terminal myristoylation (M-G-X-X-X(S/small uncharged)-X) and C-terminal prenylation (-CCLA) membrane anchoring motifs (Fig. 3, boxes) indicating that perhaps ePLA2 γ is not membrane anchored.

ePLA2 γ is Expression is limited to Ovaries.

An EST database search using the ePLA2 γ ORF nucleotide sequence found multiple identical or near identical matches with ovary, unfertilized egg, fertilized egg, two cell, and four cell cDNAs. Single matches were also obtained with cDNA clones derived from adult male colon (GenBank accession no. **BB624941**) and infiltrating ductal carcinoma (GenBank accession no. **BX524494**) somatic tissues. This information led to the prediction that ePLA2 γ may be an egg-abundant transcript whose expression persists throughout the early stages of cleavage. ePLA2 γ expression was then examined in various tissues using northern blot analysis (Fig. 4). Two blots, one containing 5 μ g multitissue poly-A+ RNA, and the other 40 μ g of total ovarian RNA, were probed with a ³²P-labeled random-primed 900 bp ePLA2 γ N-terminal DNA fragment. The probe, while not binding to RNA from somatic and testicular RNA, did recognize a ~ 2.7 Kb ovarian transcript (Fig. 4). Subsequent multiple tissue northern blotting experiments using a single membrane showed similar results (data not shown).

A 67.8 kDa (pI 5.7) protein spot which contains ePLA2 γ peptide sequence is specifically recognized by anti- ePLA2 γ antibodies.

A cDNA encoding an N-terminal ~110 amino acid region of ePLA2 γ was cloned into a bacterial expression vector, expressed, purified, and used as an immunogen for ePLA2 γ antisera production in guinea pigs. The reactivity of ePLA2 γ antisera with the recombinant protein and with the appropriately sized egg protein was first confirmed by 1-D western blotting. The preimmune serum was not reactive with recombinant ePLA2 γ or with egg proteins. The immune serum was reactive with recombinant ePLA2 γ and an

egg protein of ~ 67.8 kDa (Fig. 5), thus confirming antibody specificity. On 2-D immunoblots, the anti- ePLA2 γ sera reacted with a ~67.8 kDa protein spot (pI 5.7) at the precise location from which the spot was originally cored (arrow), while the preimmune sera was not reactive with egg proteins (Fig. 6). This finding provided an immunological proof that specific antibodies had been generated against the original cored protein. Interestingly the ePLA2 γ immune sera was also reactive with two other slightly more acidic protein spots of similar molecular weight (asterisk in Fig. 6 and Fig. 1) indicating that ePLA2 γ may be post-translationally modified.

Expression and Subcellular Localization of ePLA2 γ

Immunohistochemical localization of ePLA2 γ in adult ovaries showed that the protein mainly localized to oocytes in preantral and antral stage follicles, as indicated by the strong brown staining (Fig. 7). Weak staining of surrounding stromal tissue was also observed. While the preimmune sera was not reactive with oocytes it was weakly reactive with stromal tissue indicating that the staining of the stromal cells with the immune sera was likely non-specific (data not shown). Adjacent sections in which the primary antibody was omitted showed no detectable staining (data not shown).

Next, the expression of ePLA2 γ protein during oocyte maturation, fertilization, and early embryonic development was investigated (Fig. 8). Immature germinal vesicle stage oocytes (GV), mature metaphase II arrested eggs (MII), and pronuclear zygotes (PN) were collected from females and immediately fixed. Pronuclear stage embryos were also collected and then cultured in vitro to the 2 cell, 4 cell, 8-16 cell, morula (Mo.), and blastocyst (Bl.) stage and fixed for immunofluorescent staining. Anti- ePLA2 γ immune sera staining indicated that ePLA2 γ protein was present in oocytes, eggs, and persisted in preimplantation embryos until at least the blastocyst stage of development. At the subcellular level, ePLA2 γ appears to mainly localize to the cortical region of oocytes as well as the nucleoplasm (Fig. 8). In metaphase II stage eggs, ePLA2 γ localizes solely to the egg cortex (Fig. 8B). In pronuclear stage zygotes and blastomeres from preimplantation stage embryos (Fig. 8C-H), intense cortical staining was seen at each stage shown, except in the morula, where the cortical staining was limited to the external surface of the blastomeres. ePLA2 γ staining was also observed in

the nucleoplasm of embryonic blastomeres at all stages and at levels above what was seen in the oocyte cytoplasm. The preimmune serum was not reactive with germinal vesicle stage oocytes (Fig. 8 I), eggs, or preimplantation stage embryos (data not shown).

5 ***Subcellular Localization of ePLA2 γ during Germinal Vesicle Breakdown.***

While investigating ePLA2 γ localization in oocytes, we noticed that the protein appeared to concentrate around the nuclear envelope in some instances and predicted that this expression pattern might correlate with germinal vesicle breakdown (GVBD). To further investigate this possibility, we collected oocytes from follicles and either
10 immediately fixed them for immunofluorescence or cultured them at 37°C and 5% CO₂ and then fixed them at different time points during spontaneous GVBD. Along with ePLA2 γ staining, we followed the status of the nuclear envelope by staining for lamin B, which is attached to the inner leaflet of the nuclear envelope. The results showed three different ePLA2 γ staining patterns during GVBD. When oocytes were fixed immediately
15 after collection, a cortical and homogeneous nuclear staining was observed (Fig. 9A). However, after ~ 1 hour of culture ePLA2 γ appeared to begin accumulating around the nuclear envelope in uniform concentrated foci (Fig. 9B). Between 1 and 3 hours of culture, ePLA2 γ began concentrating at regions of the nuclear envelope which appeared ruffled (Fig. 9 C-E, arrows). Close examination of the oocyte nucleus revealed that the
20 ePLA2 γ accumulation appeared to occur on the cytoplasmic side of the lamin staining indicating that it was likely either co-localizing with, or was adjacent to, the nuclear envelope. Yellow staining pattern indicates co-localization of lamin B with ePLA2 γ . Following dissolution of the nuclear envelope, (as evaluated by a lack of lamin staining and presence of condensed chromatin) after 24 h hours culture, ePLA2 γ is diffusely
25 distributed throughout the cytoplasm (Fig. 9F). When the nuclear envelope reforms at the pronuclear stage, diffuse nuclear staining returns and persist through the subsequent developmental stages until the blastocyst stage as showed in Fig. 8.

DISCUSSION

Described herein is the characterization of ePLA2 γ , a novel 67.8 kDa (pI 5.7) egg and embryo abundant protein that is ~56% identical to human cPLA2 γ . Northern blot analysis shows that ePLA2 γ expression appears to be limited to the ovary.

Immunohistochemical analysis of ovarian cross-sections suggests that ePLA2 γ

5 expression may, in fact, be restricted to the oocyte. The finding that ePLA2 γ continues to be expressed until at least the blastocyst stage of development, suggest that ePLA2 γ may represent a previously uncharacterized maternal effect gene. At the subcellular level, ePLA2 γ mainly localizes to the cortex and nucleoplasm. However, during germinal vesicle breakdown (GVBD), we show that ePLA2 γ appears to aggregate at the nuclear
10 envelope.

As with human cPLA2 γ , ePLA2 γ , lacks, the cPLA2 α calcium-dependent lipid-binding domain yet retains the cPLA2 α catalytic triad (Fig. 3, arrows). As opposed to human cPLA2 γ , ePLA2 γ lacks obvious membrane anchoring motifs (Fig. 3, boxes) indicating that perhaps ePLA2 γ is not membrane anchored. This prediction is supported
15 by our immunofluorescence analysis showing that a portion of ePLA2 γ appears to be diffusely distributed throughout both the cytoplasm and nucleoplasm. In several other cell types, PLA2's have been shown to translocate to either the nuclear envelope or nuclear fraction upon stimulation with calcium ionophore and therefore our finding that ePLA2 γ partially localizes to the nucleus is not without precedent. The data demonstrate
20 that within the cytoplasm, in certain instances following 1 to 3 h culture, ePLA2 γ appeared to concentrate into either one or two large foci (Fig 9D, black arrow). A review of the literature found that this staining pattern was highly reminiscent of a novel oocyte structure termed multivesicular aggregates (MVA). These large (~10 μ m) γ -tubulin-positive structures contain a variety of vesicular structures and migrate towards the GV
25 and then break into smaller units, some of which then mature into microtubule organizing centers.

With regard to tissue distribution, human cPLA2 γ and ePLA2 γ appear to be quite different. Human cPLA2 γ is highly expressed in the heart and skeletal muscle and absent from the ovary, while ePLA2 γ expression appears to be limited to ovaries and is not
30 expressed in heart or skeletal muscle. When this finding is coupled with the fact that

ePLA2 γ partially localizes to the nucleoplasm while human cPLA2 γ subcellular localization appears limited to the cytoplasm, it seems possible that ePLA2 γ is not the ortholog of human cPLA2 γ and represents a previously uncharacterized cPLA2 γ isoform.

Identification of a related ePLA2 γ

Identification of novel egg membrane targets in the mouse oocyte proteome.

5 Mature, metaphase II eggs were collected from superovulated 25-30 gram ICR female mice (10 IU PMSG followed by 10 IU hCG 48 hours later). Twelve hours after hCG injection, oocytes were released from the oviducts into TYH media containing 0.05% hyaluronidase to remove cumulus cells. Zonae pellucida were loosened by
10 treating eggs with TYH containing 10 μ g/ml chymotrypsin for 1 minute. The eggs were then washed 5 times in TYH and the zonae were removed with by mechanical agitation using a pulled Pasteur pipette. The eggs were allowed to recover from chymotrypsin treatment for 3 h in TYH at 37°C and 5% CO₂. The oocytes were then washed 10 times in PBS containing 0.1% PVA and stored at -80°C. Eight hundred and eighty mouse
15 oocytes were processed to prepare Triton X-114 soluble fraction for 2-D gel analysis. Extracted proteins were cleaned up and enriched using the kit from Pierce. Proteins were separated on Criterion 2D gel system. Isoelectric focusing was performed in the pH range 3 to 10, and for the second dimension proteins were separated on a gradient gel 8-16 % acrylamide. The gel was silver stained and five spots from pH range 5-9 were
20 cored and sent for mass spec analysis. Four spots were successfully sequenced and one spot needs to be resubmitted due to technical difficulties with HPLC column during the sequencing procedure.

Peptides from four spots were analyzed by database searching using the Sequest search algorithm. Three known proteins were identified: phospholipase A2, elongation
25 factor 1 alpha, lactate dehydrogenase 2. The fourth spot produced five prominent peptides of an unknown protein with Riken cDNA from *mus musculus*. An EST search showed that this cDNA is present only in fertilized mouse oocytes. According to Riken sequence the molecular weight of the protein is 18.5 kDa and pI is 5.8 which is consistent with experimental data. Analysis of the protein sequence showed no glycosylation sites, six
30 serine and two threonine possible phosphorylation sites and no transmembrane domains.

Hydrophobic Egg Proteins. To further search for egg membrane contraceptive targets, the mouse egg proteome was fractionated by hydrophobicity in order to enrich for oolemmal proteins. Proteins from 600 mouse eggs were separated using Triton X-114 detergent into hydrophobic and hydrophilic fractions and the hydrophobic fraction was enriched for protein on a micro column and then separated on a 2D gel (pH 3-10 and acrylamide gradient 8-16 %). Five proteins from the membrane fraction of mouse oocytes were micro sequenced and identified. Three proteins turned out to be previously known – pyruvate kinase 3, elongation factor, and lactate dehydrogenase. Two proteins were novel and highly embryo- and egg- specific (MOP31 and MOP26).

The gene corresponding to proteome spots MOP31/MOP8 was cloned. MOP31/MOP8 spots contained similar protein microsequences upon tandem mass spectrometry. Analysis of the peptides obtained from the cored protein yielded a novel sequence that appeared to only match ovary, egg, and preimplantation embryo cDNAs. MOP31/MOP8 is a fragment of ePLA2 γ . MOP31 is a novel protein similar to phospholipase A2 catalytic domain. MOP31 is also called MOP8 herein. The nucleic acid (SEQ ID NO:8) and amino acid (SEQ ID NO:7) sequences of MOP31/MOP8 are as follows:

```

1 gaaagaactt ggctgttgg atgctgtcac atacctcgca ggggtctctg ggtccacttg
20   K E L  G L L D A V T Y L A G V S G S T W 20
61 ggcaactgtct tcactctaca ccaagaatgg aaatatggaa gggatagaag aggagctgaa
   A L S  S L Y T K N G N M E G I E E E L K 40
121 acatcggtat gagaagaatg agtgggactt tcatgagagc ctggagaaag ccatccaggc
   H R Y E K N E W D F H E S L E K A I Q A 60
25 181 atcaaagagg gagaattact ccctgactga cttttgggcc tatttaattg tttccaggca
   S K R E N Y S L T D F W A Y L I V S R Q 80
241 aatcagagaa cttcaggatt cgaatttgct cagtctaaag aagcaagtgg aagaaggagt
   I R E L Q D S N L S S L K K Q V E E G V 100
301 gctgccctat cccatctttg cagccattga tgaggacctt ctggctgatt ggagggagag
30   L P Y P I F A A I D E D L L A D W R E R 120
361 gaaaactcag aattcctggt ttgaattcac tcctcatcat gctggctacc ctgcacttgg
   K T Q N S W F E F T P H H A G Y P A L G 140
421 ggcttatgtc ccctcacag agtttgaag cagatttgag aatgggaaac tggttaaatc
   A Y V P I T E F G S R F E N G K L V K S 160
35 481 tgagcctgag agagatttga ctttcctgag aggtttatgg ggaagtgcct ttgctgatat
   E P E R D L T F L R G L W G S A F A D I 180

```

541 taaagaaatt aagaattata ttttgaacta cttcaggaac ccctttggaa aattgaagtt
 K E I K N Y I L N Y F R N P F G K L K F 200
 601 tatagaagga ccagtgcacat actcagaagc accaaggatg aatgtggatg caatgctctt
 I E G P V T Y S E A P R M N V D A M L L 220
 5 661 ggatttagtg atggcttatt tcacagatat gaatgacccc agcatcaagg ataagctctg
 D L V M A Y F T D M N D P S I K D K L C 240
 721 tgcccttcag caggctctgg gtactgagac agatgaatgt ggcatagaga tggctgagat
 A L Q Q A L G T E T D E F G I E M A E I 260
 781 catccagaat tggaatgaga cctccgcaga gaagaaggag cagtttctcg accatctgtt
 10 I Q N W N E T S A E K K E Q F L D H L L 280
 841 ggatcgcttc aagaagacac aagaagacac caccacatac agtttgatga actggaacac
 D R F K K T Q E D T T T Y S L M N W N T 300
 901 gggctctagtt tgggaccggt gcgtttttgt gaatgaaact cgcaagtgtg tctccaaatg
 G L V W D R C V F V N E T R K C V S K W 320
 15 961 gcagtgggga actgtttaca acttcctcta taaacacggt aaaattgcag atgagaccat
 Q W G T V Y N F L Y K H G K I A D E T M 340
 1021 gtgcagccga gagcttctcc atctgggtga tgctggtttt gccatcaaca ctccctatcc
 C S R E L L H L V D A G F A I N T P Y P 360
 1081 acttgttctg cctcctgtgc gtgaaactca cctcatcctc tcatttgact tcagtgtctg
 20 L V L P P V R E T H L I L S F D F S A G 380
 1141 ggaccacta gagaccatca gggccacagc agactactgc caacgccatg aaatcccctt
 D P L E T I R A T A D Y C Q R H E I P F 400
 1201 tcctgaggtg agcgaggatc agctgaagga atggggccaaa gccccagcaa gctgctatgt
 P E V S E D Q L K E W A K A P A S C Y V 420
 25 1261 cctcagaggg gaaacaggac ctgttgtcat gcactttact ctgttcaaca aagacaactg
 L R G E T G P V V M H F T L F N K D N C 440
 1321 tggagatgat attgagacat ggagaaaaaa atatgggaca gtaaaaactat ctgactcata
 G D D I E T W R K K Y G T V K L S D S Y 460
 1381 cacaccagac ctggtgagag atttgctgag ggtatccaag gagaatgtga agaaaaacaa
 30 T P D L V R D L L R V S K E N V K K N K 480
 1441 aattaatatc ctcatgaga tgaggaaagt ggctgggaat cctgggaaca tccaagagt
 I N I L S E M R K V A G N P G N I P R V 500
 1501 gaacaaggag gcctgcttgg gagacagagt gaaggatccc caaggctctc agactgtgga
 N K E A C L G D R V K D P Q G S Q T V E 520
 35 1561 gtttaagaaa tcccacaaca tatctaagga
 F K K S H N I S K D

The overall homology of MOP31/MOP8 with the catalytic unit of cytosolic
 lysophospholipase A2 and phospholipase B (stretches with more than 3 homologous
 40 amino acids are underlined) is less than 50 %.

The amino terminal fragment of mouse ePLA2 γ which is not in MOP31 is
MELSSGVCPATRLQEAEKAAVHKRSPKVLEALRKLNQADQAPVIAVLGS
GGGLRAHIACLGVLSEL (SEQ ID NO:17).

5 **Example 2- ePAD Genes**

Methods and results

Isolation of cat and dog ovarian RNA

Cat ovaries were recovered from spayed cats and stored at -80°C until use. The
pooled ovaries were ground then extracted with chloroform and Trizol. After
10 centrifugation, the aqueous phase was precipitated with isopropanol and centrifuged
again. The resulting pellet was washed with 75% ethanol and air dried for 15 min.
Finally, the recovered RNA was dissolved in water and stored at -80°C until use.

This protocol was repeated for the isolation of dog ovarian RNA. After isolation,
the electrophoresis of RNA through agarose gels containing formaldehyde was performed
15 (Figures 10 and 11).

RT-PCR and SMART-RACE for the full length cDNA sequence

Several pairs of primers were synthesized based on alignment comparison of all
the PAD gene and amino acid sequences in mice, rats and humans. Using the RT-PCR
technique, we identified the first strand of cDNA from the cat RNA templates and the
20 ePAD homologue gene fragment, an estimated 500 bp. The gene fragment was cloned
using the TOPO clone vector system and submitted for DNA sequencing. Results
indicated that it has 72 % similarity to the mouse ePAD. Based on this information, the
gene specific primers for 5' and 3' SMART-RACE were designed and synthesized, and
the 5', 3' terminal unknown sequences were cloned. In cat, a total of 2374 bp cDNA
25 (SEQ ID NO:10) were identified, and deduced 691 amino acids (SEQ ID NO:9) in length
with a predicted mass of 77.5 and pI of 5.10. The similarity between the cat and mouse
gene sequence was 77%, and the deduced protein similarity was 80%.

In the dog, a total of 2374 cDNA (SEQ ID NO:12) was identified and deduced
691 amino acids (SEQ ID NO:11) in length with a predicted mass of 77.7 and pI of 5.19.
30 The similarity between the dog and mouse gene sequence was 77%, and the deduced
protein similarity was 80% (Fig. 12).

The human and mouse ePAD genes are both present as single copy genes in their respective genome and the mouse and human proteins shares approximately 61% sequence identity and 76% conserved sequence identity. The deduced amino acid sequence of Epad is 691 amino acids in length with a predicted mass of 73.5 and pI of 5.5. Blast homology search demonstrates that EPAD is most similar to the peptidylarginine deiminase family of enzymes. PADs are post-translation modification enzymes which convert arginine residues on proteins to citrulline residues in the presence of calcium. Mouse Epad is approximately 40% identical to the four known mouse PADs.

Expression of the cat and dog ePAD truncated protein

The cat ePAD sequence was initially cloned and expressed using a PET 22 vector, resulting in a 69 kD protein. The construct was expressed in several different bacteria strains, including BL21(DE3), BL21 pLyS(DE3), Nova blue(DE3), Nova blue, Origami B(DE3) and Rosetta-gami B pLyS(DE3), resulting in selection of the BL21(DE3) and Origami B(DE3) as the ideal competent cells. After analysis by sequencing and Western blot, the recombinant protein was confirmed to be insoluble (Fig. 13).

The dog ePAD sequence was initially cloned and expressed using a PET 28vector, resulting in a 38 KD protein. The construct was expressed in several different bacteria strains, as described above for the cat ePAD experiments. The BL21(DE3) and Origami B(DE3) were selected as the ideal competent cells. After analysis by sequencing and Western blot, the recombinant protein was also confirmed to be insoluble (Fig. 14).

Nickel-NTA resins were used for cat ePAD truncate protein purification since the two kinds of reconstructed proteins were engineered with 6-His tags. In the process, denaturing conditions were used to extract the inclusion bodies from the bacterial lysate. The resulting products were dialyzed in PBS buffer (pH = 8.0) and lyophilized. Western blot analysis with anti-His antibody confirmed the results. The recombinant cat ePAD protein was concentrated after gel purification, and the recombinant dog ePAD protein was concentrated after Prep Cell purification (Fig. 15). The silver stained gel showed the cat and dog truncated proteins were a clear single band with an estimated concentration of 0.32 mg/ml and 0.35 mg/ml, respectively, as determined by a Bio-Rad RC/DC modified Lowry spectrophotometer assay.

