

(19) AUSTRALIAN PATENT OFFICE

(54) Title
Gene regulation in transgenic animals using a transposon-based vector

(51)⁶ International Patent Classification(s)
C12N 15/85 (2006.01) 20060101ALI2005100
A01K 67/027 ^{8BMEP} **C07K**
(2006.01) 16/02
C07K 14/47 (2006.01) 20060101ALI2005100
C07K 14/62 (2006.01) ^{8BMEP} **C07K**
C07K 16/00 (2006.01) 16/10
C07K 16/02 (2006.01) 20060101ALI2005100
C07K 16/10 (2006.01) ^{8BMEP} **C07K**
C07K 16/18 (2006.01) 16/18
C07K 16/26 (2006.01) 20060101ALI2006031
C12N 5/10 (2006.01) ^{OBMJP} **C07K**
C12N 15/09 (2006.01) 16/26
A61K 39/00 (2006.01) 20060101ALI2005100
A61K 48/00 (2006.01) ^{8BMEP} **C12N**
C12N 15/85 5/10
20060101AFI2005100 20060101ALI2006031
^{8BMEP} **A01K** ^{OBMJP} **C12N**
67/027 15/09
20060101ALI2006072 20060101ALI2006031
^{2BMEP} **C07K** ^{OBMJP} **A61K**
14/47 39/00
20060101ALI2006031 20060101ALN200510
^{OBMJP} **C07K** ^{08BMEP} **A61K**
14/62 48/00
20060101ALI2005100 20060101ALN200510
^{8BMEP} **C07K** ^{08BMEP}
16/00 PCT/US2003/020389

(21) Application No: 2003261096

(22) Application Date: 2003 .06 .26

(87) WIPO No: W004/003157

(30) Priority Data

(31) Number	(32) Date	(33) Country
60/441,392	2003 .01 .21	US
60/441,377	2003 .01 .21	US
60/441,405	2003 .01 .21	US
60/441,381	2003 .01 .21	US
60/441,447	2003 .01 .21	US
60/441,502	2003 .01 .21	US
60/392,415	2002 .06 .26	US

(43) Publication Date : 2004 .01 .19

(43) Publication Journal Date : 2004 .03 .04

(71) Applicant(s)

TransGenRx, Inc., The Board of Supervisors of Louisiana State University and Agricultural and Mechanical College

(72) Inventor(s)

Deboer, Kenneth F.; Cadd, Gary G.; Fioretti, William C.; Cooper, Richard K.

(74) Agent/Attorney

FB Rice & Co, Level 23 200 Queen Street, Melbourne, VIC, 3000

(56) Related Art

US 6, 218, 185 B1
Kozak et al; J Mol Bio (1987); Vol 196: 947-50
WO 1997/047739 A1
Schneider et al; Gene (1997); Vol 197: 337-41

5

GENE REGULATION IN TRANSGENIC ANIMALS USING A TRANSPOSON-BASED VECTOR

10 The U.S. Government has certain rights in this invention. The development of
this invention was partially funded by the United States Government under a HATCH
grant from the United States Department of Agriculture, partially funded by the
United States Government with Formula 1433 funds from the United States
Department of Agriculture and partially funded by the United States Government
15 under contract DAAD 19-02016 awarded by the Army.

FIELD OF THE INVENTION

 The present invention relates generally to cell-specific gene regulation in
transgenic animals. Animals may be made transgenic through administration of a
20 transposon-based vector through any method of administration including pronuclear
injection, or intraembryonic, intratesticular, intraoviductal or intravenous
administration. These transgenic animals contain the gene of interest in all cells,
including germ cells. Animals may also be made transgenic by targeting specific cells
for uptake and gene incorporation of the transposon-based vectors. Stable
25 incorporation of a gene of interest into cells of the transgenic animals is demonstrated
by expression of the gene of interest in a cell, wherein expression is regulated by a
promoter sequence. The promoter sequence may be provided as a transgene along
with the gene of interest or may be endogenous to the cell. The promoter sequence
may be constitutive or inducible, wherein inducible promoters include tissue-specific
30 promoters, developmentally regulated promoters and chemically inducible promoters.

BACKGROUND OF THE INVENTION

 Transgenic animals are desirable for a variety of reasons, including their
potential as biological factories to produce desired molecules for pharmaceutical,
35 diagnostic and industrial uses. This potential is attractive to the industry due to the
inadequate capacity in facilities used for recombinant production of desired molecules
and the increasing demand by the pharmaceutical industry for use of these facilities.
Numerous attempts to produce transgenic animals have met several problems,

ATLLIB02 133492.1

including low rates of gene incorporation and unstable gene incorporation. Accordingly, improved gene technologies are needed for the development of transgenic animals for the production of desired molecules.

Improved gene delivery technologies are also needed for the treatment of disease in animals and humans. Many diseases and conditions can be treated with gene-delivery technologies, which provide a gene of interest to a patient suffering from the disease or the condition. An example of such disease is Type 1 diabetes. Type 1 diabetes is an autoimmune disease that ultimately results in destruction of the insulin producing β -cells in the pancreas. Although patients with Type 1 diabetes may be treated adequately with insulin injections or insulin pumps, these therapies are only partially effective. Insulin replacement, such as via insulin injection or pump administration, cannot fully reverse the defect in the vascular endothelium found in the hyperglycemic state (Pieper et al., 1996. Diabetes Res. Clin. Pract. Suppl. S157-S162). In addition, hyper- and hypoglycemia occurs frequently despite intensive home blood glucose monitoring. Finally, careful dietary constraints are needed to maintain an adequate ratio of consumed calories consumed. This often causes major psychosocial stress for many diabetic patients. Development of gene therapies providing delivery of the insulin gene into the pancreas of diabetic patients could overcome many of these problems and result in improved life expectancy and quality of life.

Several of the prior art gene delivery technologies employed viruses that are associated with potentially undesirable side effects and safety concerns. The majority of current gene-delivery technologies useful for gene therapy rely on virus-based delivery vectors, such as adeno and adeno-associated viruses, retroviruses, and other viruses, which have been attenuated to no longer replicate. (Kay, M.A., et al. 2001. Nature Medicine 7:33-40).