The nucleic acid (SEQ ID NO:10) sequence and the deduced amino acid sequence (SEQ ID NO:9) of cat ePAD are as follows.

```

1   ggagatcggg gttatctgag gccgctgtgc tgaggaggagg ttgtcgaaca ggcgggccacc
61  ttggaggcct ttggcatgtc cttccagagc ctcgtccacc tgtccctgga cgcccccatc
5
      M S F Q S L V H L S L D G P I
121 catgccctct gcgtgctggg cgtggacatc tgcttggatc tcagtgggtg tgccccggag
      H A L C V L G V D I C L D L S G C A P E
181 aaatgcaagt catttacat cagtggctct ccgggggtct tggttgacat ccacaacacg
      K C K S F T I S G S P G V L V D I H N T
10 241 tccccgttaa tgaccaagga ggagatggcc acgaccgggt ggcctctgtc tgatcccatg
      S P L M T K E E M A T T R W P L S D P M
301 gatgtcctgg tgaatatggt ctccccagc tctgccccg atggtgacaa ggttctggtc
      D V L V N M V S P S S A P D G D K V L V
361 tcctactatc tgctgatga ggaagtccca gtggccacgg ctgtgctctg cctcactggc
15
      S Y Y L P D E E V P V A T A V L C L T G
421 attgtggtct ccctagacgt ggacatctac cgcagcgggc aagtcgaggt ggccagtgac
      I V V S L D V D I Y R S G Q V E V A S D
481 aagcaggcta agaaaaactg ggtctgggga cccagcggct ggggcgccat cctgcttgtg
      K Q A K K N W V W G P S G W G A I L L V
20 541 aactgcagcc ctaccgacaa gggccagagc atggacaaga agaccacgaa ggtgttcttt
      N C S P T D K G Q S M D K K T T K V F F
601 cccgaggaga tcaagagtct gtcccagatg accctgaatg ttcaagggcc cagctgcact
      P E E I K S L S Q M T L N V Q G P S C T
661 ttaaagaaat accggctggt tctccacacc tccaaggaag aggcagagaa ggcaagagtc
25
      L K K Y R L V L H T S K E E A E K A R V
721 taccggcccc aaaaagacag ctcgagcacc tttgaggtac tgctggggcc tggccagcac
      Y R P Q K D S S S T F E V L L G P G Q H
781 acctacacct ttgccccct cgagaaccac ctgaaggaaa ctttctacgt ggaagctgtg

```

T Y T F A P L E N H L K E T F Y V E A V
 841 gaattcccggt ctgccgactt ctcgggcctg atttcctact ctgtctccct ggtgcaagaa
 E F P S A D F S G L I S Y S V S L V Q E
 901 tctccggacc cgtccattcc agagaccctg gtatacagag acaccgtgggt attccgggtg
 5 S P D P S I P E T L V Y R D T V V F R V
 961 gccccttgcg tcttcgttcc cagcacgcag atgcctctag aggtttacct gtgcagagag
 A P C V F V P S T Q M P L E V Y L C R E
 1021 ctgcaggtcc aaggttttgt gaacacagtg atggagctga gtgagaagag taacagccaa
 L Q V Q G F V N T V M E L S E K S N S Q
 10 1081 gtggcatctg tctacgagga cccaatcgc ctgggcaggt ggctccagga cgagatggcc
 V A S V Y E D P N R L G R W L Q D E M A
 1141 ttctgttaca ccaggtctcc ccacaagacg ctctcctttg tctcgcacac ccctcggggcc
 F C Y T Q A P H K T L S F V L D T P R A
 1201 ctcacgctgg aagatttccc catgaaatac tcaactgagcc caggggtcgg ttacgtgatc
 15 L T L E D F P M K Y S L S P G V G Y V I
 1261 cagtgcaccc aggaccacag ggtggccagc atggattcca ttgggaacct gatggtgtcc
 Q C T Q D H R V A S M D S I G N L M V S
 1321 ccaccggtca aggtgggagg gaaagagtac cccctaggca gggtcctcat tggcagctgc
 P P V K V G G K E Y P L G R V L I G S C
 20 1381 ttttacccca gcaaggaggg ccgagacatg agtaaggccc tccgggactt cctctacgcg
 F Y P S K E G R D M S K A L R D F L Y A
 1441 cagcgagtcc aggccccggt ggagctcttc tcggattggc tgatggtcgg ccacatagac
 Q R V Q A P V E L F S D W L M V G H I D
 1501 gagttcatgt gctttatccc cacacaggac aagagtgagg gcgaaaaggg cttccggctg
 25 E F M C F I P T Q D K S E G E K G F R L
 1561 ctcctggcca gcccacgcgc ctgctacagg ctgtttgagg agaaacagag agagggtat
 L L A S P S A C Y R L F E E K Q R E G Y
 1621 ggagacgtga ctctgtttga ggaggtcaga gaggatcagc tcctctccaa cgggagggag

G D V T L F E E V R E D Q L L S N G R E
 1681 gccaatacca tccatcaact tctagctgat gaaaacatga gaaagcagaa cgaatatgtg
 A N T I H Q L L A D E N M R K Q N E Y V
 1741 gagaagtgca ttaacctgaa ccgtgacatc gtgaagaagg agctgggcct ggccgagagg
 5 E K C I N L N R D I V K K E L G L A E R
 1801 gacatcatcg acatcccaca gctcttctgc ctggagcagc tgacgaacgt cccctccgac
 D I I D I P Q L F C L E Q L T N V P S D
 1861 cagcagacgg ggaagttctt cgcgaggccg tacttccctg acctgctgca gatgatggtg
 Q Q T G K F F A R P Y F P D L L Q M M V
 10 1921 atgggcaaaa atctgggcat cccaagcct tttgggcccc agatcaaggg tacctgctgc
 M G K N L G I P K P F G P Q I K G T C C
 1981 ctggaaaaaa aggtctgcca gttgctagag cccctgggct tcaagtgcac cttcattgat
 L E K K V C Q L L E P L G F K C T F I D
 2041 gactttgact gctacctgac tgaagtcggg gacttctgtg cctgtgccaa catccgccgg
 15 D F D C Y L T E V G D F C A C A N I R R
 2101 gtgccctttg cttcaaatg gtggaggatg gtaccggagc cccaggccta gcgtgccage
 V P F A F K W W R M V P E P Q A
 2161 cctgccccag caggaacggc ccattaccac tcaccccctg cctctttggg aggtactgca
 2221 ccccttcctt acctgttcac cctgtccctc agtatccaga gcgatgggtca aactgccag
 20 2281 cctgaacccc tcagaaaacg gtcttgagct gagaacccat taaagtacga gctgttctga
 2341 atgcaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa

The nucleic acid sequence (SEQ ID NO:12) and the deduced amino acid sequence (SEQ ID NO:11) of dog ePAD are as follows:

25 1 ggttatttga ggctgctgtg ctgacctcgg gttgtcgtgt aggtctgagg gtagtcggct
 61 tggagagccg tgccatgtct ttccagagca tcatccacct gtccctggac agccctgtcc
 M S F Q S I I H L S L D S P V
 121 gtgccctgtg tgtgctgggc atggaaatct gcttggatgt caatgggtgc gccccggaga

R A L C V L G M E I C L D V N G C A P E
181 ggtgtgaggc gtttaccatc agcagctccc caggggtctc ggttgacctc cacggcacgc
R C E A F T I S S S P G V S V D L H G T
241 tcccggggac gggccaggag gggacggccg tgacccgggtg gcctctgtcg aatcccacgg
5 L P G T G Q E G T A V T R W P L S N P T
301 atgtgctggt gaagatgacc tgccccagct ccgccagtga cggagaccag gttctggtct
D V L V K M T C P S S A S D G D Q V L V
361 cctactatct gcccaatgag gacgccccg tggccacggc cgtgctccgc ctcaccggga
S Y Y L P N E D A P V A T A V L R L T G
10 421 ttgtggtctc cctggacgtg gacatctacc gcagtgggca ggtggagata gcgagtgaca
I V V S L D V D I Y R S G Q V E I A S D
481 agcaggccaa gaaaaactgg gtctggggtc ctagcggttg gggtgccatc ctgcttgtga
K Q A K K N W V W G P S G W G A I L L V
541 actgcagccc tgccgacaag ggccagatca tagaccagaa gaccaccaag gtgttctttc
15 N C S P A D K G Q I I D Q K T T K V F F
601 cagaggaaat aaagagcctg tcccagatga ccctgaacgt tacaagggcc cggcgccact
P E E I K S L S Q M T L N V T R A R R H
661 taaagaaata ccggctggtt ctccacacct ccgaggaaga ggcgaggaag gcccgagtct
L K K Y R L V L H T S E E E A R K A R V
20 721 atgggcccc aagagacagc tcgagcacct ttgaggtgct gctggggcca ggccagtgca
Y R P Q R D S S S T F E V L L G P G Q C
781 cctacacgtt cgccccctg gagaacaacc tgaaggaaac cttctacgtg gaagccatcg
T Y T F A P L E N N L K E T F Y V E A I
841 agttcccgtc tgccgacttc tcgggectga tctcctactc cgtctccctg gtggaagaac
25 E F P S A D F S G L I S Y S V S L V E E
901 ctcaggaccc gtccattcca gagaccctgg tgtacaaaga cacggtagtg tttcgggtgg
P Q D P S I P E T L V Y K D T V V F R V
961 cccctgcgt ctttattccc agcaccaga tgcctctgga ggtctatctg tgcagagagc

A P C V F I P S T Q M P L E V Y L C R E
 1021 tgcagggtcca gggttttgtg aacacgggtga tggagctgag cgagaagagt aacagccagg
 L Q V Q G F V N T V M E L S E K S N S Q
 1081 tggcatctgt ctatgaggac cccaaccgcc tgggcagggtg gctccaggac gagatggcct
 5 V A S V Y E D P N R L G R W L Q D E M A
 1141 tctgttacac ccaggcgccc cacaggacgc tctccttggt cctggacacc cctcgggtcc
 F C Y T Q A P H R T L S L V L D T P R V
 1201 tcacgttgga agacttcccc atgaaatact cactgagccc cggggttggc tacgtgatcc
 L T L E D F P M K Y S L S P G V G Y V I
 10 1261 agtgcaccaa ggaccacagg gtggccagca tggattccat tgggaacctg atggtgtccc
 Q C T K D H R V A S M D S I G N L M V S
 1321 cacctgtcaa ggtggaaggg aaagagtacc ccctaggcag gatcctcatt ggcagctgct
 P P V K V E G K E Y P L G R I L I G S C
 1381 tttaccccag caaggagggc cgagacatga gcaaggccct ccgggacttc ctctacgccc
 15 F Y P S K E G R D M S K A L R D F L Y A
 1441 agcgagtcca ggccccggtg gagctcttct cggactggct gatggtgggc cacatggatg
 Q R V Q A P V E L F S D W L M V G H M D
 1501 agttcatgtg cttcatcccc acaccggaca agagtgaggg tgaaaagggc ttccggctgc
 E F M C F I P T P D K S E G E K G F R L
 20 1561 tactggccag ccccagctcc tgccacagac tgttcgagga gaaacagaag gagggtacg
 L L A S P S S C H R L F E E K Q K E G Y
 1621 gggacatggc tctgtttgaa gaggtccggg aggaccagct cctttctaac gggcggcagg
 G D M A L F E E V R E D Q L L S N G R Q
 1681 ccaacacccat caatcagctt ctggctgaca aaaacatgag aaagcagaat gactacgtag
 25 A N T I N Q L L A D K N M R K Q N D Y V
 1741 agaagtgcac caacctgaac cgcgacatcc tgaagaagga gctgggcctg gttgagaggg
 E K C I N L N R D I L K K E L G L V E R
 1801 acatcatcga catcccgag ctcttctgcc tggagcagct gaccaacgtg ccctccagcg

D I I D I P Q L F C L E Q L T N V P S S
 1861 agcagaccgg caagttcttc gcgaggccct acttcctga cctgctgcag atgattgtga
 E Q T G K F F A R P Y F P D L L Q M I V
 1921 tgggcaagaa tctgggcata cccaagcctt ttgggcccc gatcaagggt acctgctgcc
 5 M G K N L G I P K P F G P Q I K G T C C
 1981 tggaagaaaa aatctgccag ttgctcgagc ccctgggctt caagtgcacc ttcattgatg
 L E E K I C Q L L E P L G F K C T F I D
 2041 acttcgactg ctacctgacc gaaatcgggg atttctgtgc ctgtgccaac atccgccggg
 D F D C Y L T E I G D F C A C A N I R R
 10 2101 tgccctttgc cttcaaatgg tggaggatgg tgcccagacc ccaggcctag tgccgccagt
 V P F A F K W W R M V P E P Q A
 2161 cctggcccgg tgcgagcggc cactgatcg cactccgca gcgttcaggt tccctgcctc
 2221 accctgtgta cctgtccctt tgtcctgagg gtggcaaatg ctgcccgtgg gaacctctc
 2281 aaacggcctt cagccaagaa ggggttacgg agtgtccatt aaagtcaag ctgttttgaa
 15 2341 tacaaaaaaa aaaaaaaaaa aaaaaaaaaa aa

Immunogenicity and Immunocontraceptive Potential of ePAD

An immunogenicity study was performed in mice using recombinant mouse
 ePAD. In two of five mice ePAD was immunogenic. The anti-ePAD antibodies
 20 generated recognized recombinant ePAD and reacted with mouse eggs in the ovary (see
 Figs. 17 and 18).

Mouse ePAD

Mouse ePAD (GenBank accession number NM_153106) amino acid (SEQ ID
 NO:13) and nucleic acid (SEQ ID NO:14) sequences are as follows:

25 MSFQNSLSLSLVNPTHALCMVGMEITLDISKCAPDKCKSFTIRGSPRILIHIS
 SSVIAGKEDTVVWRSMNHPTVALVRMVAPSPTVDEDKVLVSFYCPDQEVPTAT
 AVLFLTGIEISLEADYRDGQLDMPSDKQAKKKWMWGMNGWGAILLVNCSPNA
 VGQPDEQSFQEGPREIQNNLSQMNVTVEGPTSILQNYQLILHTSEEEAKKTRVYW

SQRGSSAYELVVGPNKPVYLLPTFENRRKEAFYVEATEFPSPSFSGLISLSLSLVEK
 AHDECIPEIPLYKDTV MFRVAPYIFMPSTQMPLEVYLCRELQLQG FVDSVTKLSE
 KSKVQVVKVYEDPNRQSKWLQDEMAFCYTQAPHKTVSLILDTPRVSKLEDFPM
 KYTLTPGSGYLIRQIEDHRVASLDSIGNLMVSPPVKAQGKDYPLGRVLIGGSFYPS
 5 SEGRDMNKGLREFVYAQQVQAPVELFSDWLMTGHMDQFMCFVPTNDKNNDQ
 KDFRLLLASPSACFELFEQKQKEGYGNVTLFEDIGAEQLLSNGRESKTISQILADK
 SFREQNTYVEKCISLNRTLKTELGLEDKDIILIPQLFCLEQLTNVPSNQQSTKLFA
 RPYFPDMLQIIVLGKNLGIKPKFGPKINGTCCLEEKVCGLLEPLGLKCTFIDDFDCY
 LANIGDVCASAIINRVPAFKWWKMTP (SEQ ID NO:13), and

10

1 ggactgctga cagtggctag cttggtaagc ccagccatgt cttttcagaa ctactcagc
 61 ctgtctctgg tcaatcccac ccatgccctc tgcattgtag gcatggaaat caccttggac
 121 atcagcaagt gtgcaccgga caagtgcaag tctttcacca tccgtgggtc cccaggatc
 181 ttgatccaca tctctagctc cgtcatcgct ggcaaagagg aactgtgggt ctggagggtca
 15 241 atgaaccatc ccacagtggc attggtgagg atggtggcgc ccagcccac tgtggatgaa
 301 gacaagggtc tggctccta cttctgtcct gaccaagaag tccccacggc cacagtgtg
 361 ctgtttctca cggcatcga gatctcctg gaggcagaca tctatcgaga tggacaactg
 421 gacatgcca gtgataagca agctaagaaa aaatggatgt ggggtatgaa cggctgggga
 481 gccatcctgc ttgtgaattg tagcccta at gctgtggg cc agcctgatga acagtccttt
 20 541 caggagggcc ccagagaaat acagaacaac ctgtctcaga tgaatgtaac tgtggagggc
 601 cccaccagca tcctacagaa ttaccagtgt atcctacata cctccgaaga agaggcgaag
 661 aagacaagag tctactggc tcagagaggc tcctctgct atgaactggg ggtgggaccc
 721 aacaagcctg tctatctcct gcctaccttt gagaaccgta ggaaagaggc tttctacgta
 781 gaagccacgg aattcccatc tccagcttc tggggcctga tctccttgct actctcccta
 25 841 gtagaaaagg ctacagacga gtgcattccc gagattccgc tctataagga tacagtgatg
 901 ttccgggttg cacttatat cttcatgccc agcaccaga tgcctctaga ggtttacctg
 961 tgcagggagc tacagctgca aggccttctg gactcagtga ccaagctgag cgagaagagc
 1021 aaagtgcagg tggtaaaggc ctatgaggac cccaaccgcc agagcaagtg gctccaggac
 1081 gagatggctt tctgctatac tcaggctcct cacaagacgg tgtcattgat ccttgacacc
 30 1141 ccaaggggtt ccaagctgga agacttcccc atgaaataca cactgacccc tggctctggc
 1201 tacctgatcc gacaattga ggaccaccgg gtggctagcc tggattccat cgggaacctg
 1261 atggtatctc cgcctgtcaa ggctcagggc aaagactacc ctctagggag ggtcctcatt
 1321 ggtggcagct tttaccccag ctctgagggc cgggacatga acaagggcct gcgagaattc
 1381 gtgtatgccc agcaggtgca ggcccctgtg gaactctct cggactggct gatgaccggt
 35 1441 cacatggatc aattcatgtg ctttgtccct accaatgata aaaacaacga ccagaaggac
 1501 ttccgcctgc tgctggccag cccagtgcc tgctttgagc tgttcgaaca gaagcagaag
 1561 gaaggctatg ggaacgtgac cctgtttgaa gacattggag cagaacagct cctttcta at
 1621 gggagggaga gcaaaactat ttcccaaatc ctggctgaca agagtcttcg agagcagaac
 1681 acctatgttg agaagtgtat cagcctgaac cgcaccctcc tgaagacaga actgggattg
 40 1741 gaggacaagg acatcatcct gatccgcgag ctcttctgcc tggagcagct gacgaatgtc
 1801 ccctccaacc agcagagcac caactcttc gcgaggccgt acttccccga catgctgcag
 1861 ataactgtgt tgggcaagaa ccttggaatc cccaagccct ttggggccaa aatcaatggc
 1921 acctgctgcc tagaagagaa agtgtgtgga ttactggagc cctgggtct caagtgcacc
 1981 ttcatgtgat attttgactg ctacctggcc aacataggg acgtctgtgc cagtgccatc
 45 2041 ataaacaggg tgccatttgc attcaagtgg tggaagatga ccccataaac ccctggccct
 2101 ggcacggcca gtccgcgcca gtacggatgg cttttgcat agatagtagt ggggtgcgagc
 2161 gttgtgtgtg cactgggttg aaggagcgg gagctgggag ttagggctct tcacatctac
 2221 cagcttgaca cttctggagg ggaaaaggga aaagagcgc tatgtaaaca aattgccata

2281 gagccaataa agcatggtat tctgaataca aaaaaaaaaa aaaaaaaaaa aaaaaa (SEQ ID NO:14)

Example 3- MOP26

5 Using techniques described above, another novel mouse gene and protein was identified and characterized. The other novel protein, called MOP26 herein, is completely unknown with Riken DNA and no domain homology with other proteins in the databases has been found. The MOP26 amino acid sequence (SEQ ID NO:15) and cDNA sequence (SEQ ID NO:16) are as follows:

10

```

1 gagctttgaa aagcgagacc agaccggctg ccaagcagcg ccccgtcctc ctgcttgcag
61 tatggcatcc cacacggctg atgctgacgc caagccagac tctgactctc agaagctgct
    M A S H T A D A D A K P D S D S Q K L L 20
121 taacgtcctg cctgtgtccc tgagacttcg cactcggccc tgggtggttcc caattcagga
15 N V L P V S L R L R T R P W W F P I Q E 40
181 agtcagtaat cctctggtgc tctacatgga agcctgggtg gcagaaaggg taataggcac
    V S N P L V L Y M E A W V A E R V I G T 60
241 agaccaagcg gaaatctcag aaatagagtg gatgtgccaa gccctgctga cagtggactc
    D Q A E I S E I E W M C Q A L L T V D S 80
20 301 cgtcaactct gggaacctag ctgaaatcac catctttgga caaccagtg cacagacgcg
    V N S G N L A E I T I F G Q P S A Q T R 100
361 aatgaaaaat attctcttga acatggcggc ctggcacaaa gaaaacgaac tccaaagagc
    M K N I L L N M A A W H K E N E L Q R A 120
421 tgtgaagggtg aaagaagttg aagagttctt gaaaattcgt gcctcctcaa tcctaagcaa
25 V K V K E V E E F L K I R A S S I L S K 140
481 gttaagtaag aaagggtctaa aactggctgg ctttccgctt ccgctggagg gaagagaaac
    L S K K G L K L A G F P L P L E G R E T 160
541 acagatggag tottaaataca ctacacgtat tttcttaccg gccttcacga gcacctcctg
    Q M E S 164
30 601 tttatttttc ttgcaaata gcaatttgta gaactattcc aatataaaaa attttttttg
661 gtatgctggt ggtttgatca catacagtaa aatttataaa cactcg

```

35 MOP26 has been demonstrated herein to be located on chromosome 9 and has 3 exons and 2 introns. Most of ESTs were found in fertilized mouse eggs or mouse embryos.

In cloning MOP26 and MOP31 (see above) mouse oocyte cDNA was reverse transcribed from RNA from 350 mouse oocytes and used for PCR. The PCR products for each gene are shown below. The PCR product for MOP31 was around 1500 bp and the
 40 PCR product for MOP 26 was about 500 bp. These bands corresponded to predicted sizes of MOP31 and MOP26. These products were put in TOPO vector, confirmed by sequencing, and the PCR products were visualized (see Fig. 19).

Both new sequences were cloned into expression vectors. MOP26 was cloned in pET28 expression vector, expressed in HMS174 cells and purified on the affinity column and by PrepCell. The purified proteins are demonstrated in Figures 20 and 21.

5 Purified recombinant MOP26 was injected into rats for antibody production. A high titer reagent was obtained. This reagent recognized recombinant MOP26 as well as endogenous MOP26 obtained from mouse eggs (see Fig. 22).