There are multiple problems associated with the use of viral vectors. First, they are not tissue-specific. In fact, a gene therapy trial using adenovirus was recently halted because the vector was present in the patient's sperm (Gene trial to proceed despite fears that therapy could change child's genetic makeup. The New York Times, December 23, 2001). Second, viral vectors are likely to be transiently incorporated, which necessitates re-treating a patient at specified time intervals. (Kay, M.A., et al. 2001. Nature Medicine 7:33-40). Third, there is a concern that a viral-based vector could revert to its virulent form and cause disease. Fourth, viral-based vectors require a dividing cell for stable integration. Fifth, viral-based vectors indiscriminately integrate into various cells and tissues, which can result in undesirable germline integration. Sixth, the required high titers needed to achieve the desired effect have resulted in the death of one patient and they are believed to be

responsible for induction of cancer in a separate study. (Science, News of the Week, October 4, 2002).

Accordingly, what is needed is a new vector to produce transgenic animals and humans with stably incorporated genes, which vector does not cause disease or other unwanted side effects. There is also a need for DNA constructs that would be stably incorporated into the tissues and cells of animals and humans, including cells in the resting state, which are not replicating. There is a further recognized need in the art for DNA constructs capable of delivering genes to specific tissues and cells of animals and humans.

When incorporating a gene of interest into an animal for the production of a desired protein or when incorporating a gene of interest in an animal or human for the treatment of a disease, it is often desirable to selectively activate incorporated genes using inducible promoters. These inducible promoters are regulated by substances either produced or recognized by the transcription control elements within the cell in which the gene is incorporated. In many instances, control of gene expression is desired in transgenic animals or humans so that incorporated genes are selectively activated at desired times and/or under the influence of specific substances. Accordingly, what is needed is a means to selectively activate genes introduced into the genome of cells of a transgenic animal or human. This can be taken a step further to cause incorporation to be tissue-specific, which prevents widespread gene incorporation throughout a patient's body (animal or human). This decreases the amount of DNA needed for a treatment, decreases the chance of incorporation in gametes, and targets gene delivery, incorporation, and expression to the desired tissue where the gene is needed to function.

SUMMARY OF THE INVENTION

The present invention addresses the problems described above by providing new, effective and efficient compositions for producing transgenic animals and for treating disease in animals or humans. Transgenic animals include all egg-laying animals and milk-producing animals. Transgenic animals further include but are not limited to avians, fish, amphibians, reptiles, insects, mammals and humans. In a preferred embodiment, the animal is an avian animal. In another preferred embodiment, the animal is a milk-producing animal, including but not limited to bovine, porcine, ovine and equine animals. Animals are made transgenic through administration of a composition comprising a transposon-based vector designed for stable incorporation of a gene of interest for production of a desired protein, together

with an acceptable carrier. A transfection reagent is optionally added to the composition before administration.

The transposon-based vectors of the present invention include a transposase, operably-linked to a first promoter, and a coding sequence for a protein or peptide of interest operably-linked to a second promoter, wherein the coding sequence for the protein or peptide of interest and its operably-linked promoter are flanked by transposase insertion sequences recognized by the transposase. The transposon-based vector also includes the following characteristics: a) one or more modified Kozak sequences comprising ACCATG (SEQ ID NO:13) at the 3' end of the first promoter to enhance expression of the transposase; b) modifications of the codons for the first several N-terminal amino acids of the transposase, wherein the nucleotide at the third base position of each codon was changed to an A or a T without changing the corresponding amino acid; c) addition of one or more stop codons to enhance the termination of transposase synthesis; and/or, d) addition of an effective polyA sequence operably-linked to the transposase to further enhance expression of the transposase gene.

Use of the compositions of the present invention results in highly efficient and stable incorporation of a gene of interest into the genome of transfected animals. For example, transgenic avians have been mated and produce transgenic progeny in the G1 generation. The transgenic progeny have been mated and produce transgenic progeny in the G2 generation.

The present invention also provides for tissue-specific incorporation and/or expression of a gene of interest. Tissue-specific incorporation of a gene of interest may be achieved by placing the transposase gene under the control of a tissue-specific promoter, whereas tissue-specific expression of a gene of interest may be achieved by placing the gene of interest under the control of a tissue-specific promoter. In some embodiments, the gene of interest is transcribed under the influence of an ovalbumin, or other oviduct specific, promoter. Linking the gene of interest to an oviduct specific promoter in an egg-laying animal results in synthesis of a desired molecule and deposition of the desired molecule in a developing egg. The present invention further provides for stable incorporation and expression of genes in the epithelial cells of the mammary gland in milk-producing animals. Transcription of the gene of interest in the epithelial cells of the mammary gland results in synthesis of a desired molecule

and deposition of the desired molecule in the milk. A preferred molecule is a protein. In some embodiments, the desired molecule deposited in the milk is an antiviral protein, an antibody, or a serum protein.

5 In other embodiments, specific incorporation of the proinsulin gene into liver cells of a diabetic animal results in the improvement of the animal's condition. Such improvement is achieved by placing a transposase gene under the control of a liver-specific promoter, which drives integration of the gene of interest in liver cells of the diabetic animal.

The present invention advantageously produces a high number of transgenic
10 animals having a gene of interest stably incorporated. These transgenic animals successfully pass the desired gene to their progeny. The transgenic animals of the present invention also produce large amounts of a desired molecule encoded by the transgene. Transgenic egg-laying animals, particularly avians, produce large amounts of a desired protein that is deposited in the egg for rapid harvest and purification.
15 Transgenic milk-producing animals produce large amounts of a desired protein that is deposited in the milk for rapid harvest and purification.

Any desired gene may be incorporated into the novel transposon-based vectors of the present invention in order to synthesize a desired molecule in the transgenic animals. Proteins, peptides and nucleic acids are preferred desired molecules to be
20 produced by the transgenic animals of the present invention. Particularly preferred proteins are antibody proteins.

This invention provides a composition useful for the production of transgenic hens capable of producing substantially high amounts of a desired protein or peptide. Entire flocks of transgenic birds may be developed very quickly in order to produce
25 industrial amounts of desired molecules. The present invention solves the problems inherent in the inadequate capacity of fermentation facilities used for bacterial production of molecules and provides a more efficient and economical way to produce desired molecules. Accordingly, the present invention provides a means to produce large amounts of therapeutic, diagnostic and reagent molecules.