10 The antibody to MOP26 was employed to stain isolated mouse eggs. Both zona pellucida intact as well as zona free mouse eggs stained with the antibody. The entire cytoplasm was fluorescent as well as zones of concentrated fluorescence reminiscent of the microvillus surface (Fig. 23).

The antibody to MOP26 was also used to stain the ovary. MOP26 localized only to the egg and not other ovarian tissues (Fig. 24).

15 In live eggs, the antisera to MOP26 stained the perivitelline space of zona intact eggs as well as the egg cytoplasm (Fig. 25). Trypsin treated dezonulated eggs did not stain.

MOP26 was compared to MOP29, MOP31, MOP32, and MOP30 by mapping the proteins electrophoretically (see Fig. 26), which were used for core sequencing. The proteins were cored and analyzed by tandem mass-spectrometry. MOP29, MOP30, and MOP32 are known proteins which are not egg or ovary specific.

20 An electrophoretic analysis of some of the PCR products was performed with mouse oocyte cDNA. For PCR, total RNA was extracted from 500 mouse eggs and reverse transcribed to make mouse egg cDNA. Specific PCR products were cloned into TOPO vector and sequenced, and subjected to electrophoretic analysis. Groups included MOP26, Control, and MOP31 (see Fig. 27).

25 Next, MOP26 expression was analyzed in HMS174 and BL21. To that end, recombinant MOP26 with a His-tag at the C-terminus was cloned into the pET-28b+ vector and expressed in HMS174 and BL21 cells (see Fig. 28).

30 The specificity of MOP26 expression was determined by Western blot analysis of a series of cells and tissues. The cells and tissues included oocyte, ovary, testes, thymus, liver, spleen, kidney, intestine, stomach, lung, and heart. To this end, 50 µg of total

protein from these cells and tissues and a total protein extract from 35 mouse eggs were probed with polyclonal antiserum directed against MOP26 (see Fig. 29).

Next, a multi-tissue RT-PCR analysis of MOP26 expression was performed in a variety of cells and tissues (heart, lung, liver, pancreas, kidney, testis, ovary, spleen
5 thymus, and oocyte). To that end, 70 ng of RNA from different tissues was used in RT-PCR to amplify MOP26 and the control β -actin (see Fig. 30).

The MOP26 transcript was registered in two tissues – ovary and thymus. There was no evidence of MOP26 transcript in other tissues. MOP26 protein was identified only in mouse oocytes. Based on RT-PCR results that show no MOP26 transcript in
10 heart tissue, the two positive bands in the western blot could be due to non-specific reactivity with pre-immune sera.

Next, MOP26 expression was examined during various stages of oocyte maturation and early development (see Fig. 31). The stages included gv, m2, pn, 2-cell, 4-8 cell, and blastocyst. The results demonstrate that MOP26 specifically stains oocytes
15 at various developmental stages, including primary and secondary follicles. In zona intact non-permeabilized eggs, MOP26 was localized to the perivitelline space. In permeabilized zona free eggs, MOP26 was localized to the egg cortex and showed an asymmetrical distribution. During the early stages of embryo development, it was demonstrated that MOP26 remained at the egg cortex and retained its asymmetrical
20 distribution. By the pronuclear stage the distribution became more symmetrical and MOP26 retained its subcellular localization in the cortex of the egg. In the blastocyst, MOP26 appeared uniformly distributed in the trophectoderm.

Phosphorylation of MOP26

Next, it was determined whether the potential MOP26 phosphorylation sites could
25 be phosphorylated. The putative sites of MOP26 include:

Protein Kinase C Phosphorylation Site:

16-18 SqK

26-28 SIR

142-144 SkK

30 Casein Kinase II (CKII) phosphorylation site:

66-69 SeiE

160-163 TqmE

N-Myristoylation site

59-64 GTdqAE

93-98 GQpsAQ

5 145-150 GLklAG

In vitro phosphorylation studies were performed to determine if MOP26 could be phosphorylated. CKII is a ubiquitous enzyme that preferentially phosphorylates Ser/Thr residues in acidic stretches of substrates. CK II plays important roles in the initiation of DNA replication and the regulation of transcription. The in vitro studies described herein suggest that DNA-binding sperm proteins function as potent activators for CK II in fertilized eggs because of the high content of basic amino acids and plural oligo-Arg clusters. The data demonstrated that MOP26 does have potential phosphorylation sites for CKII and PKA (cat) (see Figure 32).

Further analysis indicated that MOP26 showed 25% identity and 54% similarity with developmental pluripotency associated protein. Developmental pluripotency proteins play the major role in maintaining pluripotency of embryonic stem cells. They allow propagation of stem cells undifferentiated and retain stem cell capacity for multilineage differentiation. Without wishing to be bound by any particular theory, MOP26 could be another pluripotency factor which is activated upon fertilization.

It was determined that the ORF for MOP26 is equal to 492, that there are 164 Amino Acids, that the predicted MW is 18.5, and that the predicted pI is 5.8. Furthermore, MOP26 produced five prominent peptides of unknown protein with Riken cDNA from mus musculus. An EST search showed that this cDNA is present mostly in fertilized mouse oocytes, although there were also detectable levels in thymus. Analysis of the protein sequence showed no glycosylation sites, but a possible four serine and one threonine phosphorylation sites, three myristoylation sites and no transmembrane domains.

The terms MOP26 and MOEP19 are used interchangeably herein.

Conclusions and Summary- MOP26 was found to be a novel protein and highly specific for fertilized eggs and embryos in mouse. It has RIKEN DNA. MOP26 was

cloned using mouse oocyte cDNA and its recombinant protein was expressed and purified. Mouse polyclonal serum was developed and proved to be highly specific for MOP26. MOP26 was localized to perivitelline space and cortical region in mouse eggs. MOP26 was demonstrated to persist in blastomeres at the cell cortex and found to segregate to the trophectoderm in blastocyst. MOP26 was found to be present in oocytes at various stages of the follicle development. Real time PCR and RT-PCR results confirmed the presence of MOP26 in oocytes, ovary and thymus at the RNA level; however, only in oocytes was MOP26 protein detectable.

Other methods which were used but not described herein are well known and within the competence of one of ordinary skill in the art of cell biology, molecular biology, and medicine.

The invention should not be construed to be limited solely to the assays and methods described herein, but should be construed to include other methods and assays as well. One of skill in the art will know that other assays and methods are available to perform the procedures described herein.

Headings are included herein for reference and to aid in locating certain sections. These headings are not intended to limit the scope of the concepts described therein under, and these concepts may have applicability in other sections throughout the entire specification.

The disclosures of each and every patent, patent application, and publication cited herein are hereby incorporated herein by reference in their entirety.

While this invention has been disclosed with reference to specific embodiments, it is apparent that other embodiments and variations of this invention may be devised by the previous description of the disclosed embodiments is provided to enable any person skilled in the art to make or use the present invention. Various modifications to these embodiments will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other embodiments without departing from the spirit or scope of the invention. The appended claims are intended to be construed to include all such embodiments and equivalent variations. Accordingly, the present invention is not intended to be limited to the embodiments shown herein but is to be

accorded the widest scope consistent with the principles and novel features disclosed herein.

CLAIMS

What is claimed is:

- 5 1. A pharmaceutical composition comprising a pharmaceutically-acceptable carrier and at least one novel egg protein, or a homolog, fragment or derivative thereof, wherein said protein is capable of inducing an immune response useful for inhibiting conception in a subject.
- 10 2. The pharmaceutical composition of claim 1, wherein said novel egg protein comprises an amino acid sequence selected from the group consisting of SEQ ID NOs:5, 7, 9, 11, 15, and 17.
- 15 3. The pharmaceutical composition of claim 2, further comprising at least one known egg protein, wherein said known egg protein is capable of inducing an immune response useful for contraception.
- 20 4. The pharmaceutical composition of claim 3, wherein said known egg protein has the amino acid sequence of SEQ ID NO:13, or a homolog, fragment or derivative thereof.
- 25 5. The pharmaceutical composition of claim 2, further comprising an adjuvant.
6. The pharmaceutical composition of claim 2, wherein said pharmaceutical composition comprises at least two different proteins, or homologs, fragments, or derivatives thereof, wherein said at least two different proteins comprise an amino acid sequence selected from the group consisting of SEQ ID NOs:5, 7, 9, 11, 15, and 17.
- 30 7. The pharmaceutical composition of claim 6, further comprising at least one known egg protein.

8. A contraceptive vaccine, said vaccine comprising the pharmaceutical composition of claim 5.

5 9. A method for inhibiting conception in a subject, said method comprising administering to said subject a pharmaceutical composition comprising a pharmaceutically-acceptable carrier and at least one novel egg protein, or a homolog, fragment or derivative thereof.

10 10. The method of claim 9, wherein said at least one novel egg protein comprises an amino acid sequence selected from the group consisting of SEQ ID NOs:5, 7, 9, 11, 15, and 17.

15 11. The method of claim 10, wherein said pharmaceutical composition further comprises at least one known egg protein, or a homolog, fragment, or derivative thereof.

12. The method of claim 11, wherein said known egg protein comprises the amino acid sequence SEQ ID NO:13.

20 13. The method of claim 12, wherein said subject is selected from the group consisting of cattle, pigs, horses, sheep, goats, birds, cats, and dogs.

25 14. The method of claim 13, wherein said pharmaceutical composition is administered via a route selected from the group consisting of topical, oral, intramuscular, and intravenous.

15. The method of claim 14, wherein said composition is administered as a controlled-release formulation.

30 16. The method of claim 9, wherein at least two different proteins, or homologs, fragments, or derivatives thereof are administered.

17. The method of claim 16, wherein the proteins, or homologs, fragments, or derivatives thereof are administered in a mixture of approximately equimolar concentrations.

5 18. The method of claim 16, wherein the amount of proteins, or homologs, fragments, or derivatives thereof administered is between about 0.001 mg/kg body weight and about 100 mg/kg body weight.

10 19. The method of claim 9, wherein at least three different proteins, or homologs, fragments, or derivatives thereof are administered.

20. An isolated nucleic acid comprising a nucleic acid sequence encoding a protein of claim 1, or a homolog or fragment thereof.

15 21. The isolated nucleic acid of claim 20, wherein said nucleic acid sequence shares at least 70% identity with a nucleic acid sequence encoding a protein, or homolog or fragment thereof, wherein said protein is selected from the group of proteins having the amino acid sequences of SEQ ID NOs: 5, 7, 9, 11, 15, and 17.

20 22. The isolated nucleic acid of claim 20, wherein said nucleic acid sequence is selected from the group of sequences having SEQ ID NOs: 6, 8, 10, 12, and 16.

23. A vector comprising the isolated nucleic acid of claim 20.

25 24. A recombinant cell comprising the isolated nucleic of claim 20.

25. A recombinant cell comprising the vector of claim 23.

30 26. A pharmaceutical composition comprising at least one isolated nucleic acid of claim 20.

27. An antibody that specifically binds with a novel egg protein of claim 2.

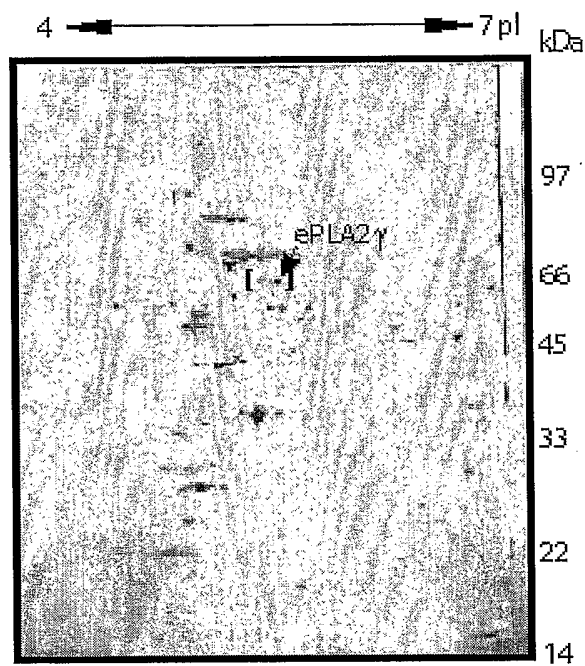
28. The antibody of claim 30, wherein said antibody is selected from the group consisting of a polyclonal antibody, a monoclonal antibody, a chimeric antibody, and a
5 synthetic antibody.

29. A method for inhibiting conception in a subject, said method comprising administering to said subject a pharmaceutical composition comprising at least one
10 antibody of claim 28.

30. The method of claim 29, wherein said antibody inhibits the function of said
novel egg protein.

31. The method of claim 29, wherein said antibody inhibits fertilization.
15

32. A kit for inhibiting conception in a subject, said kit comprising the pharmaceutical composition of claim 5, an applicator, and an instructional material for the use thereof.

**FIG. 1**

gactcaatattaggggtctgagcgccgcccgggaggtatggatttctgcttgcctctgct
 ccttgacatgcttgaggttccagtcctgaagtccttgggtgataaacacagcagcattggaaacta
 M E L
 AGCTCTGCGGCTCTGCGCTGCCACTAGACTCCAGGAGCCAGAAAAGCCAGCTGTGCACAAA
 S S G V C P A T R L Q E A E K A A V H K
 CGCAGTCCCAAGGTTTGGAGGCTCTACGAAAGCTCAACATCCAAGCTGACAGGCTCCCA
 R S P K V L E A L R K L N I Q A D Q A P
 GTCATTGCTGTCTTGGGCTCTGCGGCGGGCTGCGGCGCCACATCGCTTGTCTTGGTGTG
 V I A V L G S S G G L R A H I A C L G V
 CTGAGTGAAGCTGAAAGACTTGGCTGTGGATGCTGTCAATACCTGCGAGCGGCTCTCT
 L S E L K E L G L L D A V T Y L A G V S
 GGGTCCACTTGGGCACTGTCTTCACTCTACACCAAGATGGGAANTATGGAGGGGATAGAA
 G S T W A L S S L Y T K N G N M E G I E
 GAGAGCTGAALCATCGGTATGAGAAATGAGTGGGACTTTCTATGAGAGCTTGGAGAA
 E E L K H R Y E K N E W D F H E S L E K
 GGCATCCAGGCTCAAGAGGGGAAATTACTTCCCTGACTGACTTTTGGGCTATTATTT
 A I Q A S K R E N Y S L T D F W A Y L I
 GTTCCAGGCAATCAGAGAACTTCAGGATTGGAATTTGTCCAGCTCTAAGAGAGCAAGTG
 V S R Q I R E L Q D S N L S S L K K Q V
 GAAGAGGAGTGTCTGCGCTTCCCTATCTTTGCGAGCCATTGATGAGGAGCTTCTGCTGTAT
 E E G V L P Y P I F A A I D E D L I A D
 TGGAGGGAGAGGAAACTCAGAACTTCCCTGCTTGAATTCCTCTCATCATGCTGGCTAC
 W R E R K T Q N S W F E F T P H H A G Y
 CCGTCACTTGGGCTTATGTCCCTCTCAGAGCTTGGAGCAGATTTGAGAAATGGGAA
 F A L G A Y V P I T E F G S R F E N G K
 CTGGTTAAATCTGAGCGCTGAGAGAGATTTGACTTTCTGAGAGGTTTATGGGAGAGTGT
 L V K S E P E R D L T F L R G L W G S A
 TTTCTGATATTAAAGAAATTAAGAAATTTATTTTGAACACTTTCAGGACCCCTTTTGA
 F A D I K E I K N Y I L N Y F R N P F G
 AAATGAAGTTTATAGAGAGCAGTGAACACTCAGAGCAGCAGCAGCAGCAGCAGCAGCAG
 K L K F I E G P V T Y S E A P R M N V D
 GCATGCTCTTGGATTTAGTGAATGGCTTATTTCAAGATATGAATGACCCCGAGCATCAG
 A N L L D L V M A Y F T D M N D P S I K
 GATAAGCTCTGTGCGCTTACAGAGCTCTGCTGCTGAGAGCAGATGAATTTGGCATAGAG
 D K L C A L Q Q A L G T S T D E F G I E
 ATGCTGAGATCATCAGAAATGGAATGAGAGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAG
 M A E I I Q N W N E T S A E K K E Q F L
 GACCATCTGTTGGATGCTTCAAGAGACACAGAGAGACACACACACACACACACACACACAC
 D N L L D R F K K T Q E D F T T Y S L M
 AACTGGAACACCGGCTCTAGATTGGGACCGTTCGCTTTTGTGAAATGAAACTGCGAGCTGT
 N W N T G L V W D R C V F V N E T R K C
 GTCCTCAATAGGCAAGTGGGAACTGTTCACACCTTCTCTATAAACAGGTAATTTGCA
 V S K W Q W G T V Y N E L Y K H G K I A
 GATGAGACCATGTGAGCCGAGAGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 D B T M C S R E L L H L V D A G F A I N
 ACTGCTCTTCCACTTGTCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 T F Y P L V L P P V R E T H L I L S F D
 TTCAGTCTGCGGAGCCCACTAGAGACCATCAGGSCCAGAGCAGACTACTGSCCAAGCCCAT
 F S A Q D P L E T I R A T A D Y C Q R H
 GAATCCCTTTTCTGAGGTCAGCGAGGATCAGTTGAGGAGATGGGCAAGGCCCGAGCA
 E I P F P E V S E D Q L K E W A K A P A
 AGCTGCTATGCTCTCAGAGGGGAAACAGGAGCTGTTGTCATGCACTTTACTCTGTTCAAC
 S C Y V L R G E T G P V V M H F T L F N
 AAAGACAACCTGTGAGATGATATTGAGACTGAGCAAAAATATGGGACAGTAAAGACTA
 K D N C G D D I E T W R K K Y G T V K L
 TCTGACTCATAACACACAGACCTGCTGAGAGATTTCTGAGGCTATCCAGGAGAGATGCTG
 S D S Y T P D L V R D L L R V S K E N V
 AAGAAAACAAATTAATATCTCTGCTGAGATGAGAAAGTGGCTGGGAATCTCTGGGAGC
 K K N K I N I L S E M R K K V A G N F G N
 ATCCCAAGAGTGAACAGGAGGCTGCTTGGAGACAGAGTAAAGGATCCCAAGGCTCT
 I P R V N K E A C L G D R V K D P Q G S
 CAGACTGTGGAGTTTAAAGATCCACACATATCTAAGGATTAAGGATGCTTGTGCACTA
 Q T V E F K K S H N I S K D *
 cctggagaggttggcatctatttagtggactcaaccactgtgtctctctccaaagtccacagt
 aagctgtctcttctgtgaagaaagactgcagtttccagggtaactgtgggactactgtcttta
 ccaggggcccagtgtaaacctggcctgatgtctctgccactcaaaaattgcatttggc
 ttagctattaccacttgagcaagctcccgagactagatgggtgtctttgatagcattatt
 taggggcaactgggtccaccagagctaaaagagggcaaggtatgtctctcagatacctcagggg
 accctaaagctagagcaactccttgatcaatcaagccacccctgtccacacccctgtg
 gttaaaattgaatgttagtgattgagcccaagctctcaaaattttctctctctctctctctct
 tgcctcattttatcaacccagctcttactaaagtctctttttttctctataaaaggatagacott
 agctcaaatcaatgagccatcaacttcagagcaatbaaaaaaacccagcaactctgtgagag
 acatgttctgtatatactgttgatctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt
 tct
 tcatgttaatatcaagggcaagttctctctcaatgaatctcttaalcatctctatgggaataa
 atgtgtctctgt
 atgtgtctctgt

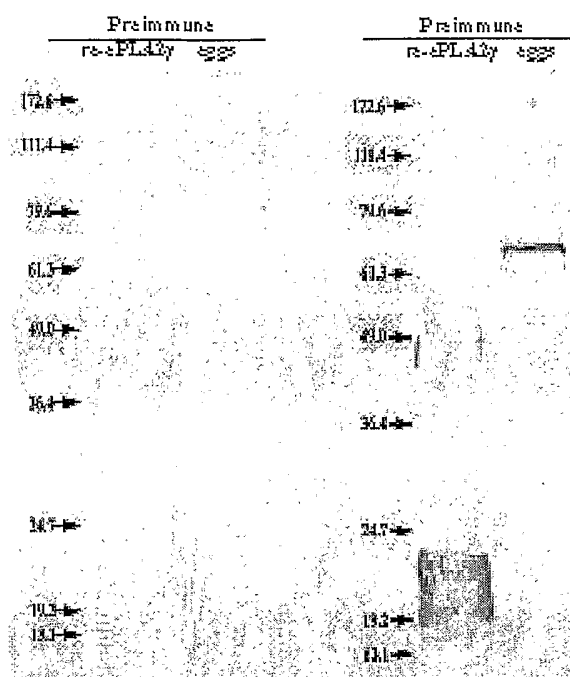
FIG. 2

ePLA2 γ	1	HELESEEDIPATYQKREKAAVHAKSPKQLSELEKNTLSDMAPVAVLGGGKIKHLSN	50
cPLA2 γ -human	1	HCSESEVSIIPGLQCESEKRAVERPRLHYLSELEKRIPEKDEAPVAVLGGGGGIRAHLC	59
ePLA2 γ	61	GGVLSHYRELHLIDAVTFLAGVSGSTWELSSLTQNEKHHGIKKELKHVYKEMEWLHES	120
cPLA2 γ -human	60	GGVLSGHEDELIDAVTFLAGVSGSTWELSSLTQNEKHHGIKKELKHVYKEMEWLHES	119
ePLA2 γ	121	LSKATQASRENTSLTDVEYLIQSDIKRELQDEKLSLEKQVEKQVLPYPIFAIDDEL	180
cPLA2 γ -human	120	LQNTIQASRENTSLTDVEYLIQSDIKRELQDEKLSLEKQVEKQVLPYPIFAIDDEL	179
ePLA2 γ	181	LADWEKRWTONWEFFTPHHAGCPALGKAPLITFEGSGEENKPLVWSEPERDLTELEGDW	240
cPLA2 γ -human	180	OPSDWEKRPETWFFTPHHAGCPALGKAPLITFEGSGEENKPLVWSEPERDLTELEGDW	239
ePLA2 γ	241	SSKADIWESENYILNYSNPFGKLEIEGPIYSEAPKRWDAHLLDLPYHFTDHNIC	300
cPLA2 γ -human	240	SSALGITEVISEKIFDQENLTLEGLVRAVANKSIGHLIFAPLEKQES--EQGEHE	296
ePLA2 γ	301	SIADKLCALQQLSTDEKFGIEHAKIDWREYKAKKEDQLDHLIDPEKATQEDTITY	360
cPLA2 γ -human	297	PPEDS-----SEKPEH--TWLTSALSKWRTSLKDEG-----PHEDPEKKSLS	339
ePLA2 γ	361	SLHMMNTGLVWRQVEVNETRKQKRWQNGTIVNELYHGMTRHICSSKELLHLWDAGF	420
cPLA2 γ -human	340	MLHD-----EUKKGIKSKKWEKGTWELYNKGSFQKIKSSKELLHLWDAGL	388
ePLA2 γ	421	SLNTIPPLVLPFFRETHLILSEDSAGDELETRATATYQKHEIPEPEVSDQLKEWEN	480
cPLA2 γ -human	389	SLNTIPPLVLPFFRETHLILSEDSAGDELETRATATYQKHEIPEPEVSDQLKEWEN	448
ePLA2 γ	481	SPASGNIAGGTHKVVHHPLEMLDAGEEDKEMEDTDDEKADQLDLDVLLIALAG	540
cPLA2 γ -human	449	SPASGNIAGGTHKVVHHPLEMLDAGEEDKEMEDTDDEKADQLDLDVLLIALAG	508
ePLA2 γ	541	KWVGGKNTLSEHKKPEKPGNIPRVKKEACLGDEVKDPOGSQTVKFKSHNISKD	594
cPLA2 γ -human	509	KWVGGKNTLSEHKKPEKPGNIPRVKKEACLGDEVKDPOGSQTVKFKSHNISKD	525