30 Transgenic chickens are excellent in terms of convenience and efficiency of manufacturing molecules such as proteins and peptides. Starting with a single transgenic rooster, thousands of transgenic offspring can be produced within a year. (In principle, up to forty million offspring could be produced in just three

generations). Each transgenic female is expected to lay at least 250 eggs/year, each potentially containing hundreds of milligrams of the selected protein. Flocks of chickens numbering in the hundreds of thousands are readily handled through established commercial systems. The technologies for obtaining eggs and
5 fractionating them are also well known and widely accepted. Thus, for each therapeutic, diagnostic, or other protein of interest, large amounts of a substantially pure material can be produced at relatively low incremental cost.

A wide range of recombinant peptides and proteins can be produced in transgenic egg-laying animals and milk-producing animals. Enzymes, hormones,
10 antibodies, growth factors, serum proteins, commodity proteins, biological response modifiers, peptides and designed proteins may all be made through practice of the present invention. For example, rough estimates suggest that it is possible to produce in bulk growth hormone, insulin, or Factor VIII, and deposit them in transgenic egg whites, for an incremental cost in the order of one dollar per gram. At such prices it
15 is feasible to consider administering such medical agents by inhalation or even orally, instead of through injection. Even if bioavailability rates through these avenues were low, the cost of a much higher effective dose would not be prohibitive.

In one embodiment, the egg-laying transgenic animal is an avian. The method of the present invention may be used in avians including Ratites, Psittaciformes,
20 Falconiformes, Piciformes, Strigiformes, Passeriformes, Coraciformes, Ralliformes, Cuculiformes, Columbiformes, Galliformes, Anseriformes, and Herodiones. Preferably, the egg-laying transgenic animal is a poultry bird. More preferably, the bird is a chicken, turkey, duck, goose or quail. Another preferred bird is a ratite, such as, an emu, an ostrich, a rhea, or a cassowary. Other preferred birds are partridge,
25 pheasant, kiwi, parrot, parakeet, macaw, falcon, eagle, hawk, pigeon, cockatoo, song birds, jay bird, blackbird, finch, warbler, canary, toucan, mynah, or sparrow.

In another embodiment, the transgenic animal is a milk-producing animal, including but not limited to bovine, ovine, porcine, equine, and primate animals. Milk-producing animals include but are not limited to cows, goats, horses, pigs,
30 buffalo, rabbits, non-human primates, and humans.

Accordingly, it is an object of the present invention to provide novel transposon-based vectors.

It is another object of the present invention to provide novel transposon-based vectors that encode for the production of desired proteins or peptides in cells.

It is an object of the present invention to produce transgenic animals through administration of a transposon-based vector.

5 Another object of the present invention is to produce transgenic animals through administration of a transposon-based vector, wherein the transgenic animals produce desired proteins or peptides.

Yet another object of the present invention is to produce transgenic animals through administration of a transposon-based vector, wherein the transgenic animals
10 produce desired proteins or peptides and deposit the proteins or peptides in eggs or milk.

It is a further object of the present invention to produce transgenic animals through intraembryonic, intratesticular or intraoviductal administration of a transposon-based vector.

15 It is further an object of the present invention to provide a method to produce transgenic animals through administration of a transposon-based vector that are capable of producing transgenic progeny.

Yet another object of the present invention is to provide a method to produce transgenic animals through administration of a transposon-based vector that are
20 capable of producing a desired molecule, such as a protein, peptide or nucleic acid.

Another object of the present invention is to provide a method to produce transgenic animals through administration of a transposon-based vector, wherein such administration results in modulation of endogenous gene expression.

It is another object of the present invention to provide transposon-vectors
25 useful for cell- or tissue-specific expression of a gene of interest in an animal or human with the purpose of gene therapy.

It is yet another object of the present invention to provide a method to produce transgenic avians through administration of a transposon-based vector that are capable of producing proteins, peptides or nucleic acids.

30 It is another object of the present invention to produce transgenic animals through administration of a transposon-based vector encoding an antibody or a fragment thereof.

Still another object of the present invention is to provide a method to produce transgenic avians through administration of a transposon-based vector that are capable of producing proteins or peptides and depositing these proteins or peptides in the egg.

Another object of the present invention is to provide transgenic avians that
5 contain a stably incorporated transgene.

Still another object of the present invention is to provide eggs containing desired proteins or peptides encoded by a transgene incorporated into the transgenic avian that produces the egg.

It is further an object of the present invention to provide a method to produce
10 transgenic milk-producing animals through administration of a transposon-based vector that are capable of producing proteins, peptides or nucleic acids.

Still another object of the present invention is to provide a method to produce transgenic milk-producing animals through administration of a transposon-based vector that are capable of producing proteins or peptides and depositing these proteins
15 or peptides in their milk.

Another object of the present invention is to provide transgenic milk-producing animals that contain a stably incorporated transgene.

Another object of the present invention is to provide transgenic milk-producing animals that are capable of producing proteins or peptides and depositing
20 these proteins or peptides in their milk.

Yet another object of the present invention is to provide milk containing desired molecules encoded by a transgene incorporated into the transgenic milk-producing animals that produce the milk.

Still another object of the present invention is to provide milk containing
25 desired proteins or peptides encoded by a transgene incorporated into the transgenic milk-producing animals that produce the milk.

A further object of the present invention to provide a method to produce transgenic sperm through administration of a transposon-based vector to an animal.

A further object of the present invention to provide transgenic sperm that
30 contain a stably incorporated transgene.

An advantage of the present invention is that transgenic animals are produced with higher efficiencies than observed in the prior art.

Another advantage of the present invention is that these transgenic animals possess high copy numbers of the transgene.

Another advantage of the present invention is that the transgenic animals produce large amounts of desired molecules encoded by the transgene.

5 Still another advantage of the present invention is that desired molecules are produced by the transgenic animals much more efficiently and economically than prior art methods, thereby providing a means for large scale production of desired molecules, particularly proteins and peptides.

According to the invention there is also provided a vector comprising:

10 a) a modified transposase gene operably-linked to a first promoter; wherein the nucleic acid sequence 3' to the first promoter comprises the sequence as set forth in SEQ ID NO:13, wherein SEQ ID NO:13 contains the Kozak sequence and a start codon for the transposase, and wherein at least one of the first twenty codons for the transposase gene are modified from the wild-type sequence by changing a nucleotide at
15 a third base position of the codon to an adenine or thymine without modifying the amino acid encoded by the codon, and

b) one or more genes of interest operably-linked to one or more additional promoters; and wherein the one or more genes of interest and their operably-linked promoters are flanked by transposase insertion sequences recognized by the transposase
20 encoded by the modified transposase gene.

According to the invention there is also provided a method of producing a transgenic animal comprising administering to the animal a vector according to the invention.

25 According to the invention there is also provided an egg produced by the transgenic avian animal according to the invention, wherein the egg contains one or more desired proteins encoded by the one or more genes of interest.

According to the invention there is also provided a transgenic sperm produced by the transgenic animal produced according to the invention.

30 According to the invention there is also provided a method for producing a desired protein comprising:

- a) administering to the animal a vector according to the invention; and
- b) isolating the desired protein produced in the animal.

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated
35 element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps. Any

discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present invention as it existed before the priority date of each claim of this application.

These and other objects, features and advantages of the present invention will become apparent after a review of the following detailed description of the disclosed embodiments and claims.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 depicts schematically a transposon-based vector containing a transposase operably linked to a first promoter and a gene of interest operably-linked to a second promoter, wherein the gene of interest and its operably-linked promoter are flanked by insertion sequences (IS) recognized by the transposase. "Pro" designates a promoter. In this and subsequent figures, the size of the actual nucleotide sequence is not necessarily proportionate to the box representing that sequence.

Figure 2 depicts schematically a transposon-based vector for targeting deposition of a polypeptide in an egg white wherein Ov pro is the ovalbumin promoter, Ov protein is the ovalbumin protein and PolyA is a polyadenylation sequence. The TAG sequence includes a spacer, the gp41 hairpin loop from HIV I and a protein cleavage site.

Figure 3 depicts schematically a transposon-based vector for targeting deposition of a polypeptide in an egg white wherein Ovo pro is the ovomucoid promoter and Ovo SS is the ovomucoid signal sequence. The TAG sequence includes a spacer, the gp41 hairpin loop from HIV I and a protein cleavage site.

Figure 4 depicts schematically a transposon-based vector for targeting deposition of a polypeptide in an egg yolk wherein Vit pro is the vitellogenin promoter and Vit targ is the vitellogenin targeting sequence.

5 Figure 5 depicts schematically a transposon-based vector for expression of antibody heavy and light chains. Prepro indicates a prepro sequence from cecropin and pro indicates a pro sequence from cecropin.

10 Figure 6 depicts schematically a transposon-based vector for expression of antibody heavy and light chains. Ent indicates an enterokinase cleavage sequence.

15 Figure 7 depicts schematically egg white targeted expression of antibody heavy and light chains from one vector in either tail-to-tail (Figure 7A) or tail-to-head (Figure 7B) configuration. In the tail-to-tail configuration, the ovalbumin signal sequence adjacent to the gene for the light chain contains on its 3' end an enterokinase cleavage site (not shown) to allow cleavage of the signal sequence from the light chain, and the ovalbumin signal sequence adjacent to the gene for the heavy chain contains on its 5' end an enterokinase cleavage site (not shown) to allow cleavage of the signal sequence from the heavy chain. In the tail-to-head configuration, the
20 ovalbumin signal sequence adjacent to the gene for the heavy chain and the light chain contains on its 3' end an enterokinase cleavage site (not shown) to allow cleavage of the signal sequence from the heavy or light chain.

DETAILED DESCRIPTION OF THE INVENTION

25 The present invention provides a new, effective and efficient method of producing transgenic animals, particularly egg-laying animals and milk-producing animals, through administration of a composition comprising a transposon-based vector designed for stable incorporation of a gene of interest for production of a desired molecule.

30 Definitions

It is to be understood that as used in the specification and in the claims, "a" or "an" can mean one or more, depending upon the context in which it is used. Thus, for example, reference to "a cell" can mean that at least one cell can be utilized.

The term "antibody" is used interchangeably with the term "immunoglobulin" and is defined herein as a protein synthesized by an animal or a cell of the immune system in response to the presence of a foreign substance commonly referred to as an "antigen" or an "immunogen". The term antibody includes fragments of antibodies.

- 5 Antibodies are characterized by specific affinity to a site on the antigen, wherein the site is referred to an "antigenic determinant" or an "epitope". Antigens can be naturally occurring or artificially engineered. Artificially engineered antigens include but are not limited to small molecules, such as small peptides, attached to haptens such as macromolecules, for example proteins, nucleic acids, or polysaccharides.
- 10 Artificially designed or engineered variants of naturally occurring antibodies and artificially designed or engineered antibodies not occurring in nature are all included in the current definition. Such variants include conservatively substituted amino acids and other forms of substitution as described in the section concerning proteins and polypeptides.

- 15 As used herein, the term "egg-laying animal" includes all amniotes such as birds, turtles, lizards and monotremes. Monotremes are egg-laying mammals and include the platypus and echidna. The term "bird" or "fowl," as used herein, is defined as a member of the Aves class of animals which are characterized as warm-blooded, egg-laying vertebrates primarily adapted for flying. Avians include, without
- 20 limitation, Ratites, Psittaciformes, Falconiformes, Piciformes, Strigiformes, Passeriformes, Coraciiformes, Ralliformes, Cuculiformes, Columbiformes, Galliformes, Anseriformes, and Herodiones. The term "Ratite," as used herein, is defined as a group of flightless, mostly large, running birds comprising several orders and including the emus, ostriches, kiwis, and cassowaries. The term "Psittaciformes",
- 25 as used herein, includes parrots and refers to a monofamilial order of birds that exhibit zygodactylism and have a strong hooked bill. A "parrot" is defined as any member of the avian family Psittacidae (the single family of the Psittaciformes), distinguished by the short, stout, strongly hooked beak. The term "chicken" as used herein denotes chickens used for table egg production, such as egg-type chickens, chickens reared for
- 30 public meat consumption, or broilers, and chickens reared for both egg and meat production ("dual-purpose" chickens). The term "chicken" also denotes chickens produced by primary breeder companies, or chickens that are the parents,

grandparents, great-grandparents, etc. of those chickens reared for public table egg, meat, or table egg and meat consumption.

The term "egg" is defined herein as a large female sex cell enclosed in a porous, calcarous or leathery shell, produced by birds and reptiles. The term "ovum" is defined as a female gamete, and is also known as an egg. Therefore, egg production in all animals other than birds and reptiles, as used herein, is defined as the production and discharge of an ovum from an ovary, or "ovulation". Accordingly, it is to be understood that the term "egg" as used herein is defined as a large female sex cell enclosed in a porous, calcarous or leathery shell, when a bird or reptile produces it, or it is an ovum when it is produced by all other animals.