FIG. 3

1	HELSSGVTPFATLQAKKAAHLEKFPKLSALGSEKIDALDQAPWIAVLGSAGGIGRAHIAK	50
1	-HGSESEVPIPGLOKREKASUERKELHULKALKRIKADSRPWVNLGSAGGIGRAHIAK	59
61	LGULSELKELGLIDAVTLAGVSGSTVALSSLYTANGVHNGIPEELKHELEKNEWLPHEK	120
60	LGULSEAKEDGLIDAVTLAGVSGSTVALSSLYTANGVHNGIPEELKHELEKNEWLPHEK	119
121	LEKATQAKSEKEMYSLIDFVAYLIVKQIKRELQDENLSELKMDJEEGULPYPIFAIDENL	180
120	LEKATQAKSEKEMYSLIDFVAYLIVKQIKRELPESELKMDJEEGULPYPIFAIDENL	179
181	LADDERKKTQMSWEEETPHHAGLPALEGLTPVTEFSGSEKEMKLVKSEPEKDLTELEGLV	240
180	OPSDERKAPETWEEETPHHAGLPALEGLTPVTEFSGSEKEMKLVKSEPEKDLTELEGLV	239
241	ESSEKADIKEIANYILNFFENPFGMLKFIKGPVLYSEAPRNVIAHLLDQRAYETDHWDE	300
240	GSALGMEIVSEYIFDQELTLKGLWRAVAKKESIGHLIFARLKLDES--SQGEHP	295
301	SITKALCALQVALSTEDREPGIEHAKIOMREYSAKESKQELHALLDEFKRTQEDTITY	360
297	PPKES-----GSEPEH--TWLDEKESKQRTSLKEDSA-----PHDEPKKESLS	339
361	SLHWMTGLVMDRCVPMREKELKPVVWSTLYNDEKAKHGMFQSTHOSKELLHLVDAEF	420
340	NLHD-----FVLMGILKSTWEGTTHWLYKHGGKDKIHSEKELHLVDAEL	388
421	SLNTIEPLVLPPEKELTILSEDEKAGDLELIRATDYQKHSPEPELSEDQKEMK	480
389	SLNTIEPLVLPPEKELTILSEDEKAGDLELIRATDYQKHSPEPELSEDQKEMK	448
481	SPACAVLAGETSPVWHEPLEMDKAGGDIKEDTQDQELDQKILVWVLLKALAK	540
449	SPACAVLAGETSPVWHEPLEMDKAGGDIKEDTQDQELDQKILVWVLLKALAK	508
541	ENVMKKNINILESEKRNAGPKNIPVNRKACLGDENVDPQGSQTVKFKSHNISM	594
509	KTVSEKAKKILRELHNAGLYPEDSERSCLLA-----	525

FIG. 4

**FIG. 5**

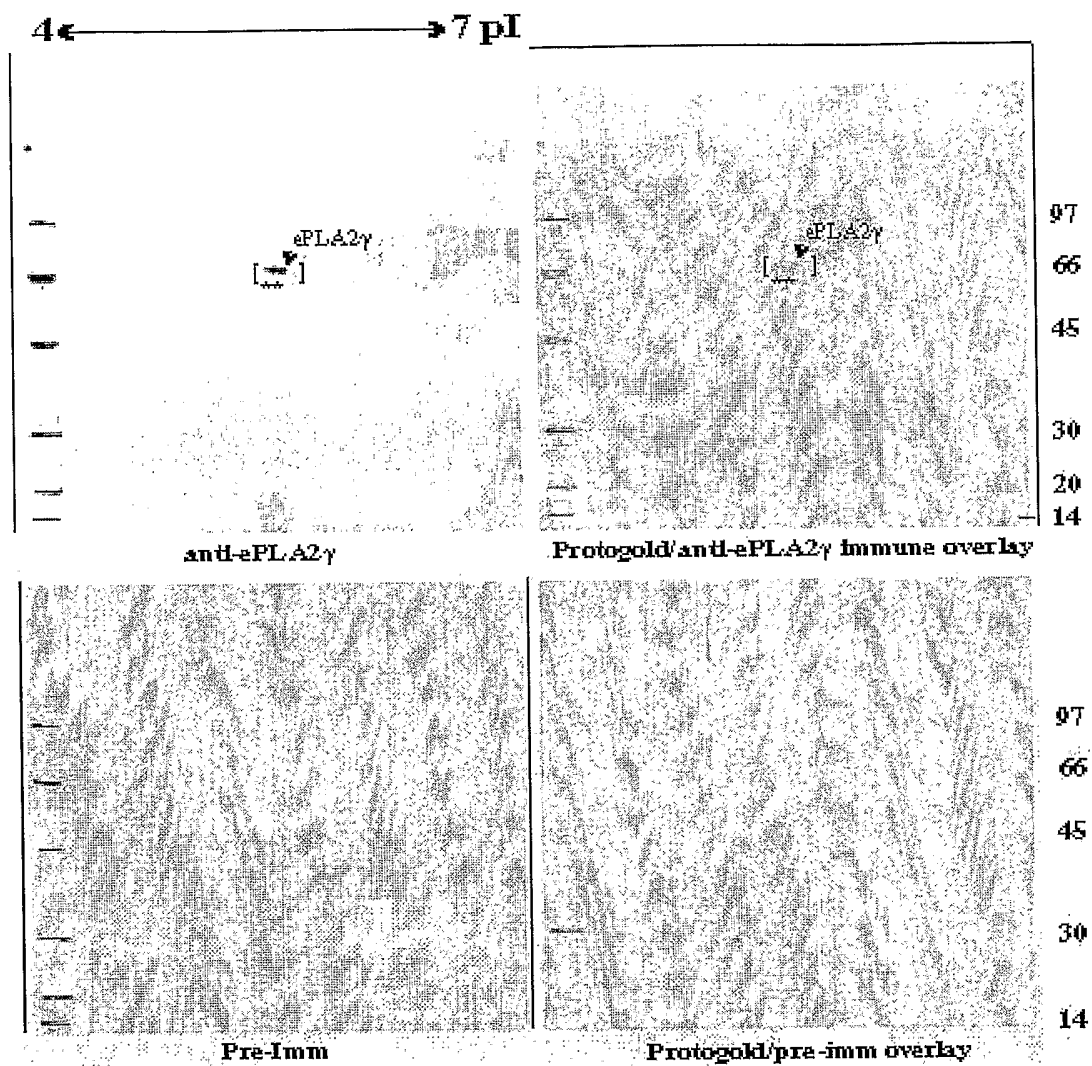


FIG. 6

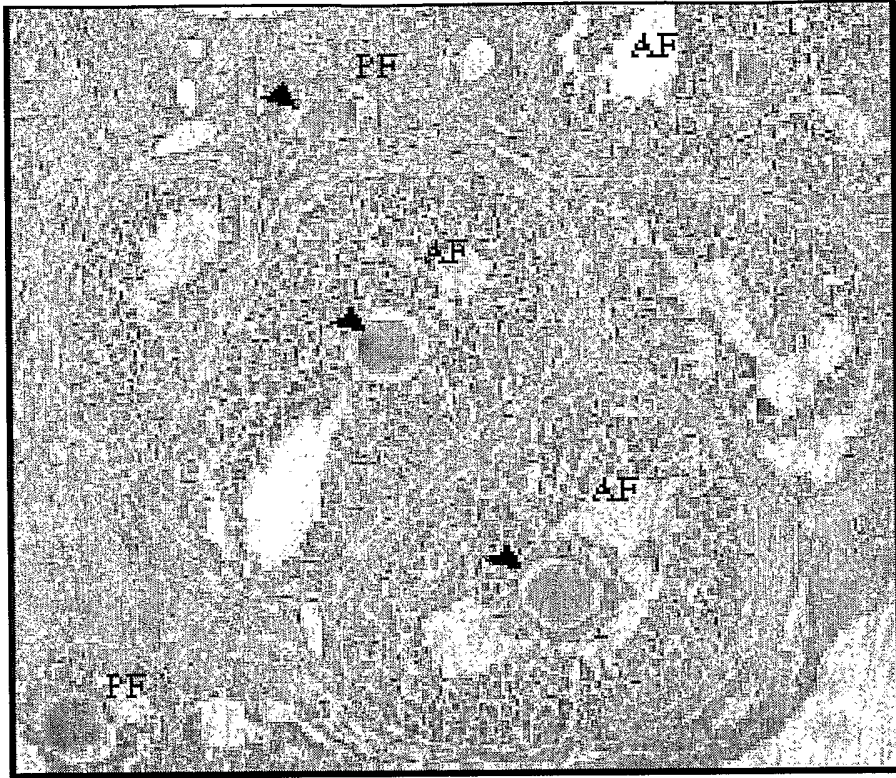


FIG. 7

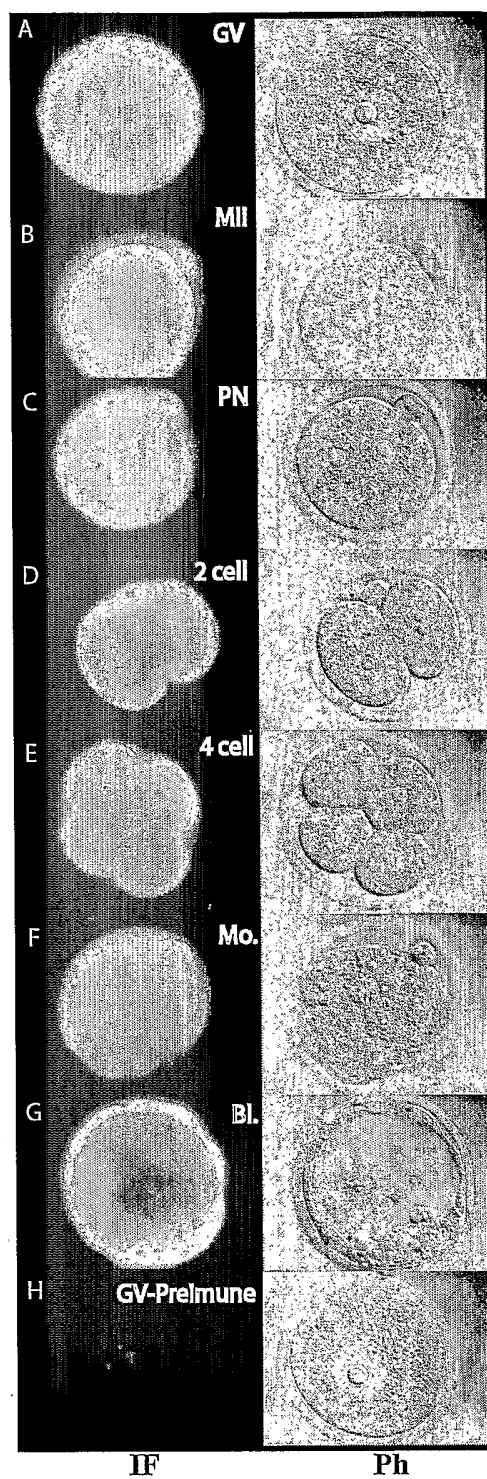


FIG. 8



FIG. 9

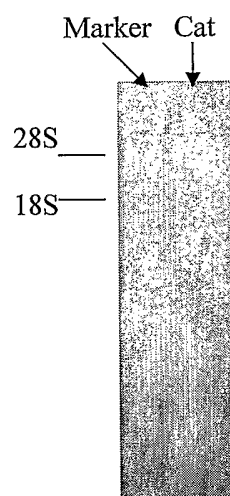


FIG. 10

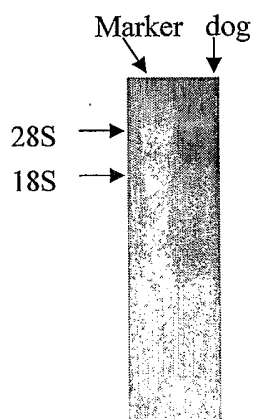


FIG. 11

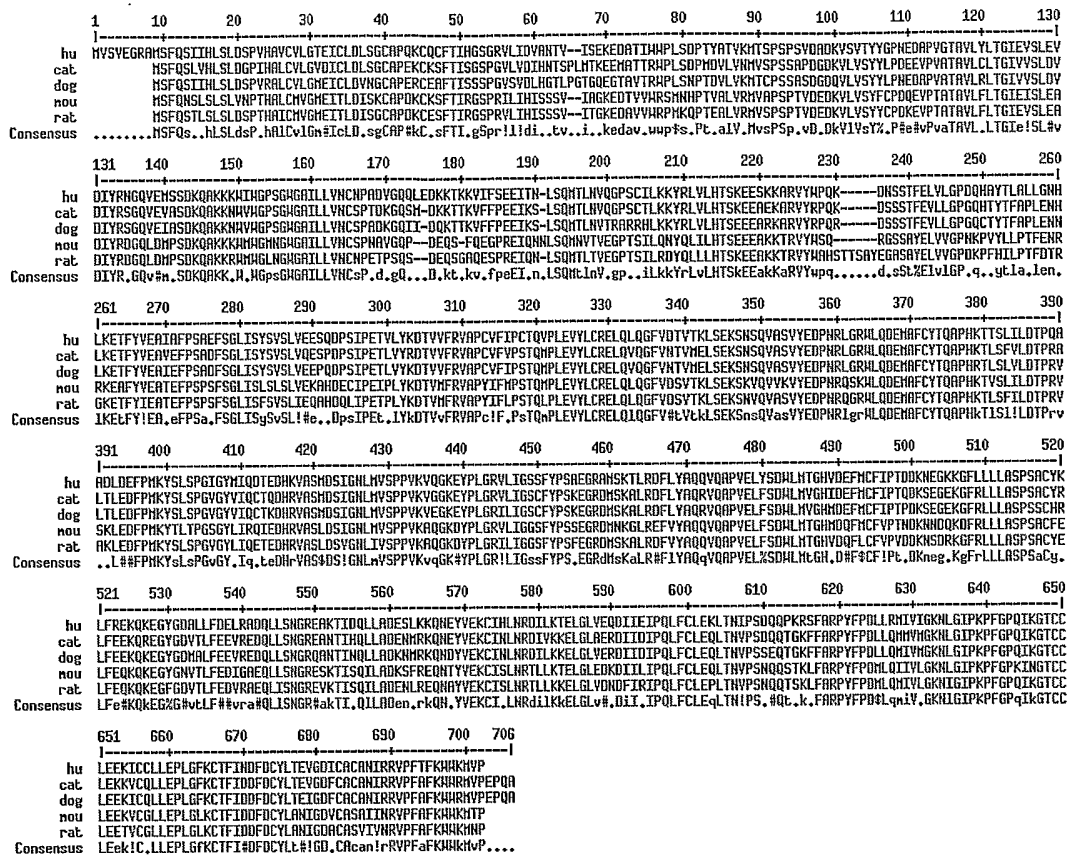
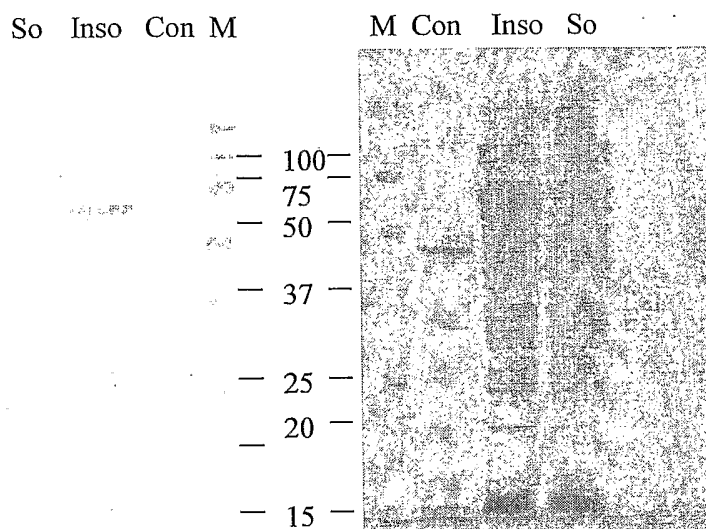
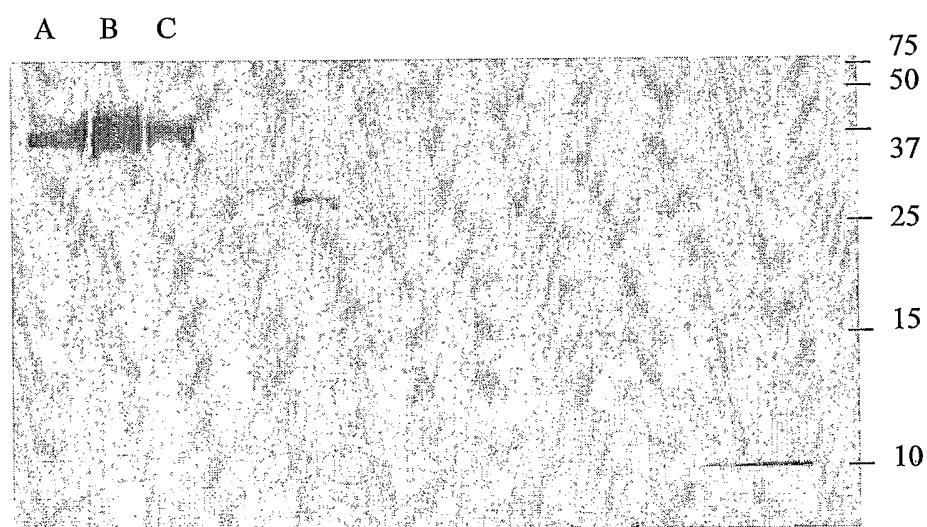
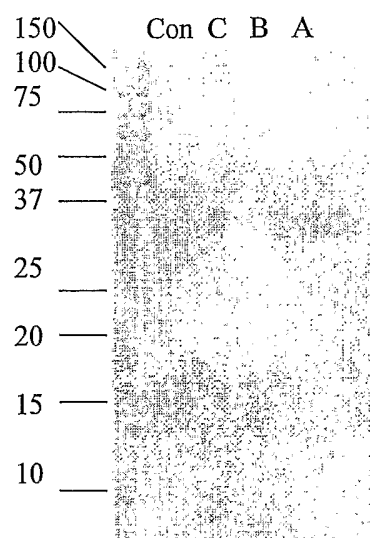
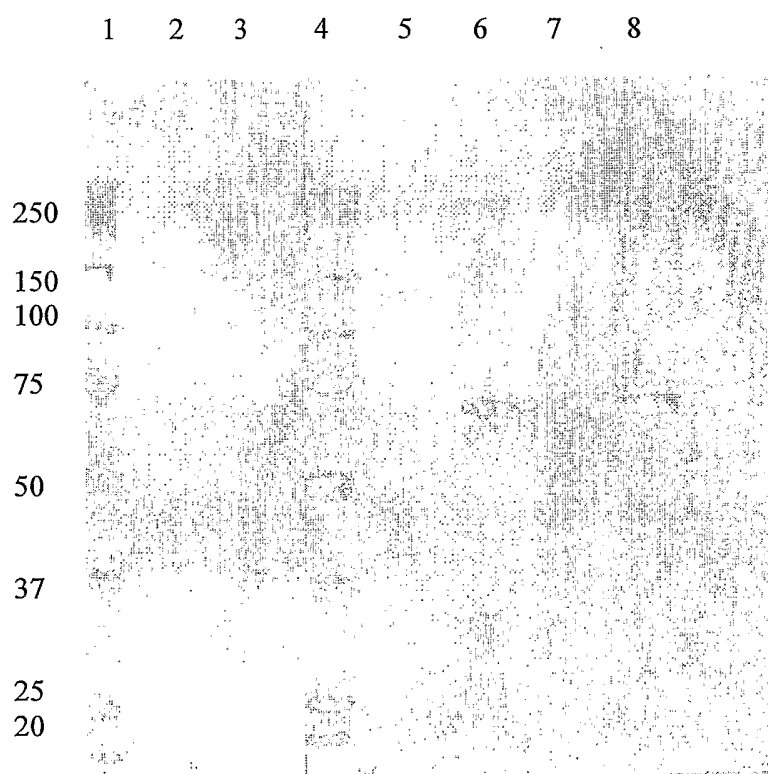