The term "milk-producing animal" refers herein to mammals including, but not limited to, bovine, ovine, porcine, equine, and primate animals. Milk-producing animals include but are not limited to cows, llamas, camels, goats, reindeer, zebu, water buffalo, yak, horses, pigs, rabbits, non-human primates, and humans.

The term "gene" is defined herein to include a coding region for a protein, peptide or polypeptide.

The term "vector" is used interchangeably with the terms "construct", "DNA construct" and "genetic construct" to denote synthetic nucleotide sequences used for manipulation of genetic material, including but not limited to cloning, subcloning, sequencing, or introduction of exogenous genetic material into cells, tissues or organisms, such as birds. It is understood by one skilled in the art that vectors may contain synthetic DNA sequences, naturally occurring DNA sequences, or both. The vectors of the present invention are transposon-based vectors as described herein.

When referring to two nucleotide sequences, one being a regulatory sequence, the term "operably-linked" is defined herein to mean that the two sequences are associated in a manner that allows the regulatory sequence to affect expression of the other nucleotide sequence. It is not required that the operably-linked sequences be directly adjacent to one another with no intervening sequence(s).

The term "regulatory sequence" is defined herein as including promoters, enhancers and other expression control elements such as polyadenylation sequences, matrix attachment sites, insulator regions for expression of multiple genes on a single construct, ribosome entry/attachment sites, introns that are able to enhance expression, and silencers.

Transposon-Based Vectors

While not wanting to be bound by the following statement, it is believed that the nature of the DNA construct is an important factor in successfully producing transgenic animals. The "standard" types of plasmid and viral vectors that have previously been almost universally used for transgenic work in all species, especially avians, have low efficiencies and may constitute a major reason for the low rates of transformation previously observed. The DNA (or RNA) constructs previously used often do not integrate into the host DNA, or integrate only at low frequencies. Other factors may have also played a part, such as poor entry of the vector into target cells.

The present invention provides transposon-based vectors that can be administered to an animal that overcome the prior art problems relating to low transgene integration frequencies. Two preferred transposon-based vectors of the present invention in which a transposase, gene of interest and other polynucleotide sequences may be introduced are termed pTnMCS (SEQ ID NO:36) and pTnMod (SEQ ID NO:1).

The transposon-based vectors of the present invention produce integration frequencies an order of magnitude greater than has been achieved with previous vectors. More specifically, intratesticular injections performed with a prior art transposon-based vector (described in U.S. Patent No. 5,719,055) resulted in 41% sperm positive roosters whereas intratesticular injections performed with the novel transposon-based vectors of the present invention resulted in 77% sperm positive roosters. Actual frequencies of integration were estimated by either or both comparative strength of the PCR signal from the sperm and histological evaluation of the testes and sperm by quantitative PCR.

The transposon-based vectors of the present invention include a transposase gene operably-linked to a first promoter, and a coding sequence for a desired protein or peptide operably-linked to a second promoter, wherein the coding sequence for the desired protein or peptide and its operably-linked promoter are flanked by transposase insertion sequences recognized by the transposase. The transposon-based vector also includes one or more of the following characteristics: a) one or more modified Kozak sequences comprising ACCATG (SEQ ID NO:13) at the 3' end of the first promoter to enhance expression of the transposase; b) modifications of the codons for the first several N-terminal amino acids of the transposase, wherein the third base of each codon was changed to an A or a T without changing the corresponding amino acid; c)

addition of one or more stop codons to enhance the termination of transposase synthesis; and, d) addition of an effective polyA sequence operably-linked to the transposase to further enhance expression of the transposase gene. Figure 1 shows a schematic representation of several components of the transposon-based vector. The present invention further includes vectors containing more than one gene of interest, wherein a second or subsequent gene of interest is operably-linked to the second promoter or to a different promoter. It is also to be understood that the transposon-based vectors shown in the Figures are representational of the present invention and that the order of the vector elements may be different than that shown in the Figures, that the elements may be present in various orientations, and that the vectors may contain additional elements not shown in the Figures.

Transposases and Insertion Sequences

In a further embodiment of the present invention, the transposase found in the transposase-based vector is an altered target site (ATS) transposase and the insertion sequences are those recognized by the ATS transposase. However, the transposase located in the transposase-based vectors is not limited to a modified ATS transposase and can be derived from any transposase. Transposases known in the prior art include those found in AC7, Tn5SEQ1, Tn916, Tn951, Tn1721, Tn 2410, Tn1681, Tn1, Tn2, Tn3, Tn4, Tn5, Tn6, Tn9, Tn10, Tn30, Tn101, Tn903, Tn501, Tn1000 ($\gamma\delta$), Tn1681, Tn2901, AC transposons, Mp transposons, Spm transposons, En transposons, Dotted transposons, Mu transposons, Ds transposons, dSpm transposons and I transposons. According to the present invention, these transposases and their regulatory sequences are modified for improved functioning as follows: a) the addition one or more modified Kozak sequences comprising ACCATG (SEQ ID NO:13) at the 3' end of the promoter operably-linked to the transposase; b) a change of the codons for the first several amino acids of the transposase, wherein the third base of each codon was changed to an A or a T without changing the corresponding amino acid; c) the addition of one or more stop codons to enhance the termination of transposase synthesis; and/or, d) the addition of an effective polyA sequence operably-linked to the transposase to further enhance expression of the transposase gene.

Although not wanting to be bound by the following statement, it is believed that the modifications of the first several N-terminal codons of the transposase gene increase transcription of the transposase gene, in part, by increasing strand

dissociation. It is preferable that between approximately 1 and 20, more preferably 3 and 15, and most preferably between 4 and 12 of the first N-terminal codons of the transposase are modified such that the third base of each codon is changed to an A or a T without changing the encoded amino acid. In one embodiment, the first ten N-terminal codons of the transposase gene are modified in this manner. It is also preferred that the transposase contain mutations that make it less specific for preferred insertion sites and thus increases the rate of transgene insertion as discussed in U.S. Patent No. 5,719,055.

In some embodiments, the transposon-based vectors are optimized for expression in a particular host by changing the methylation patterns of the vector DNA. For example, prokaryotic methylation may be reduced by using a methylation deficient organism for production of the transposon-based vector. The transposon-based vectors may also be methylated to resemble eukaryotic DNA for expression in a eukaryotic host.

Transposases and insertion sequences from other analogous eukaryotic transposon-based vectors that can also be modified and used are, for example, the *Drosophila* P element derived vectors disclosed in U.S. Patent No. 6,291,243; the *Drosophila* mariner element described in Sherman et al. (1998); or the sleeping beauty transposon. See also Hackett et al. (1999); D. Lampe et al., 1999. Proc. Natl. Acad. Sci. USA, 96:11428-11433; S. Fischer et al., 2001. Proc. Natl. Acad. Sci. USA, 98:6759-6764; L. Zagoraïou et al., 2001. Proc. Natl. Acad. Sci. USA, 98:11474-11478; and D. Berg et al. (Eds.), Mobile DNA, Amer. Soc. Microbiol. (Washington, D.C., 1989). However, it should be noted that bacterial transposon-based elements are preferred, as there is less likelihood that a eukaryotic transposase in the recipient species will recognize prokaryotic insertion sequences bracketing the transgene.

Many transposases recognize different insertion sequences, and therefore, it is to be understood that a transposase-based vector will contain insertion sequences recognized by the particular transposase also found in the transposase-based vector. In a preferred embodiment of the invention, the insertion sequences have been shortened to about 70 base pairs in length as compared to those found in wild-type transposons that typically contain insertion sequences of well over 100 base pairs.

While the examples provided below incorporate a "cut and insert" Tn10 based vector that is destroyed following the insertion event, the present invention also

encompasses the use of a "rolling replication" type transposon-based vector. Use of a rolling replication type transposon allows multiple copies of the transposon/transgene to be made from a single transgene construct and the copies inserted. This type of transposon-based system thereby provides for insertion of multiple copies of a transgene into a single genome. A rolling replication type transposon-based vector may be preferred when the promoter operably-linked to gene of interest is endogenous to the host cell and present in a high copy number or highly expressed. However, use of a rolling replication system may require tight control to limit the insertion events to non-lethal levels. Tn1, Tn2, Tn3, Tn4, Tn5, Tn9, Tn21, Tn501, Tn551, Tn951, Tn1721, Tn2410 and Tn2603 are examples of a rolling replication type transposon, although Tn5 could be both a rolling replication and a cut and insert type transposon.

Stop Codons and PolyA Sequences

In one embodiment, the transposon-based vector contains two stop codons operably-linked to the transposase and/or to the gene of interest. In an alternate embodiment, one stop codon of UAA or UGA is operably linked to the transposase and/or to the gene of interest. As used herein an "effective polyA sequence" refers to either a synthetic or non-synthetic sequence that contains multiple and sequential nucleotides containing an adenine base (an A polynucleotide string) and that increases expression of the gene to which it is operably-linked. A polyA sequence may be operably-linked to any gene in the transposon-based vector including, but not limited to, a transposase gene and a gene of interest. In one embodiment, a polyA sequence comprises the polynucleotide sequence provided in SEQ ID NO:28. A preferred polyA sequence is optimized for use in the host animal or human. In one embodiment, the polyA sequence is optimized for use in a bird, and more specifically, a chicken. The chicken optimized polyA sequence generally contains a minimum of 60 base pairs, and more preferably between approximately 60 and several hundred base pairs, that precede the A polynucleotide string and thereby separate the stop codon from the A polynucleotide string. A chicken optimized polyA sequence may also have a reduced amount of CT repeats as compared to a synthetic polyA sequence. In one embodiment of the present invention, the polyA sequence comprises a conalbumin polyA sequence as provided in SEQ ID NO:33 and as taken from GenBank accession # Y00407, base pairs 10651-11058.

Promoters and Enhancers

The first promoter operably-linked to the transposase gene and the second promoter operably-linked to the gene of interest can be a constitutive promoter or an inducible promoter. Constitutive promoters include, but are not limited to, immediate early cytomegalovirus (CMV) promoter, herpes simplex virus 1 (HSV1) immediate early promoter, SV40 promoter, lysozyme promoter, early and late CMV promoters, early and late HSV promoters, β -actin promoter, tubulin promoter, Rous-Sarcoma virus (RSV) promoter, and heat-shock protein (HSP) promoter. Inducible promoters include tissue-specific promoters, developmentally-regulated promoters and chemically inducible promoters. Examples of tissue-specific promoters include the glucose 6 phosphate (G6P) promoter, vitellogenin promoter, ovalbumin promoter, ovomucoid promoter, conalbumin promoter, ovotransferrin promoter, prolactin promoter, kidney uromodulin promoter, and placental lactogen promoter. In one embodiment, the vitellogenin promoter includes a polynucleotide sequence of SEQ ID NO:17. The G6P promoter sequence may be deduced from a rat G6P gene untranslated upstream region provided in GenBank accession number U57552.1. Examples of developmentally-regulated promoters include the homeobox promoters and several hormone induced promoters. Examples of chemically inducible promoters include reproductive hormone induced promoters and antibiotic inducible promoters such as the tetracycline inducible promoter and the zinc-inducible metallothionein promoter.

Other inducible promoter systems include the Lac operator repressor system inducible by IPTG (isopropyl beta-D-thiogalactoside) (Cronin, A. et al. 2001. *Genes and Development*, v. 15), ecdysone-based inducible systems (Hoppe, U. C. et al. 2000. *Mol. Ther.* 1:159-164); estrogen-based inducible systems (Braselmann, S. et al. 1993. *Proc. Natl. Acad. Sci.* 90:1657-1661); progesterone-based inducible systems using a chimeric regulator, GLVP, which is a hybrid protein consisting of the GAL4 binding domain and the herpes simplex virus transcriptional activation domain, VP16, and a truncated form of the human progesterone receptor that retains the ability to bind ligand and can be turned on by RU486 (Wang, et al. 1994. *Proc. Natl. Acad. Sci.* 91:8180-8184); CID-based inducible systems using chemical inducers of dimerization (CIDs) to regulate gene expression, such as a system wherein rapamycin induces dimerization of the cellular proteins FKBP12 and FRAP (Belshaw, P. J. et al. 1996. *J.*

Chem. Biol. 3:731-738; Fan, L. et al. 1999. Hum. Gene Ther. 10:2273-2285; Shariat, S.F. et al. 2001. Cancer Res. 61:2562-2571; Spencer, D.M. 1996. Curr. Biol. 6:839-847). Chemical substances that activate the chemically inducible promoters can be administered to the animal containing the transgene of interest via any method known to those of skill in the art.