FIG. 12

**FIG. 13****FIG. 14**

**FIG. 15****FIG. 16**

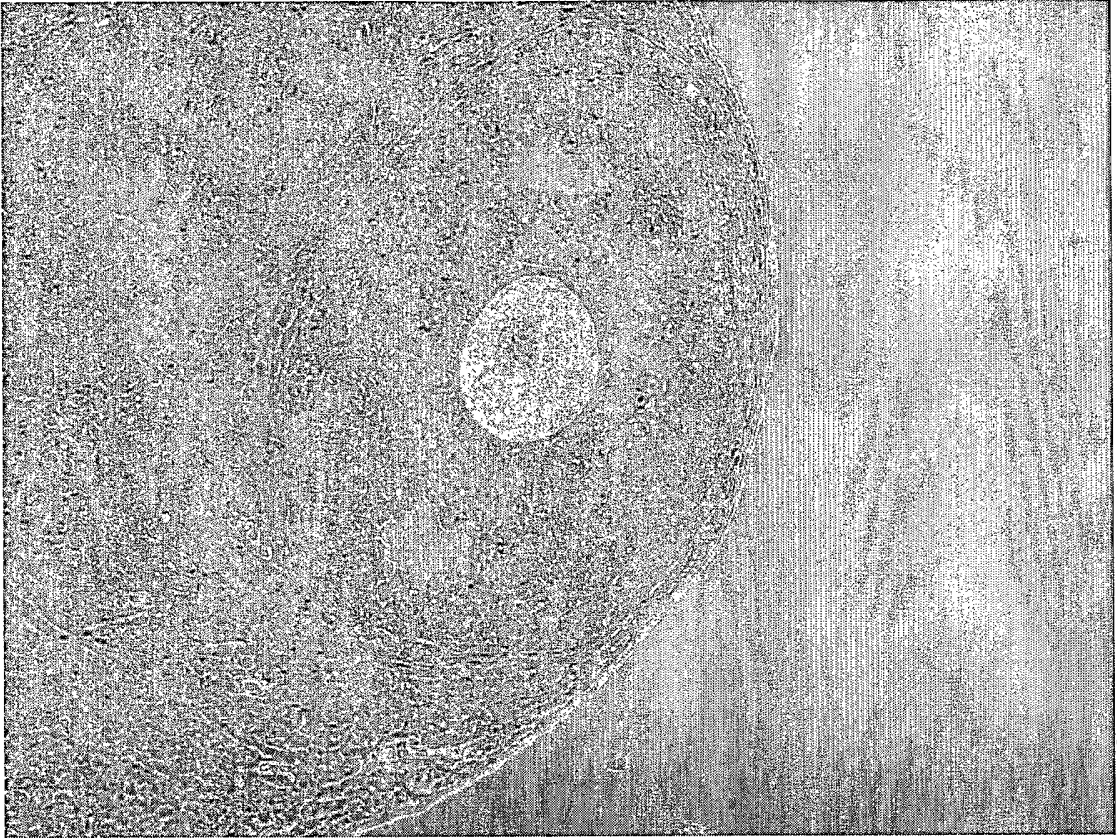


FIG. 17

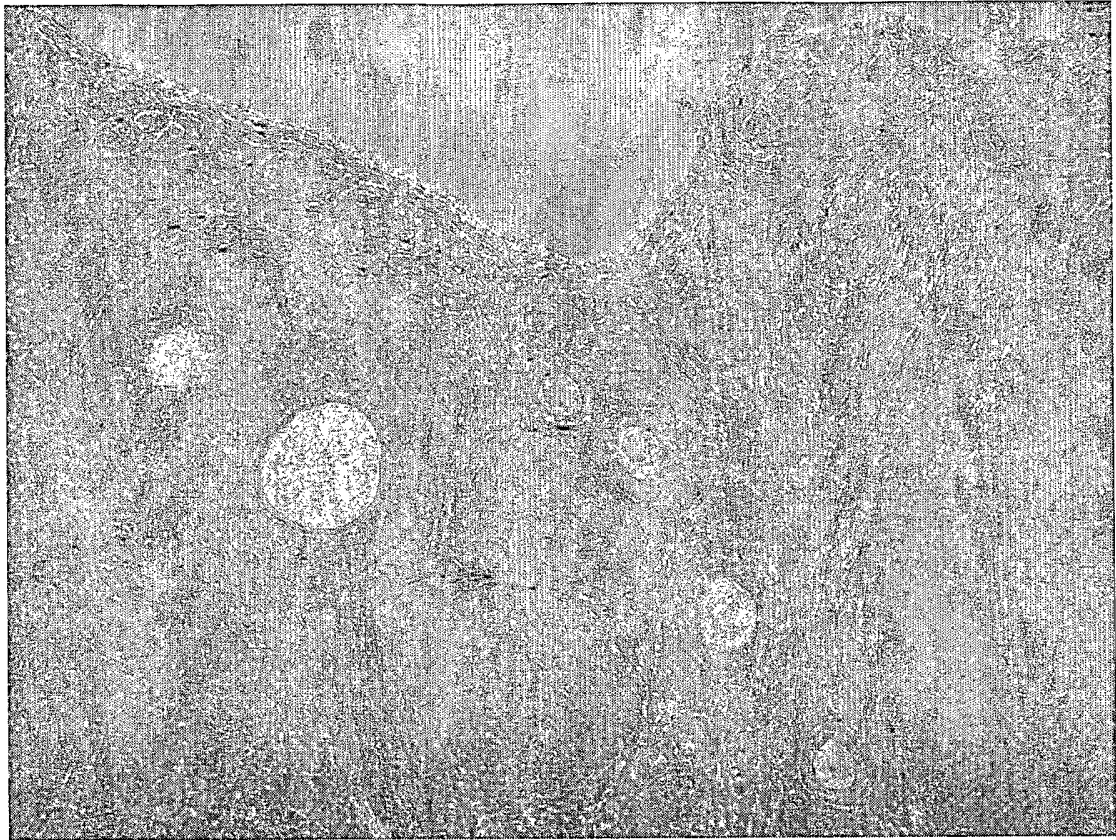


FIG. 18

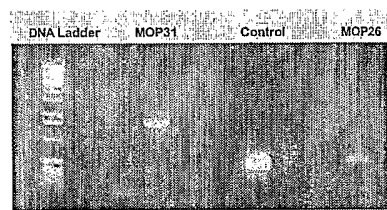


FIG. 19

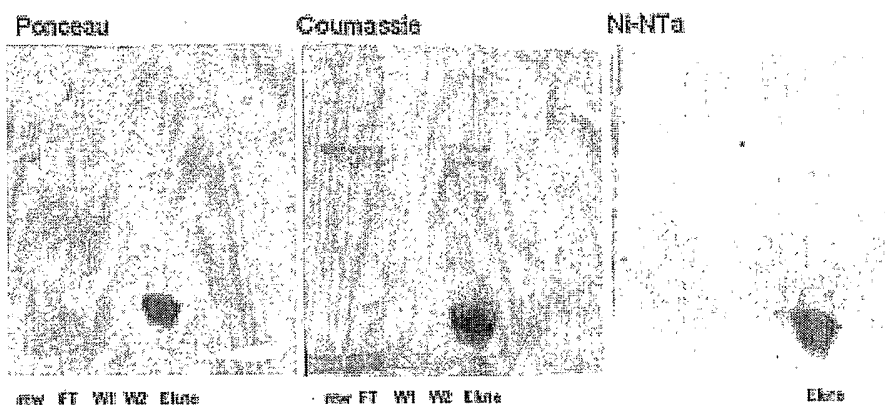


FIG. 20

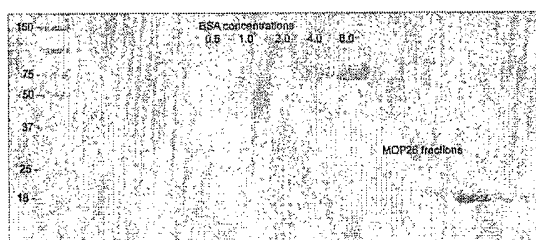
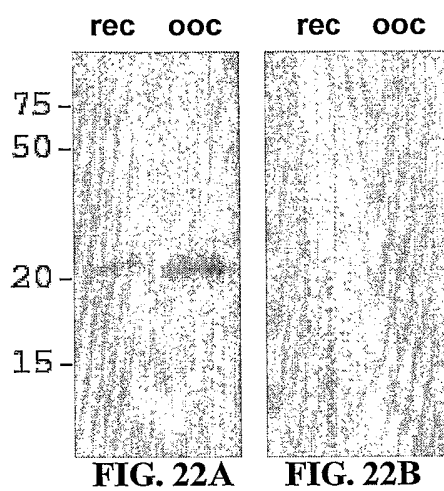
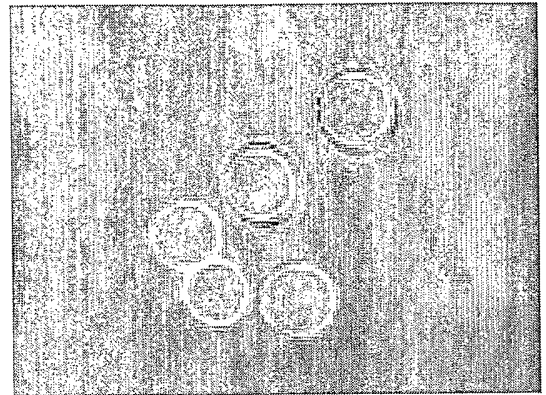
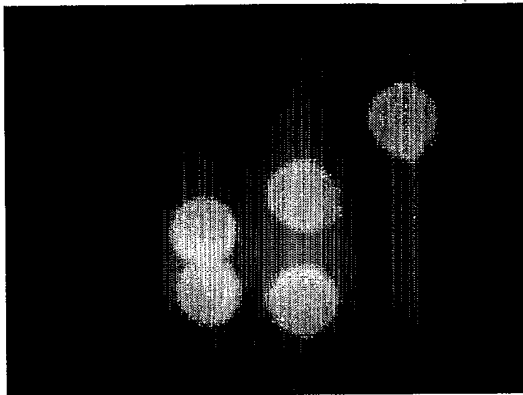


FIG. 21



A



B

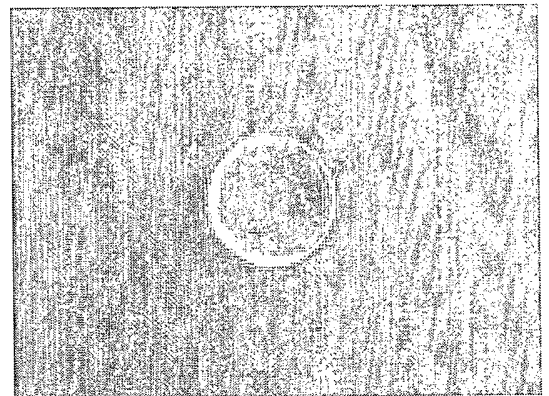
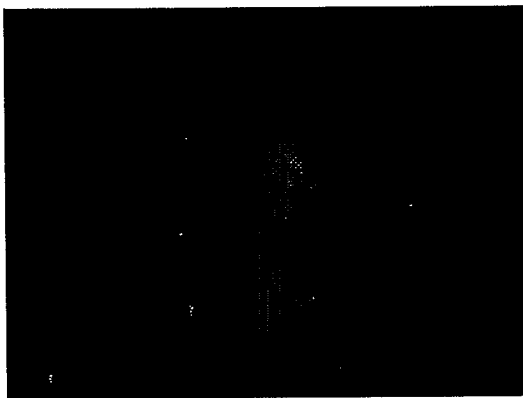