Other examples of cell or tissue-specific and constitutive promoters include but are not limited to smooth-muscle SM22 promoter, including chimeric SM22alpha/telokin promoters (Hoggatt A.M. et al., 2002. Circ Res. 91(12):1151-9); ubiquitin C promoter (Biochim Biophys Acta, 2003. Jan. 3;1625(1):52-63); Hsf2 promoter; murine COMP (cartilage oligomeric matrix protein) promoter; early B cell-specific mb-1 promoter (Sigvardsson M., et al., 2002. Mol. Cell Biol. 22(24):8539-51); prostate specific antigen (PSA) promoter (Yoshimura I. et al., 2002. J. Urol. 168(6):2659-64); exorh promoter and pineal expression-promoting element (Asaoka Y., et al., 2002. Proc. Natl. Acad. Sci. 99(24):15456-61); neural and liver ceramidase gene promoters (Okino N. et al., 2002. Biochem. Biophys. Res. Commun. 299(1):160-6); PSP94 gene promoter/enhancer (Gabril M.Y. et al., 2002. Gene Ther. 9(23):1589-99); promoter of the human FAT/CD36 gene (Kuriki C., et al., 2002. Biol. Pharm. Bull. 25(11):1476-8); VL30 promoter (Staplin W.R. et al., 2002. Blood October 24, 2002); IL-10 promoter (Brenner S., et al., 2002. J. Biol. Chem. December 18, 2002).

Examples of avian promoters include, but are not limited to, promoters controlling expression of egg white proteins, such as ovalbumin, ovotransferrin (conalbumin), ovomucoid, lysozyme, ovomucin, g2 ovoglobulin, g3 ovoglobulin, ovoflavoprotein, ovostatin (ovomacroglobin), cystatin, avidin, thiamine-binding protein, glutamyl aminopeptidase minor glycoprotein 1, minor glycoprotein 2; and promoters controlling expression of egg-yolk proteins, such as vitellogenin, very low-density lipoproteins, low density lipoprotein, cobalamin-binding protein, riboflavin-binding protein, biotin-binding protein (Awade, 1996. Z. Lebensm. Unters. Forsch. 202:1-14). An advantage of using the vitellogenin promoter is that it is active during the egg-laying stage of an animal's life-cycle, which allows for the production of the protein of interest to be temporally connected to the import of the protein of interest into the egg yolk when the protein of interest is equipped with an appropriate targeting sequence.

Liver-specific promoters of the present invention include, but are not limited to, the following promoters, vitellogenin promoter, G6P promoter, cholesterol-7-alpha-hydroxylase (CYP7A) promoter, phenylalanine hydroxylase (PAH) promoter, protein C gene promoter, insulin-like growth factor I (IGF-I) promoter, bilirubin
5 UDP-glucuronosyltransferase promoter, aldolase B promoter, furin promoter, metallothioneine promoter, albumin promoter, and insulin promoter.

Also included in the present invention are promoters that can be used to target expression of a protein of interest into the milk of a milk-producing animal including, but not limited to, β lactoglobulin promoter, whey acidic protein promoter, lactalbumin
10 promoter and casein promoter.

Promoters associated with cells of the immune system may also be used. Acute phase promoters such as interleukin (IL)-1 and IL-2 may be employed. Promoters for heavy and light chain Ig may also be employed. The promoters of the T cell receptor components CD4 and CD8, B cell promoters and the promoters of
15 CR2 (complement receptor type 2) may also be employed. Immune system promoters are preferably used when the desired protein is an antibody protein.

Also included in this invention are modified promoters/enhancers wherein elements of a single promoter are duplicated, modified, or otherwise changed. In one embodiment, steroid hormone-binding domains of the ovalbumin promoter are moved
20 from about -6.5 kb to within approximately the first 1000 base pairs of the gene of interest. Modifying an existing promoter with promoter/enhancer elements not found naturally in the promoter, as well as building an entirely synthetic promoter, or drawing promoter/enhancer elements from various genes together on a non-natural backbone, are all encompassed by the current invention.

Accordingly, it is to be understood that the promoters contained within the transposon-based vectors of the present invention may be entire promoter sequences or fragments of promoter sequences. For example, in one embodiment, the promoter operably linked to a gene of interest is an approximately 900 base pair fragment of a chicken ovalbumin promoter (SEQ ID NO:40). The constitutive and inducible
30 promoters contained within the transposon-based vectors may also be modified by the addition of one or more modified Kozak sequences of ACCATG (SEQ ID NO:13).

As indicated above, the present invention includes transposon-based vectors containing one or more enhancers. These enhancers may or may not be operably-

linked to their native promoter and may be located at any distance from their operably-linked promoter. A promoter operably-linked to an enhancer is referred to herein as an "enhanced promoter." The enhancers contained within the transposon-based vectors are preferably enhancers found in birds, and more preferably, an ovalbumin enhancer, but are not limited to these types of enhancers. In one embodiment, an approximately 675 base pair enhancer element of an ovalbumin promoter is cloned upstream of an ovalbumin promoter with 300 base pairs of spacer DNA separating the enhancer and promoter. In one embodiment, the enhancer used as a part of the present invention comprises base pairs 1-675 of a Chicken Ovalbumin enhancer from GenBank accession #S82527.1. The polynucleotide sequence of this enhancer is provided in SEQ ID NO:37.

Also included in some of the transposon-based vectors of the present invention are cap sites and fragments of cap sites. In one embodiment, approximately 50 base pairs of a 5' untranslated region wherein the capsite resides are added on the 3' end of an enhanced promoter or promoter. An exemplary 5' untranslated region is provided in SEQ ID NO:38. A putative cap-site residing in this 5' untranslated region preferably comprises the polynucleotide sequence provided in SEQ ID NO:39.

In one embodiment of the present invention, the first promoter operably-linked to the transposase gene is a constitutive promoter and the second promoter operably-linked to the gene of interest is a tissue-specific promoter. In this embodiment, use of the first constitutive promoter allows for constitutive activation of the transposase gene and incorporation of the gene of interest into virtually all cell types, including the germline of the recipient animal. Although the gene of interest is incorporated into the germline generally, the gene of interest is only expressed in a tissue-specific manner. It should be noted that cell- or tissue-specific expression as described herein does not require a complete absence of expression in cells or tissues other than the preferred cell or tissue. Instead, "cell-specific" or "tissue-specific" expression refers to a majority of the expression of a particular gene of interest in the preferred cell or tissue, respectively.