FIG. 23

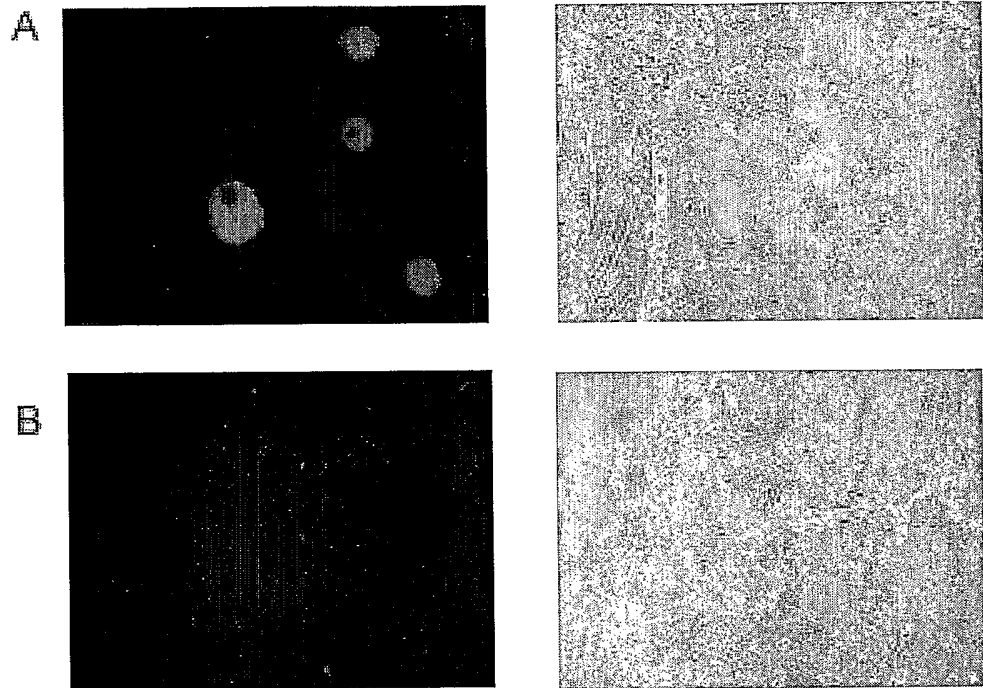


FIG. 24

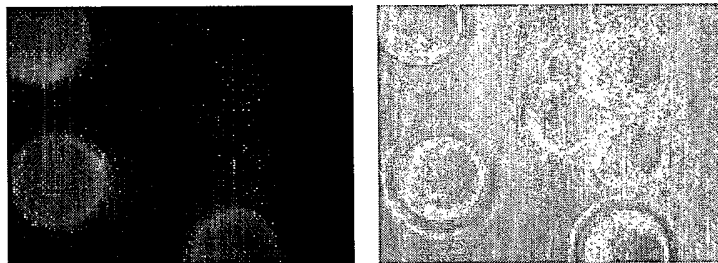


FIG. 25

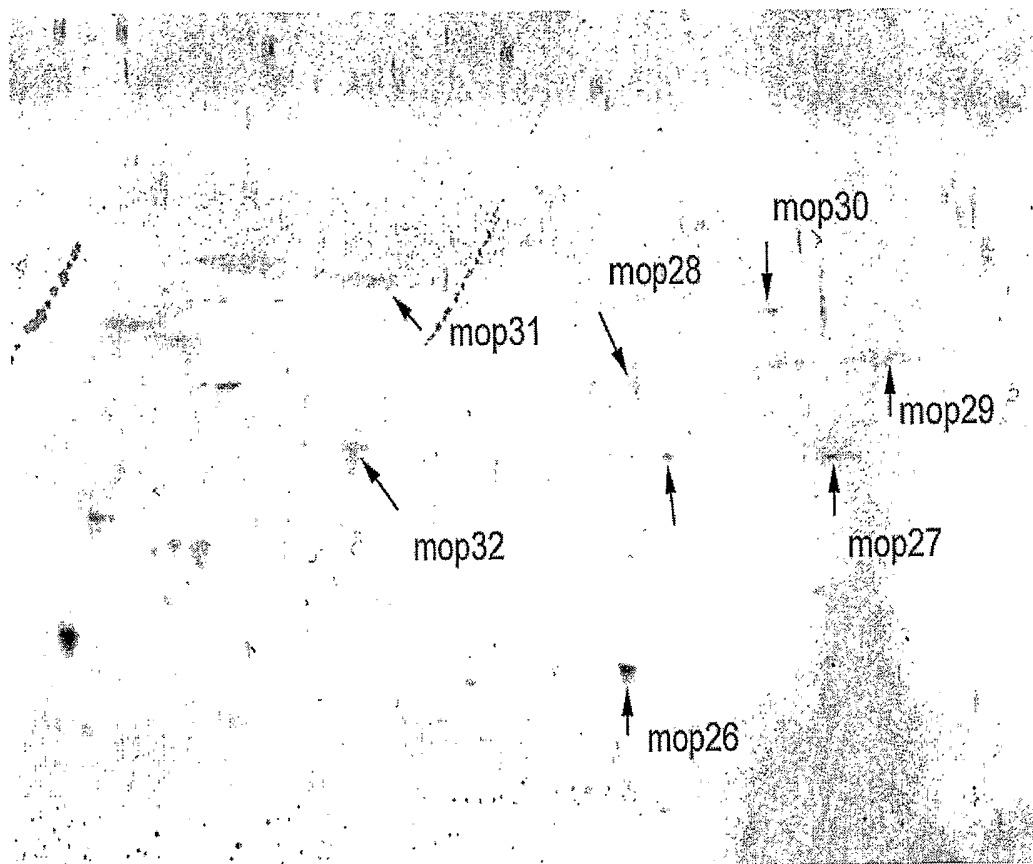


FIG. 26

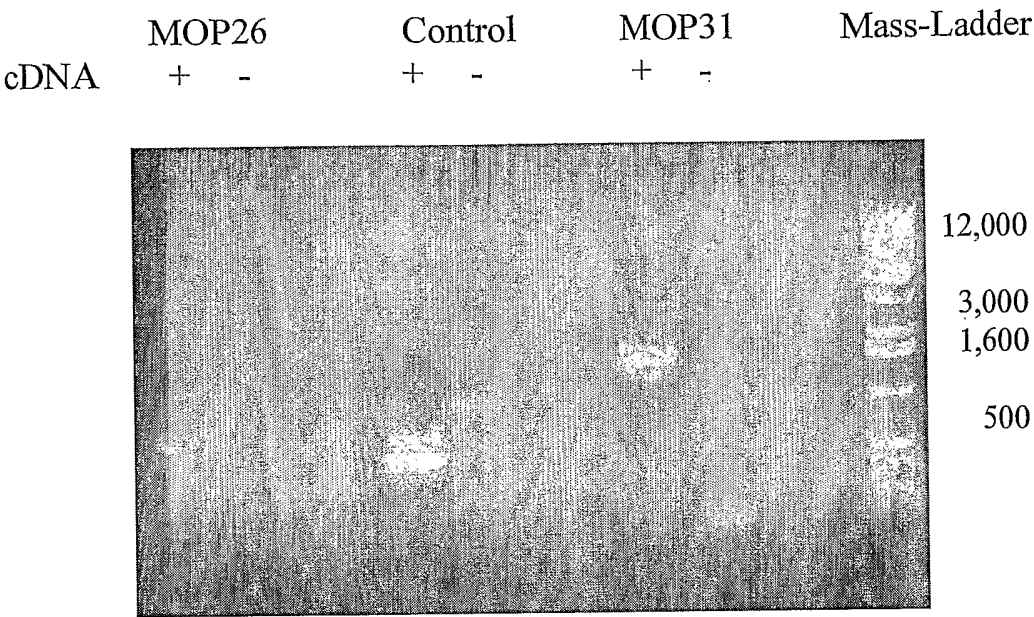


FIG. 27

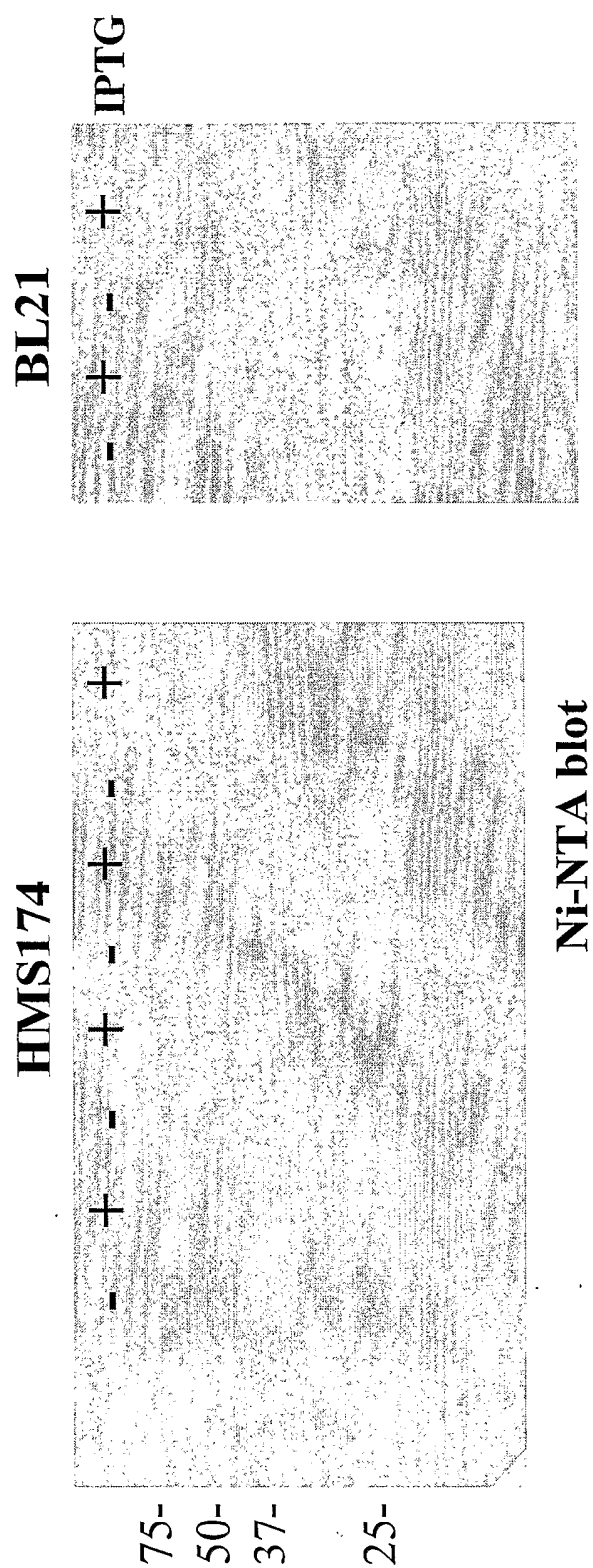


FIG. 28

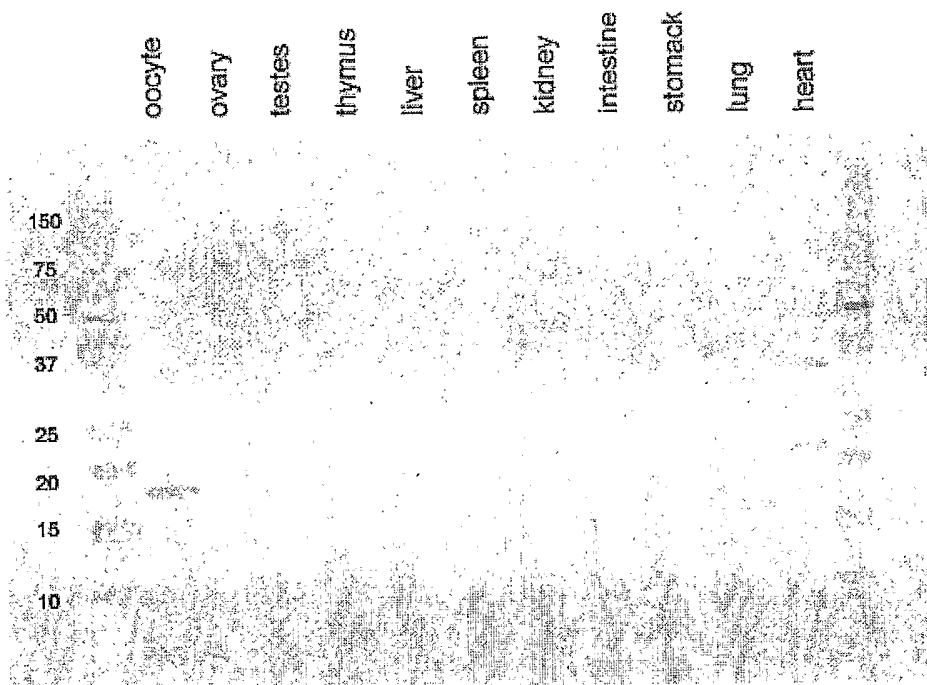


FIG. 29

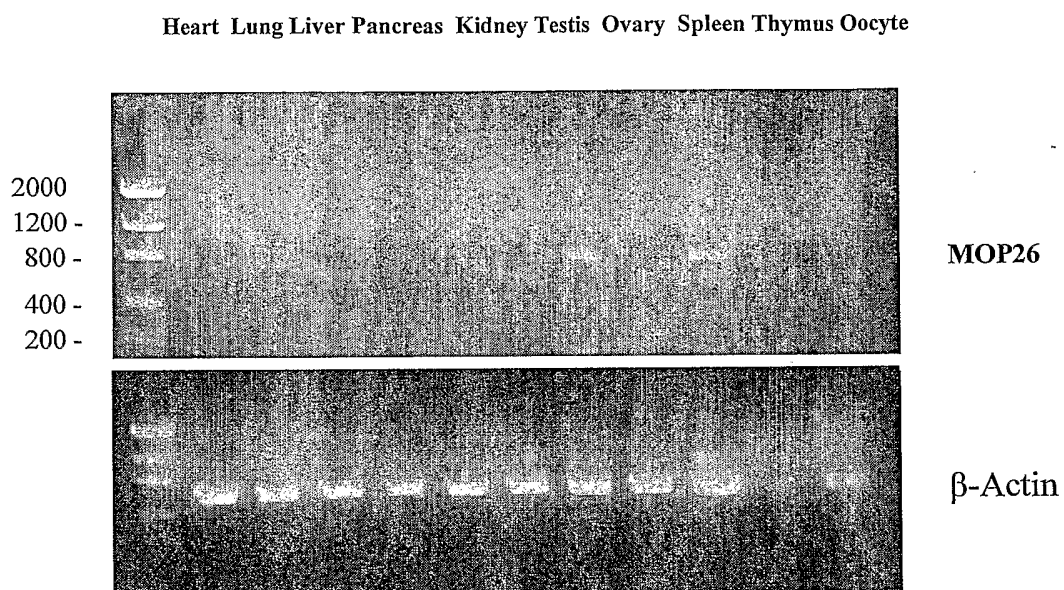


FIG. 30

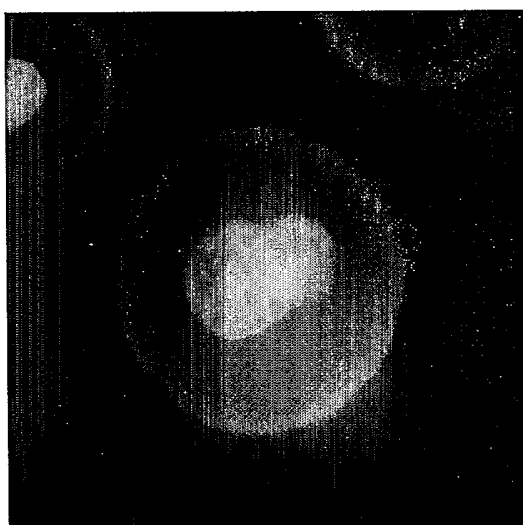


FIG. 31A

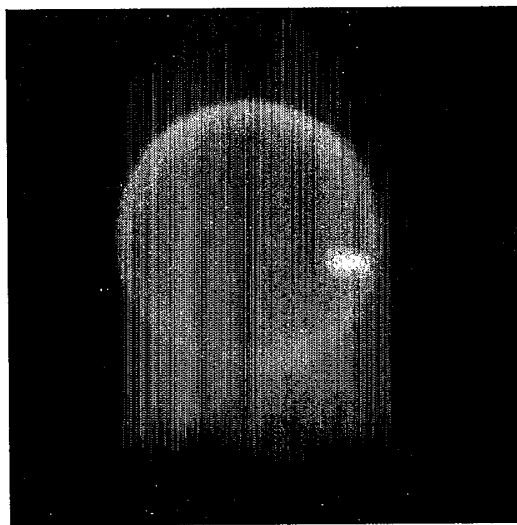


FIG. 31B

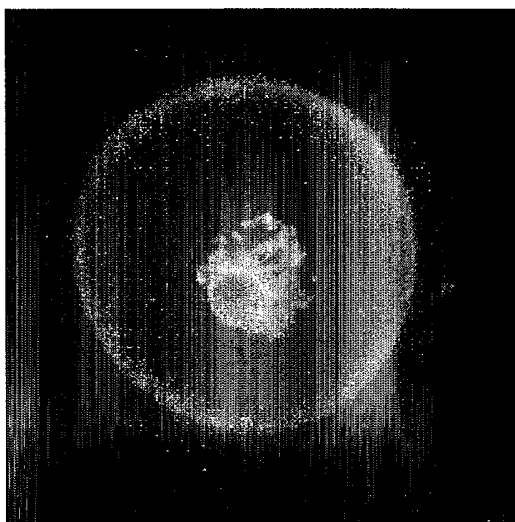


FIG. 31C

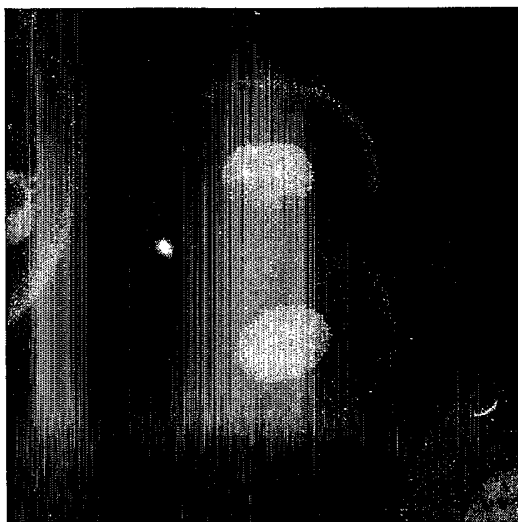


FIG. 31D

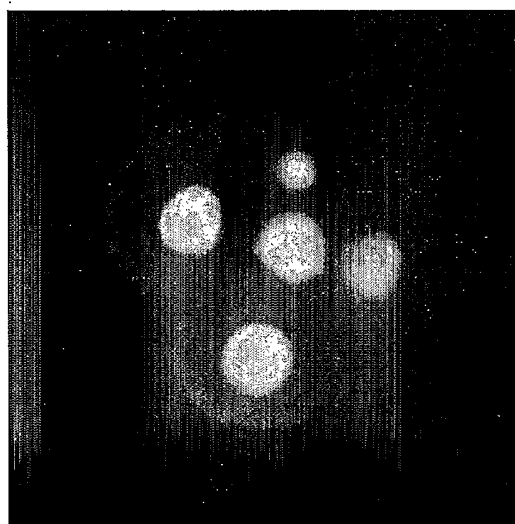


FIG. 31E

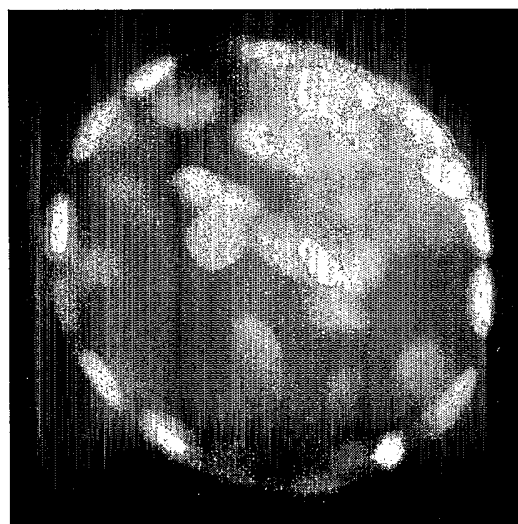


FIG. 31F

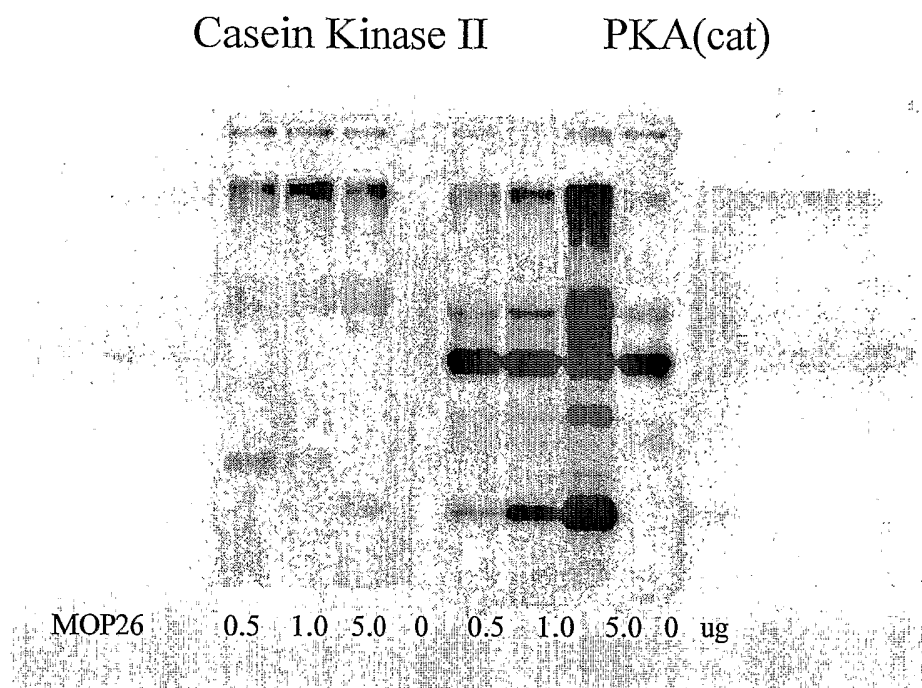


FIG. 32

HerrVETVAC01065-03.txt
SEQUENCE LISTING

<110> University of Virginia Patent Foundation
Chertihin, Olga
He, Wei
Herr, John C

<120> Contraceptive Vaccines for Dogs and Cats Based on Egg Membrane Antigens

<130> 01065-03

<150> 60/630507
<151> 2004-11-23

<150> 60/630310
<151> 2004-11-23

<150> 60/635885
<151> 2004-12-14

<150> 60/725871
<151> 2005-10-11

<160> 17

<170> PatentIn version 3.3

<210> 1
<211> 21
<212> DNA
<213> mus musculus

<400> 1
atggaactaa gctctgggggt c 21

<210> 2
<211> 21
<212> DNA
<213> mus musculus

<400> 2
agggtgttgt atagattcct a 21

<210> 3
<211> 21
<212> DNA
<213> mus musculus

<400> 3
tccagtctaa agaagcaagt g 21

<210> 4
<211> 13
<212> PRT
<213> mus musculus

<400> 4
Phe Ile Glu Gly Pro Val Thr Tyr Ser Glu Ala Pro Arg
1 5 10

<210> 5
<211> 597

HerrVETVAC01065-03.txt

<212> PRT

<213> mus musculus

<400> 5

Met Glu Leu Ser Ser Gly Val Cys Pro Ala Thr Arg Leu Gln Glu Ala
 1 5 10 15

Glu Lys Ala Ala Val His Lys Arg Ser Pro Lys Val Leu Glu Ala Leu
 20 25 30

Arg Lys Leu Asn Ile Gln Ala Asp Gln Ala Pro Val Ile Ala Val Leu
 35 40 45

Gly Ser Gly Gly Gly Leu Arg Ala His Ile Ala Cys Leu Gly Val Leu
 50 55 60

Ser Glu Leu Lys Glu Leu Gly Leu Leu Asp Ala Val Thr Tyr Leu Ala
 65 70 75 80

Gly Val Ser Gly Ser Thr Trp Ala Leu Ser Ser Leu Tyr Thr Lys Asn
 85 90 95

Gly Asn Met Glu Gly Ile Glu Glu Glu Leu Lys His Arg Tyr Glu Lys
 100 105 110

Asn Glu Trp Asp Phe His Glu Ser Leu Glu Lys Ala Ile Gln Ala Ser
 115 120 125

Lys Arg Glu Asn Tyr Ser Leu Thr Asp Phe Trp Ala Tyr Leu Ile Val
 130 135 140

Ser Arg Gln Ile Arg Glu Leu Gln Asp Ser Asn Leu Ser Ser Leu Lys
 145 150 155 160

Lys Gln Val Glu Glu Gly Val Leu Pro Tyr Pro Ile Phe Ala Ala Ile
 165 170 175

Asp Glu Asp Leu Leu Ala Asp Trp Arg Glu Arg Lys Thr Gln Asn Ser
 180 185 190

Trp Phe Glu Phe Thr Pro His His Ala Gly Tyr Pro Ala Leu Gly Ala
 195 200 205

Tyr Val Pro Ile Thr Glu Phe Gly Ser Arg Phe Glu Asn Gly Lys Leu
 210 215 220

Val Lys Ser Glu Pro Glu Arg Asp Leu Thr Phe Leu Arg Gly Leu Trp
 225 230 235 240

Gly Ser Ala Phe Ala Asp Ile Lys Glu Ile Lys Asn Tyr Ile Leu Asn
 245 250 255

HerrVETVAC01065-03.txt

Tyr Phe Arg Asn Pro Phe Gly Lys Leu Lys Phe Ile Glu Gly Pro Val
 260 265 270
 Thr Tyr Ser Glu Ala Pro Arg Met Asn Val Asp Ala Met Leu Leu Asp
 275 280 285
 Leu Val Met Ala Tyr Phe Thr Asp Met Asn Asp Pro Ser Ile Lys Asp
 290 300
 Lys Leu Cys Ala Leu Gln Gln Ala Leu Gly Thr Glu Thr Asp Glu Phe
 305 310 315 320
 Gly Ile Glu Met Ala Glu Ile Ile Gln Asn Trp Asn Glu Thr Ser Ala
 325 330 335
 Glu Lys Lys Glu Gln Phe Leu Asp His Leu Leu Asp Arg Phe Lys Lys
 340 345 350
 Thr Gln Glu Asp Thr Thr Thr Tyr Ser Leu Met Asn Trp Asn Thr Gly
 355 360 365
 Leu Val Trp Asp Arg Cys Val Phe Val Asn Glu Thr Arg Lys Cys Val
 370 375 380
 Ser Lys Trp Gln Trp Gly Thr Val Tyr Asn Phe Leu Tyr Lys His Gly
 385 390 395 400
 Lys Ile Ala Asp Glu Thr Met Cys Ser Arg Glu Leu Leu His Leu Val
 405 410 415
 Asp Ala Gly Phe Ala Ile Asn Thr Pro Tyr Pro Leu Val Leu Pro Pro
 420 425 430
 Val Arg Glu Thr His Leu Ile Leu Ser Phe Asp Phe Ser Ala Gly Asp
 435 440 445
 Pro Leu Glu Thr Ile Arg Ala Thr Ala Asp Tyr Cys Gln Arg His Glu
 450 455 460
 Ile Pro Phe Pro Glu Val Ser Glu Asp Gln Leu Lys Glu Trp Ala Lys
 465 470 475 480
 Ala Pro Ala Ser Cys Tyr Val Leu Arg Gly Glu Thr Gly Pro Val Val
 485 490 495
 Met His Phe Thr Leu Phe Asn Lys Asp Asn Cys Gly Asp Asp Ile Glu
 500 505 510
 Thr Trp Arg Lys Lys Tyr Gly Thr Val Lys Leu Ser Asp Ser Tyr Thr
 515 520 525

HerrVETVAC01065-03.txt

Pro Asp Leu Val Arg Asp Leu Leu Arg Val Ser Lys Glu Asn Val Lys
 530 535 540

Lys Asn Lys Ile Asn Ile Leu Ser Glu Met Arg Lys Val Ala Gly Asn
 545 550 555 560

Pro Gly Asn Ile Pro Arg Val Asn Lys Glu Ala Cys Leu Gly Asp Arg
 565 570 575

Val Lys Asp Pro Gln Gly Ser Gln Thr Val Glu Phe Lys Lys Ser His
 580 585 590

Asn Ile Ser Lys Asp
 595

<210> 6
 <211> 2663
 <212> DNA
 <213> mus musculus

<220>
 <221> misc_feature
 <222> (2300)..(2300)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (2303)..(2303)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (2329)..(2329)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (2345)..(2345)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (2410)..(2410)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (2416)..(2416)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (2444)..(2444)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (2472)..(2472)
 <223> n is a, c, g, or t

<220>

HerrVETVAC01065-03.txt

<221> misc_feature
<222> (2488)..(2488)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2519)..(2519)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2540)..(2540)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2543)..(2543)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2593)..(2593)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (2629)..(2629)
<223> n is a, c, g, or t

<400> 6
atggccttctg cgcttgctcct gctccctgac atgcttgagt tccagtcctg acgtccttgg 60
tgataaacag cagcatggaa ctaagctctg gggctctgcc tgccactaga ctccaggaag 120
cagaaaaggc agctgtgcac aaacgcagtc ccaaggtttt ggaggctcta cgaaagctca 180
acatccaagc tgaccaggct ccagtcattg ctgtcttggg ctctggcggg gggctgcggg 240
cccacatcgc ttgtcttggg gtgctgagtg agctgaaaga acttggcctg ttggatgctg 300
tcacatacct cgcagggggtc tctgggtcca cttgggcact gtcttcactc tacaccaaga 360
atggaaatat ggaagggata gaagaggagc tgaaacatcg gtatgagaag aatgagtggg 420
actttcatga gagcctggag aaagccatcc aggcatacaa gagggagaat tactccctga 480
ctgacttttg ggcctattta attgtttcca ggcaaatcag agaacttcag gattcgaatt 540
tgtccagtct aaagaagcaa gtggaagaag gagtgtgtcc ctatcccatc ttgagagcca 600
ttgatgagga ccttctggct gattggaggg agaggaaaac tcagaattcc tgggttgaat 660
tcactcctca tcatgtctggc taccctgcac ttggggctta tgtcccatc acagagtttg 720
gaagcagatt tgagaatggg aaactgggta aatctgagcc tgagagagat ttgactttcc 780
tgagagggtt atggggaagt gcttttgctg atattaaaga aattaagaat tatattttga 840
actacttcag gaaccctttt ggaaaattga agtttataga aggaccagtg acatactcag 900
aagcaccaag gatgaatgtg gatgcaatgc tcttggattt agtgatggct tatttcacag 960
atatgaatga cccagcatc aaggataagc tctgtgccct tcagcaggct ctgggtactg 1020
agacagatga atttggcata gagatggctg agatcatcca gaattggaat gagacctccg 1080
cagagaagaa ggagcagttt ctcgaccatc tggttgatcg cttcaagaag acacaagaag 1140

HerrVETVAC01065-03.txt

```

acaccaccac atacagtttg atgaactgga acacgggtct agtttgggac cgttgcgttt 1200
ttgtgaatga aactcgcaag tgtgtctcca aatggcagtg gggaactgtt tacaacttcc 1260
tctataaaca cggtaaaatt gcagatgaga ccatgtgcag ccgagagctt ctccatctgg 1320
tggatgctgg ttttgccatc aacactccct atccacttgt tctgcctcct gtgcbtgaaa 1380
ctcacctcat cctctcatth gacttcagtg ctggggaccc actagagacc atcagggcca 1440
cagcagacta ctgccaacgc catgaaatcc cctttcctga ggtgagcgag gatcagttga 1500
aggaatgggc caaagcccca gcaagctgct atgtcctcag aggggaaaca ggacctgttg 1560
tcatgcactt tactctgttc aacaaagaca actgtggaga tgatattgag acatggagaa 1620
aaaaatatgg gacagtaaaa ctatctgact catacacacc agacctgggtg agagatttgc 1680
tgagggtatc caaggagaat gtgaagaaaa acaaaattaa tatcctcagt gagatgagga 1740
aagtggctgg gaatcctggg aacatcccaa gagtgaacaa ggaggcctgc ttgggagaca 1800
gagtaaagga tccccaaggc tctcagactg tggagtttaa gaaatccac aacatatcta 1860
aggattaagg atcctttgca ctacctggag agttggcatc tattagtgga ctcaccacct 1920
gtgctctctc caaagtccac cgtaagctgt ctctttgtga agaagaactg cagttttcag 1980
gggtactgtg gcctactgct ttaccagggc ccaagtgtca acctggcctg atgttctctg 2040
ccactcataa aattgcattt gccttagcta ttaccattga gcaagtctcc agagacctag 2100
atggtgtctt tgatagcatt atttagggca ctgggtcacc agagctaaaa gagggcaaagt 2160
atgtttttca gatactcag gggaccctaa agctagagca caactccttg atcatcaaag 2220
ccaccctgtc caacatccat gtgggttaaaa tgaatgtagt gattgagacc aaagcctcat 2280
aaatttttct ctattcttn atntgctcat tttatacacc cagtcttant aatttttttt 2340
tttntataa aaggatagac cttagctcaa atcatgaccc atcacttcag agcaattaaa 2400
aaaaccagn caactntgag aagacatgtt ctgtatatgc ttgngtatct gctgccttag 2460
tgcctccttc cngcttgctt acttcctntc cttccatcc cttccacctg taactttnt 2520
ttgatttga ttccacaatn ttntcattgt aatatcacag gcaagttcct tcacatgaat 2580
tcctttaatc atntatggaa taaatgtgca ttgtggttgt gccctggtnt ttccaaaaaa 2640
aaaaaaaaa aaaaaaaaaa aaa 2663

```

<210> 7
 <211> 530
 <212> PRT
 <213> mus musculus

<400> 7

Lys Glu Leu Gly Leu Leu Asp Ala Val Thr Tyr Leu Ala Gly Val Ser
 1 5 10 15

Gly Ser Thr Trp Ala Leu Ser Ser Leu Tyr Thr Lys Asn Gly Asn Met
 20 25 30

HerrVETVAC01065-03.txt

Glu Gly Ile Glu Glu Glu Leu Lys His Arg Tyr Glu Lys Asn Glu Trp
 35 40 45

Asp Phe His Glu Ser Leu Glu Lys Ala Ile Gln Ala Ser Lys Arg Glu
 50 55 60

Asn Tyr Ser Leu Thr Asp Phe Trp Ala Tyr Leu Ile Val Ser Arg Gln
 65 70 75 80

Ile Arg Glu Leu Gln Asp Ser Asn Leu Ser Ser Leu Lys Lys Gln Val
 85 90 95

Glu Glu Gly Val Leu Pro Tyr Pro Ile Phe Ala Ala Ile Asp Glu Asp
 100 105 110

Leu Leu Ala Asp Trp Arg Glu Arg Lys Thr Gln Asn Ser Trp Phe Glu
 115 120 125

Phe Thr Pro His His Ala Gly Tyr Pro Ala Leu Gly Ala Tyr Val Pro
 130 135 140

Ile Thr Glu Phe Gly Ser Arg Phe Glu Asn Gly Lys Leu Val Lys Ser
 145 150 155 160

Glu Pro Glu Arg Asp Leu Thr Phe Leu Arg Gly Leu Trp Gly Ser Ala
 165 170 175

Phe Ala Asp Ile Lys Glu Ile Lys Asn Tyr Ile Leu Asn Tyr Phe Arg
 180 185 190

Asn Pro Phe Gly Lys Leu Lys Phe Ile Glu Gly Pro Val Thr Tyr Ser
 195 200 205

Glu Ala Pro Arg Met Asn Val Asp Ala Met Leu Leu Asp Leu Val Met
 210 215 220

Ala Tyr Phe Thr Asp Met Asn Asp Pro Ser Ile Lys Asp Lys Leu Cys
 225 230 235 240

Ala Leu Gln Gln Ala Leu Gly Thr Glu Thr Asp Glu Phe Gly Ile Glu
 245 250 255

Met Ala Glu Ile Ile Gln Asn Trp Asn Glu Thr Ser Ala Glu Lys Lys
 260 265 270

Glu Gln Phe Leu Asp His Leu Leu Asp Arg Phe Lys Lys Thr Gln Glu
 275 280 285

Asp Thr Thr Thr Tyr Ser Leu Met Asn Trp Asn Thr Gly Leu Val Trp
 290 295 300

HerrVETVAC01065-03.txt

Asp Arg Cys Val Phe Val Asn Glu Thr Arg Lys Cys Val Ser Lys Trp
 305 310 315 320

Gln Trp Gly Thr Val Tyr Asn Phe Leu Tyr Lys His Gly Lys Ile Ala
 325 330 335

Asp Glu Thr Met Cys Ser Arg Glu Leu Leu His Leu Val Asp Ala Gly
 340 345 350

Phe Ala Ile Asn Thr Pro Tyr Pro Leu Val Leu Pro Pro Val Arg Glu
 355 360 365

Thr His Leu Ile Leu Ser Phe Asp Phe Ser Ala Gly Asp Pro Leu Glu
 370 375 380

Thr Ile Arg Ala Thr Ala Asp Tyr Cys Gln Arg His Glu Ile Pro Phe
 385 390 395 400

Pro Glu Val Ser Glu Asp Gln Leu Lys Glu Trp Ala Lys Ala Pro Ala
 405 410 415

Ser Cys Tyr Val Leu Arg Gly Glu Thr Gly Pro Val Val Met His Phe
 420 425 430

Thr Leu Phe Asn Lys Asp Asn Cys Gly Asp Asp Ile Glu Thr Trp Arg
 435 440 445

Lys Lys Tyr Gly Thr Val Lys Leu Ser Asp Ser Tyr Thr Pro Asp Leu
 450 455 460

Val Arg Asp Leu Leu Arg Val Ser Lys Glu Asn Val Lys Lys Asn Lys
 465 470 475 480

Ile Asn Ile Leu Ser Glu Met Arg Lys Val Ala Gly Asn Pro Gly Asn
 485 490 495

Ile Pro Arg Val Asn Lys Glu Ala Cys Leu Gly Asp Arg Val Lys Asp
 500 505 510

Pro Gln Gly Ser Gln Thr Val Glu Phe Lys Lys Ser His Asn Ile Ser
 515 520 525

Lys Asp
 530

<210> 8
 <211> 1590
 <212> DNA
 <213> mus musculus

<400> 8
 gaaagaactt ggcctgttgg atgctgtcac atacctcgca ggggtctctg ggtccacttg

60

HerrVETVAC01065-03.txt

```

ggcactgtct tcactctaca ccaagaatgg aaatatggaa gggatagaag aggagctgaa 120
acatcggtat gagaagaatg agtgggactt tcatgagagc ctggagaaaag ccatccaggc 180
atcaaagagg gagaattact ccctgactga cttttgggcc tatttaattg tttccaggca 240
aatcagagaa cttcaggatt cgaatttgtc cagtctaaag aagcaagtgg aagaaggagt 300
gctgccctat cccatctttg cagccattga tgaggacctt ctggctgatt ggagggagag 360
gaaaactcag aattcctggt ttgaattcac tcctcatcat gctggctacc ctgcacttgg 420
ggcttatgtc cccatcacag agtttggag cagatttgag aatgggaaac tggttaaatc 480
tgagcctgag agagatttga ctttcctgag aggtttatgg ggaagtgctt ttgctgatat 540
taaagaaatt aagaattata ttttgaacta cttcaggaac ccctttggaa aattgaagtt 600
tatagaagga ccagtgcacat actcagaagc accaaggatg aatgtggatg caatgctctt 660
ggatttagtg atggcttatt tcacagatat gaatgacccc agcatcaagg ataagctctg 720
tgcccttcag caggctctgg gtactgagac agatgaattt ggcatagaga tggctgagat 780
catccagaat tggaatgaga cctccgcaga gaagaaggag cagtttctcg accatctggt 840
ggatcgcttc aagaagacac aagaagacac caccacatac agtttgatga actggaacac 900
gggtctagtt tgggaccgtt gcgtttttgt gaatgaaact cgcaagtgtg tctccaaatg 960
gcagtgggga actgtttaca acttcctcta taaacacggt aaaattgcag atgagaccat 1020
gtgcagccga gagcttctcc atctggtgga tgctggtttt gccatcaaca ctccctatcc 1080
acttgttctg cctcctgtgc gtgaaactca cctcatcctc tcatttgact tcagtgtctg 1140
ggaccacta gagaccatca gggccacagc agactactgc caacgccatg aaatcccctt 1200
tcctgaggtg agcgaggatc agctgaagga atgggcaaaa gcccagcaa gctgctatgt 1260
cctcagaggg gaaacaggac ctgttgtcat gcactttact ctgttcaaca aagacaactg 1320
tgagatgat attgagacat ggagaaaaaa atatgggaca gtaaaactat ctgactcata 1380
cacaccagac ctggtgagag atttgctgag ggtatccaag gagaatgtga agaaaaacaa 1440
aattaatatc ctcagtgaga tgaggaaagt ggctgggaat cctgggaaca tcccaagagt 1500
gaacaaggag gcctgcttgg gagacagagt gaaggatccc caaggctctc agactgtgga 1560
gtttaagaaa tcccacaaca tatctaagga 1590

```

<210> 9
 <211> 691
 <212> PRT
 <213> feline

<400> 9

Met Ser Phe Gln Ser Leu Val His Leu Ser Leu Asp Gly Pro Ile His
 1 5 10 15

Ala Leu Cys Val Leu Gly Val Asp Ile Cys Leu Asp Leu Ser Gly Cys
 20 25 30

HerrVETVAC01065-03.txt

Ala Pro Glu Lys Cys Lys Ser Phe Thr Ile Ser Gly Ser Pro Gly Val
 35 40 45
 Leu Val Asp Ile His Asn Thr Ser Pro Leu Met Thr Lys Glu Glu Met
 50 55 60
 Ala Thr Thr Arg Trp Pro Leu Ser Asp Pro Met Asp Val Leu Val Asn
 65 70 75 80
 Met Val Ser Pro Ser Ser Ala Pro Asp Gly Asp Lys Val Leu Val Ser
 85 90 95
 Tyr Tyr Leu Pro Asp Glu Glu Val Pro Val Ala Thr Ala Val Leu Cys
 100 105 110
 Leu Thr Gly Ile Val Val Ser Leu Asp Val Asp Ile Tyr Arg Ser Gly
 115 120 125
 Gln Val Glu Val Ala Ser Asp Lys Gln Ala Lys Lys Asn Trp Val Trp
 130 135 140
 Gly Pro Ser Gly Trp Gly Ala Ile Leu Leu Val Asn Cys Ser Pro Thr
 145 150 155 160
 Asp Lys Gly Gln Ser Met Asp Lys Lys Thr Thr Lys Val Phe Phe Pro
 165 170 175
 Glu Glu Ile Lys Ser Leu Ser Gln Met Thr Leu Asn Val Gln Gly Pro
 180 185 190
 Ser Cys Thr Leu Lys Lys Tyr Arg Leu Val Leu His Thr Ser Lys Glu
 195 200 205
 Glu Ala Glu Lys Ala Arg Val Tyr Arg Pro Gln Lys Asp Ser Ser Ser
 210 215 220
 Thr Phe Glu Val Leu Leu Gly Pro Gly Gln His Thr Tyr Thr Phe Ala
 225 230 235 240
 Pro Leu Glu Asn His Leu Lys Glu Thr Phe Tyr Val Glu Ala Val Glu
 245 250 255
 Phe Pro Ser Ala Asp Phe Ser Gly Leu Ile Ser Tyr Ser Val Ser Leu
 260 265 270
 Val Gln Glu Ser Pro Asp Pro Ser Ile Pro Glu Thr Leu Val Tyr Arg
 275 280 285
 Asp Thr Val Val Phe Arg Val Ala Pro Cys Val Phe Val Pro Ser Thr
 290 295 300

HerrVETVAC01065-03.txt

Gln Met Pro Leu Glu Val Tyr Leu Cys Arg Glu Leu Gln Val Gln Gly
 305 310 315 320

Phe Val Asn Thr Val Met Glu Leu Ser Glu Lys Ser Asn Ser Gln Val
 325 330 335

Ala Ser Val Tyr Glu Asp Pro Asn Arg Leu Gly Arg Trp Leu Gln Asp
 340 345 350

Glu Met Ala Phe Cys Tyr Thr Gln Ala Pro His Lys Thr Leu Ser Phe
 355 360 365

Val Leu Asp Thr Pro Arg Ala Leu Thr Leu Glu Asp Phe Pro Met Lys
 370 375 380

Tyr Ser Leu Ser Pro Gly Val Gly Tyr Val Ile Gln Cys Thr Gln Asp
 385 390 395 400

His Arg Val Ala Ser Met Asp Ser Ile Gly Asn Leu Met Val Ser Pro
 405 410 415

Pro Val Lys Val Gly Gly Lys Glu Tyr Pro Leu Gly Arg Val Leu Ile
 420 425 430

Gly Ser Cys Phe Tyr Pro Ser Lys Glu Gly Arg Asp Met Ser Lys Ala
 435 440 445

Leu Arg Asp Phe Leu Tyr Ala Gln Arg Val Gln Ala Pro Val Glu Leu
 450 455 460

Phe Ser Asp Trp Leu Met Val Gly His Ile Asp Glu Phe Met Cys Phe
 465 470 475 480

Ile Pro Thr Gln Asp Lys Ser Glu Gly Glu Lys Gly Phe Arg Leu Leu
 485 490 495

Leu Ala Ser Pro Ser Ala Cys Tyr Arg Leu Phe Glu Glu Lys Gln Arg
 500 505 510

Glu Gly Tyr Gly Asp Val Thr Leu Phe Glu Glu Val Arg Glu Asp Gln
 515 520 525

Leu Leu Ser Asn Gly Arg Glu Ala Asn Thr Ile His Gln Leu Leu Ala
 530 535 540

Asp Glu Asn Met Arg Lys Gln Asn Glu Tyr Val Glu Lys Cys Ile Asn
 545 550 555 560

Leu Asn Arg Asp Ile Val Lys Lys Glu Leu Gly Leu Ala Glu Arg Asp
 565 570 575

HerrVETVAC01065-03.txt

Ile Ile Asp Ile Pro Gln Leu Phe Cys Leu Glu Gln Leu Thr Asn Val
 580 585 590

Pro Ser Asp Gln Gln Thr Gly Lys Phe Phe Ala Arg Pro Tyr Phe Pro
 595 600 605

Asp Leu Leu Gln Met Met Val Met Gly Lys Asn Leu Gly Ile Pro Lys
 610 615 620

Pro Phe Gly Pro Gln Ile Lys Gly Thr Cys Cys Leu Glu Lys Lys Val
 625 630 635 640

Cys Gln Leu Leu Glu Pro Leu Gly Phe Lys Cys Thr Phe Ile Asp Asp
 645 650 655

Phe Asp Cys Tyr Leu Thr Glu Val Gly Asp Phe Cys Ala Cys Ala Asn
 660 665 670

Ile Arg Arg Val Pro Phe Ala Phe Lys Trp Trp Arg Met Val Pro Glu
 675 680 685

Pro Gln Ala
 690

<210> 10
 <211> 2374
 <212> DNA
 <213> feline

<400> 10
 ggagatcggg gttatctgag gccgctgtgc tgagggaggg ttgtcgaaca ggcggccacc 60
 ttggaggcct ttggcatgtc cttccagagc ctcgtccacc tgtccctgga cggccccatc 120
 catgccctct gcgtgctggg cgtggacatc tgcttgatc tcagtgggtg tgccccggag 180
 aaatgcaagt catttaccat cagtggctct ccgggggtct tggttgacat ccacaacacg 240
 tccccgttaa tgaccaagga ggagatggcc acgacccggt ggcctctgtc tgatcccatg 300
 gatgtcctgg tgaatatggt ctccccagc tctgcccccg atggtgacaa ggttctggtc 360
 tcctactatc tgcctgatga ggaagtcca gtggccacgg ctgtgctctg cctcactggc 420
 attgtgggtct ccctagacgt ggacatctac cgcagcgggc aagtcgaggt ggccagtgc 480
 aagcaggcta agaaaaactg ggtctgggga cccagcggct ggggcgccat cctgcttggtg 540
 aactgcagcc ctaccgacaa gggccagagc atggacaaga agaccacgaa ggtgttcttt 600
 cccgaggaga tcaagagtct gtcccagatg accctgaatg ttcaagggcc cagctgcact 660
 ttaaagaaat accggctgggt tctccacacc tccaaggaag aggcagagaa ggcaagagtc 720
 taccggcccc aaaaagacag ctcgagcacc tttgaggtac tgctggggcc tggccagcac 780
 acctacacct ttgccccct cgagaaccac ctgaaggaaa ccttctacgt ggaagctgtg 840
 gaattcccgt ctgccgactt ctcgggcctg atttctact ctgtctccct ggtgcaagaa 900

HerrVETVAC01065-03.txt

```

tctccggacc cgtccattcc agagaccctg gtatacagag acaccgtggt attccgggtg 960
gcccttgcg tcttcgttcc cagcacgcag atgcctctag aggtttacct gtgcagagag 1020
ctgcagggtcc aaggttttgt gaacacagtg atggagctga gtgagaagag taacagccaa 1080
gtggcatctg tctacgagga cccaatcgc ctgggcagggt ggctccagga cgagatggcc 1140
ttctgttaca cccagggtcc ccacaagacg ctctcctttg tcctcgacac ccctcggggc 1200
ctcacgtggt aagattttccc catgaaatac tctctgagcc caggggtcgg ttacgtgatc 1260
cagtgcaccc aggaccacag ggtggccagc atggattcca ttgggaacct gatggtgtcc 1320
ccaccggtca aggtggggagg gaaagagtac cccctaggca gggtcctcat tggcagctgc 1380
ttttaccca gcaaggaggg ccgagacatg agtaaggccc tccgggactt cctctacgcg 1440
cagcgagtcc agggccccgt ggagctcttc tcggattggc tgatgggtcgg ccacatagac 1500
gagttcatgt gctttatccc cacacaggac aagagtgagg gcgaaaaggg cttccgggtg 1560
ctcctggcca gccccagcgc ctgctacagg ctgtttgagg agaaacagag agagggctat 1620
ggagacgtga ctctgtttga ggaggtcaga gaggatcagc tcctctccaa cgggagggag 1680
gccaatacca tccatcaact tctagctgat gaaaacatga gaaagcagaa cgaatatgtg 1740
gagaagtgca ttaacctgaa ccgtgacatc gtgaagaagg agctgggcct ggccgagagg 1800
gacatcatcg acatcccaca gctcttctgc ctggagcagc tgacgaacgt cccctccgac 1860
cagcagacgg ggaagttctt cgcgaggccg tacttccctg acctgctgca gatgatggtg 1920
atgggcaaaa atctgggcat cccaagcct tttgggcccc agatcaaggg tacctgctgc 1980
ctggaaaaaa aggtctgcc a gttgctagag cccctgggct tcaagtgcac cttcattgat 2040
gactttgact gctacctgac tgaagtcggg gacttctgtg cctgtgcca catccgccg 2100
gtgccctttg ctttcaaag gtggaggatg gtaccggagc cccaggccta gcgtgccagc 2160
cctgccccag caggaacggc ccattaccac tcacccctg cctctttggg aggtactgca 2220
ccccttcctt acctgttcac cctgtccctc agtatccaga gcgatggtca aactgccag 2280
cctgaacccc tcagaaaacg gtcttgagct gagaacccat taaagtacga gctgttctga 2340
atgcaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 2374

```

<210> 11
 <211> 691
 <212> PRT
 <213> canine

<400> 11

Met Ser Phe Gln Ser Ile Ile His Leu Ser Leu Asp Ser Pro Val Arg
 1 5 10 15

Ala Leu Cys Val Leu Gly Met Glu Ile Cys Leu Asp Val Asn Gly Cys
 20 25 30

Ala Pro Glu Arg Cys Glu Ala Phe Thr Ile Ser Ser Ser Pro Gly Val
 35 40 45

HerrVETVAC01065-03.txt

Ser Val Asp Leu His Gly Thr Leu Pro Gly Thr Gly Gln Glu Gly Thr
 50 55 60
 Ala Val Thr Arg Trp Pro Leu Ser Asn Pro Thr Asp Val Leu Val Lys
 65 70 75 80
 Met Thr Cys Pro Ser Ser Ala Ser Asp Gly Asp Gln Val Leu Val Ser
 85 90 95
 Tyr Tyr Leu Pro Asn Glu Asp Ala Pro Val Ala Thr Ala Val Leu Arg
 100 105 110
 Leu Thr Gly Ile Val Val Ser Leu Asp Val Asp Ile Tyr Arg Ser Gly
 115 120 125
 Gln Val Glu Ile Ala Ser Asp Lys Gln Ala Lys Lys Asn Trp Val Trp
 130 135 140
 Gly Pro Ser Gly Trp Gly Ala Ile Leu Leu Val Asn Cys Ser Pro Ala
 145 150 155 160
 Asp Lys Gly Gln Ile Ile Asp Gln Lys Thr Thr Lys Val Phe Phe Pro
 165 170 175
 Glu Glu Ile Lys Ser Leu Ser Gln Met Thr Leu Asn Val Thr Arg Ala
 180 185 190
 Arg Arg His Leu Lys Lys Tyr Arg Leu Val Leu His Thr Ser Glu Glu
 195 200 205
 Glu Ala Arg Lys Ala Arg Val Tyr Arg Pro Gln Arg Asp Ser Ser Ser
 210 215 220
 Thr Phe Glu Val Leu Leu Gly Pro Gly Gln Cys Thr Tyr Thr Phe Ala
 225 230 235 240
 Pro Leu Glu Asn Asn Leu Lys Glu Thr Phe Tyr Val Glu Ala Ile Glu
 245 250 255
 Phe Pro Ser Ala Asp Phe Ser Gly Leu Ile Ser Tyr Ser Val Ser Leu
 260 265 270
 Val Glu Glu Pro Gln Asp Pro Ser Ile Pro Glu Thr Leu Val Tyr Lys
 275 280 285
 Asp Thr Val Val Phe Arg Val Ala Pro Cys Val Phe Ile Pro Ser Thr
 290 295 300
 Gln Met Pro Leu Glu Val Tyr Leu Cys Arg Glu Leu Gln Val Gln Gly
 305 310 315 320

HerrVETVAC01065-03.txt

Phe Val Asn Thr Val Met Glu Leu Ser Glu Lys Ser Asn Ser Gln Val
 325 330 335
 Ala Ser Val Tyr Glu Asp Pro Asn Arg Leu Gly Arg Trp Leu Gln Asp
 340 345 350
 Glu Met Ala Phe Cys Tyr Thr Gln Ala Pro His Arg Thr Leu Ser Leu
 355 360 365
 Val Leu Asp Thr Pro Arg Val Leu Thr Leu Glu Asp Phe Pro Met Lys
 370 375 380
 Tyr Ser Leu Ser Pro Gly Val Gly Tyr Val Ile Gln Cys Thr Lys Asp
 385 390 395 400
 His Arg Val Ala Ser Met Asp Ser Ile Gly Asn Leu Met Val Ser Pro
 405 410 415
 Pro Val Lys Val Glu Gly Lys Glu Tyr Pro Leu Gly Arg Ile Leu Ile
 420 425 430
 Gly Ser Cys Phe Tyr Pro Ser Lys Glu Gly Arg Asp Met Ser Lys Ala
 435 440 445
 Leu Arg Asp Phe Leu Tyr Ala Gln Arg Val Gln Ala Pro Val Glu Leu
 450 455 460
 Phe Ser Asp Trp Leu Met Val Gly His Met Asp Glu Phe Met Cys Phe
 465 470 475 480
 Ile Pro Thr Pro Asp Lys Ser Glu Gly Glu Lys Gly Phe Arg Leu Leu
 485 490 495
 Leu Ala Ser Pro Ser Ser Cys His Arg Leu Phe Glu Glu Lys Gln Lys
 500 505 510
 Glu Gly Tyr Gly Asp Met Ala Leu Phe Glu Glu Val Arg Glu Asp Gln
 515 520 525
 Leu Leu Ser Asn Gly Arg Gln Ala Asn Thr Ile Asn Gln Leu Leu Ala
 530 535 540
 Asp Lys Asn Met Arg Lys Gln Asn Asp Tyr Val Glu Lys Cys Ile Asn
 545 550 555 560
 Leu Asn Arg Asp Ile Leu Lys Lys Glu Leu Gly Leu Val Glu Arg Asp
 565 570 575
 Ile Ile Asp Ile Pro Gln Leu Phe Cys Leu Glu Gln Leu Thr Asn Val
 580 585 590

HerrVETVAC01065-03.txt

Pro Ser Ser Glu Gln Thr Gly Lys Phe Phe Ala Arg Pro Tyr Phe Pro
595 600 605

Asp Leu Leu Gln Met Ile Val Met Gly Lys Asn Leu Gly Ile Pro Lys
610 615 620

Pro Phe Gly Pro Gln Ile Lys Gly Thr Cys Cys Leu Glu Glu Lys Ile
625 630 635 640

Cys Gln Leu Leu Glu Pro Leu Gly Phe Lys Cys Thr Phe Ile Asp Asp
645 650 655

Phe Asp Cys Tyr Leu Thr Glu Ile Gly Asp Phe Cys Ala Cys Ala Asn
660 665 670

Ile Arg Arg Val Pro Phe Ala Phe Lys Trp Trp Arg Met Val Pro Glu
675 680 685

Pro Gln Ala
690

<210> 12
<211> 2372
<212> DNA
<213> canine

<400> 12
ggttatttga ggctgctgtg ctgacctcgg gttgtcgtgt aggtctgagg gtagtcggct 60
tggagagccg tgccatgtct ttccagagca tcatccacct gtccctggac agccctgtcc 120
gtgccctgtg tgtgctgggc atggaaatct gcttgatgt caatgggtgc gccccggaga 180
ggtgtgaggc gtttaccatc agcagctccc caggggtctc ggttgacctc cacggcacgc 240
tccccgggac gggccaggag gggacggccg tgacctggtg gcctctgtcg aatccacagg 300
atgtgctggt gaagatgacc tgccccagct ccgccagtga cggagaccag gttctggtct 360
cctactatct gccaatgag gacgcccccg tggccacggc cgtgctccgc ctcaccggga 420
ttgtggtctc cctggacgtg gacatctacc gcagtgggca ggtggagata gcgagtgaca 480
agcaggccaa gaaaaactgg gtctggggtc ctagcggttg gggtgccatc ctgcttgtga 540
actgcagccc tgccgacaag ggccagatca tagaccagaa gaccaccaag gtgttctttc 600
cagaggaaat aaagagcctg tcccagatga ccctgaacgt tacaagggcc cggcgccact 660
taaagaaata ccggctgggt ctccacacct ccgaggaaga ggcgaggaag gcccagagtct 720
atcggtccca aagagacagc tcgagcacct ttgaggtgct gctggggcca ggccagtga 780
cctacacgtt cgccccctg gagaacaacc tgaaggaaac cttctacgtg gaagccatcg 840
agttcccgtc tgccgacttc tcgggcctga tctcctactc cgtctccctg gtggaagaac 900
ctcaggaccc gtccattcca gagaccctgg tgtacaaaga cacggtagtg tttcgggtgg 960

HerrVETVAC01065-03.txt

```

ccccctgcgt ctttattccc agcaccacaga tgcctctgga ggtctatctg tgcagagagc 1020
tgcagggtcca ggggttttgtg aacacgggtga tggagctgag cgagaagagt aacagccagg 1080
tggcatctgt ctatgaggac cccaaccgcc tgggcagggtg gctccaggac gagatggcct 1140
tctgttacac ccaggcgccc cacaggacgc tctccttggt cctggacacc cctcgggtcc 1200
tcacgctgga agacttcccc atgaaatact cactgagccc cggggttggc tacgtgatcc 1260
agtgcaccaa ggaccacagg gtggccagca tggattccat tgggaacctg atggtgtccc 1320
cacctgtcaa ggtggaaggg aaagagtacc ccctaggcag gatcctcatt ggcagctgct 1380
tttaccgccag caaggagggg cgagacatga gcaaggccct ccgggacttc ctctaccccc 1440
agcgagtcca ggccccgggtg gagctcttct cggactggct gatggtgggc cacatggatg 1500
agttcatgtg cttcatcccc acaccggaca agagtgaggg tgaaaagggc ttccggctgc 1560
tactggccag cccagctcc tgccacagac tgctcgagga gaaacagaag gagggtctacg 1620
gggacatggc tctgtttgaa gaggtccggg aggaccagct cttttctaac gggcggcagg 1680
ccaacaccat caatcagctt ctggctgaca aaaacatgag aaagcagaat gactacgtag 1740
agaagtgcac caacctgaac cgcgacatcc tgaagaagga gctgggcctg gttgagaggg 1800
acatcatcga catcccgag ctcttctgcc tggagcagct gaccaacgtg ccctccagcg 1860
agcagaccgg caagttcttc gcgaggccct acttccctga cctgctgcag atgattgtga 1920
tgggcaagaa tctgggcatc cccaagcctt ttgggccccca gatcaagggg acctgctgcc 1980
tggaagaaaa aatctgccag ttgctcgagc ccctgggctt caagtgcacc ttcacgatg 2040
acttcgactg ctacctgacc gaaatcgggg atttctgtgc ctgtgccaac atccgccggg 2100
tgccctttgc cttcaaatgg tggaggatgg tgcccagacc ccaggcctag tgccgccagt 2160
cctggccccg tgcgagcggc ccactgatcg ccactccgca gcgttcaggt tccctgcctc 2220
accctgtgta cctgtccctt tgtcctgagg gtggcaaatg ctgcccgtgg gaacctcctc 2280
aaacggcctt cagccaagaa ggggttaccg agtgtccatt aaagttcaag ctgttctgaa 2340
tacaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 2372

```

<210> 13
 <211> 683
 <212> PRT
 <213> mus musculus

<400> 13

Met Ser Phe Gln Asn Ser Leu Ser Leu Ser Leu Val Asn Pro Thr His
 1 5 10 15

Ala Leu Cys Met Val Gly Met Glu Ile Thr Leu Asp Ile Ser Lys Cys
 20 25 30

Ala Pro Asp Lys Cys Lys Ser Phe Thr Ile Arg Gly Ser Pro Arg Ile
 35 40 45

HerrVETVAC01065-03.txt

Leu Ile His Ile Ser Ser Ser Val Ile Ala Gly Lys Glu Asp Thr Val
 50 55 60
 Val Trp Arg Ser Met Asn His Pro Thr Val Ala Leu Val Arg Met Val
 65 70 75 80
 Ala Pro Ser Pro Thr Val Asp Glu Asp Lys Val Leu Val Ser Tyr Phe
 85 90 95
 Cys Pro Asp Gln Glu Val Pro Thr Ala Thr Ala Val Leu Phe Leu Thr
 100 105 110
 Gly Ile Glu Ile Ser Leu Glu Ala Asp Ile Tyr Arg Asp Gly Gln Leu
 115 120 125
 Asp Met Pro Ser Asp Lys Gln Ala Lys Lys Lys Trp Met Trp Gly Met
 130 135 140
 Asn Gly Trp Gly Ala Ile Leu Leu Val Asn Cys Ser Pro Asn Ala Val
 145 150 155 160
 Gly Gln Pro Asp Glu Gln Ser Phe Gln Glu Gly Pro Arg Glu Ile Gln
 165 170 175
 Asn Asn Leu Ser Gln Met Asn Val Thr Val Glu Gly Pro Thr Ser Ile
 180 185 190
 Leu Gln Asn Tyr Gln Leu Ile Leu His Thr Ser Glu Glu Glu Ala Lys
 195 200 205
 Lys Thr Arg Val Tyr Trp Ser Gln Arg Gly Ser Ser Ala Tyr Glu Leu
 210 215 220
 Val Val Gly Pro Asn Lys Pro Val Tyr Leu Leu Pro Thr Phe Glu Asn
 225 230 235 240
 Arg Arg Lys Glu Ala Phe Tyr Val Glu Ala Thr Glu Phe Pro Ser Pro
 245 250 255
 Ser Phe Ser Gly Leu Ile Ser Leu Ser Leu Ser Leu Val Glu Lys Ala
 260 265 270
 His Asp Glu Cys Ile Pro Glu Ile Pro Leu Tyr Lys Asp Thr Val Met
 275 280 285
 Phe Arg Val Ala Pro Tyr Ile Phe Met Pro Ser Thr Gln Met Pro Leu
 290 295 300
 Glu Val Tyr Leu Cys Arg Glu Leu Gln Leu Gln Gly Phe Val Asp Ser
 305 310 315 320

HerrVETVAC01065-03.txt

Val Thr Lys Leu Ser Glu Lys Ser Lys Val Gln Val Val Lys Val Tyr
 325 330 335

Glu Asp Pro Asn Arg Gln Ser Lys Trp Leu Gln Asp Glu Met Ala Phe
 340 345 350

Cys Tyr Thr Gln Ala Pro His Lys Thr Val Ser Leu Ile Leu Asp Thr
 355 360 365

Pro Arg Val Ser Lys Leu Glu Asp Phe Pro Met Lys Tyr Thr Leu Thr
 370 375 380

Pro Gly Ser Gly Tyr Leu Ile Arg Gln Ile Glu Asp His Arg Val Ala
 385 390 395 400

Ser Leu Asp Ser Ile Gly Asn Leu Met Val Ser Pro Pro Val Lys Ala
 405 410 415

Gln Gly Lys Asp Tyr Pro Leu Gly Arg Val Leu Ile Gly Gly Ser Phe
 420 425 430

Tyr Pro Ser Ser Glu Gly Arg Asp Met Asn Lys Gly Leu Arg Glu Phe
 435 440 445

Val Tyr Ala Gln Gln Val Gln Ala Pro Val Glu Leu Phe Ser Asp Trp
 450 455 460

Leu Met Thr Gly His Met Asp Gln Phe Met Cys Phe Val Pro Thr Asn
 465 470 475 480

Asp Lys Asn Asn Asp Gln Lys Asp Phe Arg Leu Leu Leu Ala Ser Pro
 485 490 495

Ser Ala Cys Phe Glu Leu Phe Glu Gln Lys Gln Lys Glu Gly Tyr Gly
 500 505 510

Asn Val Thr Leu Phe Glu Asp Ile Gly Ala Glu Gln Leu Leu Ser Asn
 515 520 525

Gly Arg Glu Ser Lys Thr Ile Ser Gln Ile Leu Ala Asp Lys Ser Phe
 530 535 540

Arg Glu Gln Asn Thr Tyr Val Glu Lys Cys Ile Ser Leu Asn Arg Thr
 545 550 555 560

Leu Leu Lys Thr Glu Leu Gly Leu Glu Asp Lys Asp Ile Ile Leu Ile
 565 570 575

Pro Gln Leu Phe Cys Leu Glu Gln Leu Thr Asn Val Pro Ser Asn Gln
 580 585 590

HerrVETVAC01065-03.txt

Gln Ser Thr Lys Leu Phe Ala Arg Pro Tyr Phe Pro Asp Met Leu Gln
 595 600 605

Ile Ile Val Leu Gly Lys Asn Leu Gly Ile Pro Lys Pro Phe Gly Pro
 610 615 620

Lys Ile Asn Gly Thr Cys Cys Leu Glu Glu Lys Val Cys Gly Leu Leu
 625 630 635 640

Glu Pro Leu Gly Leu Lys Cys Thr Phe Ile Asp Asp Phe Asp Cys Tyr
 645 650 655

Leu Ala Asn Ile Gly Asp Val Cys Ala Ser Ala Ile Ile Asn Arg Val
 660 665 670

Pro Phe Ala Phe Lys Trp Trp Lys Met Thr Pro
 675 680

<210> 14
 <211> 2336
 <212> DNA
 <213> mus musculus

<400> 14
 ggactgctga cagtggctag cttggtaagc ccagccatgt cttttcagaa ctcaactcagc 60
 ctgtctcttg tcaatccac ccatgccctc tgcattgtag gcatggaaat caccttggac 120
 atcagcaagt gtgcaccgga caagtgaag tctttcacca tccgtgggtc ccccaggatc 180
 ttgatccaca tctctagctc cgtcatcgct ggcaaagagg aactgtgggt ctggagggtca 240
 atgaaccatc ccacagtggc attggtgagg atggtggcgc ccagccccac tgtggatgaa 300
 gacaaggtgc tgggtctccta cttctgtcct gaccaagaag tccccacggc cacagctgtg 360
 ctgtttctca ccggcatcga gatctccctg gaggcagaca tctatcgaga tggacaactg 420
 gacatgccaa gtgataagca agctaagaaa aaatggatgt ggggtatgaa cggctgggga 480
 gccatcctgc ttgtgaattg tagccctaatt gctgtgggccc agcctgatga acagtccttt 540
 caggagggcc ccagagaaat acagaacaac ctgtctcaga tgaatgtaac tgtggagggc 600
 cccaccagca tctacagaa ttaccagttg atctacata cctccgaaga agaggcgaag 660
 aagacaagag tctactgggtc tcagagaggc tctctgcgt atgaactgggt ggtgggaccc 720
 aacaagcctg tctatctcct gcctaccttt gagaaccgta ggaaagaggc tttctacgta 780
 gaagccacgg aattcccatc tcccagcttc tcgggcctga tctcctgtc actctcccta 840
 gtagaaaagg ctacagacga gtgcattcca gagattccgc tctataagga tacagtgatg 900
 ttccgggttg caccttatat cttcatgccc agcaccagga tgcctctaga ggtttacctg 960
 tgcaggggagc tacagctgca aggcctttgtg gactcagtga ccaagctgag cgagaagagc 1020
 aaagtgcagg tggtaaaggc ctatgaggac cccaaccgcc agagcaagtg gctccaggac 1080
 gagatggctt tctgctatac tcaggctcct cacaagacgg tgtcattgat ccttgacacc 1140

HerrVETVAC01065-03.txt

```

ccaaggggttt ccaagctgga agacttcccc atgaaatata cactgacccc tggctctggc 1200
tacctgatcc gacaaattga ggaccaccgg gtggctagcc tggattccat cgggaacctg 1260
atggtatctc cgcctgtcaa ggctcagggc aaagactacc ctctagggag ggtcctcatt 1320
ggtggcagct tttaacccag ctctgagggc cgggacatga acaagggcct gcgagaattc 1380
gtgtatgccc agcagggtgca ggcccctgtg gaactcttct cggactggct gatgaccggt 1440
cacatggatc aattcatgtg ctttgtccct accaatgata aaaacaacga ccagaaggac 1500
ttccgcctgc tgctggccag cccagtgcc tgctttgagc tggtcgaaca gaagcagaag 1560
gaaggctatg ggaacgtgac cctgtttgaa gacattggag cagaacagct cttttctaata 1620
gggagggaga gcaaaaactat ttcccaaadc ctggctgaca agagttttcg agagcagaac 1680
acctatgttg agaagtgtat cagcctgaac cgcaccctcc tgaagacaga actgggattg 1740
gaggacaagg acatcatcct gatcccgag ctcttctgcc tggagcagct gacgaatgtc 1800
ccctccaacc agcagagcac caaactctc gcgaggccgt acttccccga catgctgcag 1860
ataatcgtgt tgggcaagaa ccttggaatc cccaagccct ttggggccaa aatcaatggc 1920
acctgctgcc tagaagagaa agtgtgtgga ttactggagc ccctgggtct caagtgcacc 1980
ttcattgatg attttgactg ctacctggcc aacatagggg acgtctgtgc cagtgccatc 2040
ataaacaggg tgccatttgc attcaagtgg tggaagatga cccataaac ccctggccct 2100
ggcacggcca gtccgcgcca gtacggatgg cttttgcat agatagtagt gggtgcgagc 2160
gttggtgttg cactgggttg aagggacgga gagctgggag ttaggggtctc tcacatctac 2220
cagcttgaca cttctggagg ggaaaaggga aaagagcgcc tatgtaaaca aattgccata 2280
gagccaataa agcatggtat tctgaatata aaaaaaaaaa aaaaaaaaaa aaaaaa 2336

```

<210> 15
 <211> 164
 <212> PRT
 <213> mus musculus

<400> 15

Met Ala Ser His Thr Ala Asp Ala Asp Ala Lys Pro Asp Ser Asp Ser
 1 5 10 15

Gln Lys Leu Leu Asn Val Leu Pro Val Ser Leu Arg Leu Arg Thr Arg
 20 25 30

Pro Trp Trp Phe Pro Ile Gln Glu Val Ser Asn Pro Leu Val Leu Tyr
 35 40 45

Met Glu Ala Trp Val Ala Glu Arg Val Ile Gly Thr Asp Gln Ala Glu
 50 55 60

Ile Ser Glu Ile Glu Trp Met Cys Gln Ala Leu Leu Thr Val Asp Ser
 65 70 75 80

HerrVETVAC01065-03.txt

Val Asn Ser Gly Asn Leu Ala Glu Ile Thr Ile Phe Gly Gln Pro Ser
 85 90 95

Ala Gln Thr Arg Met Lys Asn Ile Leu Leu Asn Met Ala Ala Trp His
 100 105 110

Lys Glu Asn Glu Leu Gln Arg Ala Val Lys Val Lys Glu Val Glu Glu
 115 120 125

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

Gly Leu Lys Leu Ala Gly Phe Pro Leu Pro Leu Glu Gly Arg Glu Thr
 145 150 155 160

Gln Met Glu Ser

<210> 16
 <211> 706
 <212> DNA
 <213> mus musculus

<400> 16
 gagctttgaa aagcgagacc agaccggctg ccaagcagcg ccccgctcctc ctgcttgcag 60
 tatggcatcc cacacggctg atgctgacgc caagccagac tctgactctc agaagctgct 120
 taacgtcctg cctgtgtccc tgagacttcg cactcggccc tgggtggttcc caattcagga 180
 agtcagtaat cctctggtgc tctacatgga agcctgggtg gcagaaaggg taataggcac 240
 agaccaagcg gaaatctcag aaatagagtg gatgtgccaa gccctgctga cagtggactc 300
 cgtcaactct gggaacctag ctgaaatcac catctttgga caaccagtg cacagacgcg 360
 aatgaaaaat attctcttga acatggcggc ctggcacaaa gaaaacgaac tccaaagagc 420
 tgtgaaggtg aaagaagttg aagagttctt gaaaattcgt gcctcctcaa tcctaagcaa 480
 gttaagtaag aaagggctaa aactggctgg ctttccgctt ccgctggagg gaagagaaac 540
 acagatggag tcttaaatca ctacacgtat tttcttacct gccttcacga gcacctcctg 600
 tttatttttc tttgcaaata gcaatttgta gaactattcc aatataaaaa attttttttg 660
 gtatgctgtt gggttgatca catacagtaa aatttataaa cactcg 706

<210> 17
 <211> 67
 <212> PRT
 <213> mus musculus

<400> 17

Met Glu Leu Ser Ser Gly Val Cys Pro Ala Thr Arg Leu Gln Glu Ala
 1 5 10 15

Glu Lys Ala Ala Val His Lys Arg Ser Pro Lys Val Leu Glu Ala Leu
 20 25 30

HerrVETVAC01065-03.txt

Arg Lys Leu Asn Ile Gln Ala Asp Gln Ala Pro Val Ile Ala Val Leu
35 40 45

Gly Ser Gly Gly Gly Leu Arg Ala His Ile Ala Cys Leu Gly Val Leu
50 55 60

Ser Glu Leu
65