When incorporation of the gene of interest into the germline is not preferred, the first promoter operably-linked to the transposase gene can be a tissue-specific promoter. For example, transfection of a transposon-based vector containing a transposase gene operably-linked to a liver-specific promoter such as the G6P

promoter or vitellogenin promoter provides for activation of the transposase gene and incorporation of the gene of interest in the cells of the liver but not into the germline and other cells generally. In this second embodiment, the second promoter operably-linked to the gene of interest can be a constitutive promoter or an inducible promoter.

- 5 In a preferred embodiment, both the first promoter and the second promoter are a G6P promoter. In embodiments wherein tissue-specific expression or incorporation is desired, it is preferred that the transposon-based vector is administered directly to the tissue of interest or to an artery leading to the tissue of interest.

- Accordingly, cell specific promoters may be used to enhance transcription in
10 selected tissues. In birds, for example, promoters that are found in cells of the fallopian tube, such as ovalbumin, conalbumin, ovomucoid and/or lysozyme, are used in the vectors to ensure transcription of the gene of interest in the epithelial cells and tubular gland cells of the fallopian tube, leading to synthesis of the desired protein encoded by the gene and deposition into the egg white. In mammals, promoters
15 specific for the epithelial cells of the alveoli of the mammary gland, such as prolactin, insulin, beta lactoglobulin, whey acidic protein, lactalbumin, casein, and/or placental lactogen, are used in the design of vectors used for transfection of these cells for the production of desired proteins for deposition into the milk. In liver cells, the G6P promoter may be employed to drive transcription of the gene of interest for protein
20 production. Proteins made in the liver of birds may be delivered to the egg yolk.

- In order to achieve higher or more efficient expression of the transposase gene, the promoter and other regulatory sequences operably-linked to the transposase gene may be those derived from the host. These host specific regulatory sequences can be tissue specific as described above or can be of a constitutive nature. For
25 example, an avian actin promoter and its associated polyA sequence can be operably-linked to a transposase in a transposase-based vector for transfection into an avian. Examples of other host specific promoters that could be operably-linked to the transposase include the myosin and DNA or RNA polymerase promoters.

Directing Sequences

- 30 In some embodiments of the present invention, the gene of interest is operably-linked to a directing sequence or a sequence that provides proper conformation to the desired protein encoded by the gene of interest. As used herein, the term "directing sequence" refers to both signal sequences and targeting sequences.

An egg directing sequence includes, but is not limited to, an ovomucoid signal sequence, an ovalbumin signal sequence and a vitellogenin targeting sequence. The term "signal sequence" refers to an amino acid sequence, or the polynucleotide sequence that encodes the amino acid sequence, that directs the protein to which it is linked to the endoplasmic reticulum in a eukaryote, and more preferably the translocational pores in the endoplasmic reticulum, or the plasma membrane in a prokaryote, or mitochondria, such as for the purpose of gene therapy of mitochondrial diseases. Signal and targeting sequences can be used to direct a desired protein into, for example, the milk, when the transposon-based vectors are administered to a milk-producing animal.

Signal sequences can also be used to direct a desired protein into, for example, a secretory pathway for incorporation into the egg yolk or the egg white, when the transposon-based vectors are administered to a bird or other egg-laying animal. One example of such a transposon-based vector is provided in Figure 3 wherein the gene of interest is operably linked to the ovomucoid signal sequence. The present invention also includes a gene of interest operably-linked to a second gene containing a signal sequence. An example of such an embodiment is shown in Figure 2 wherein the gene of interest is operably-linked to the ovalbumin gene that contains an ovalbumin signal sequence. Other signal sequences that can be included in the transposon-based vectors include, but are not limited to the ovotransferrin and lysozyme signal sequences.

As also used herein, the term "targeting sequence" refers to an amino acid sequence, or the polynucleotide sequence encoding the amino acid sequence, which amino acid sequence is recognized by a receptor located on the exterior of a cell. Binding of the receptor to the targeting sequence results in uptake of the protein or peptide operably-linked to the targeting sequence by the cell. One example of a targeting sequence is a vitellogenin targeting sequence that is recognized by a vitellogenin receptor (or the low density lipoprotein receptor) on the exterior of an oocyte. In one embodiment, the vitellogenin targeting sequence includes the polynucleotide sequence of SEQ ID NO:18. In another embodiment, the vitellogenin targeting sequence includes all or part of the vitellogenin gene. Other targeting sequences include VLDL and Apo E, which are also capable of binding the vitellogenin receptor. Since the ApoE protein is not endogenously expressed in birds,

its presence may be used advantageously to identify birds carrying the transposon-based vectors of the present invention.

Genes of Interest Encoding Desired Proteins

A gene of interest selected for stable incorporation is designed to encode any
 5 desired protein or peptide or to regulate any cellular response. In some embodiments, the desired proteins or peptides are deposited in an egg or in milk. It is to be understood that the present invention encompasses transposon-based vectors containing multiple genes of interest. The multiple genes of interest may each be operably-linked to a separate promoter and other regulatory sequence(s) or may all be
 10 operably-linked to the same promoter and other regulatory sequence(s). In one embodiment, multiple gene of interest are linked to a single promoter and other regulatory sequence(s) and each gene of interest is separated by a cleavage site or a portion of a signal sequence.

Protein and peptide hormones are a preferred class of proteins in the present
 15 invention. Such protein and peptide hormones are synthesized throughout the endocrine system and include, but are not limited to, hypothalamic hormones and hypophysiotropic hormones, anterior, intermediate and posterior pituitary hormones, pancreatic islet hormones, hormones made in the gastrointestinal system, renal hormones, thymic hormones, parathyroid hormones, adrenal cortical and medullary
 20 hormones. Specifically, hormones that can be produced using the present invention include, but are not limited to, chorionic gonadotropin, corticotropin, erythropoietin, glucagons, IGF-1, oxytocin, platelet-derived growth factor, calcitonin, follicle-stimulating hormone, leutinizing hormone, thyroid-stimulating hormone, insulin, gonadotropin-releasing hormone and its analogs, vasopressin, octreotide,
 25 somatostatin, prolactin, adrenocorticotrophic hormone, antidiuretic hormone, thyrotropin-releasing hormone (TRH), growth hormone-releasing hormone (GHRH), dopamine, melatonin, thyroxin (T_4), parathyroid hormone (PTH), glucocorticoids such as cortisol, mineralocorticoids such as aldosterone, androgens such as testosterone, adrenaline (epinephrine), noradrenaline (norepinephrine), estrogens such
 30 as estradiol, progesterone, glucagons, calcitriol, calciferol, atrial-natriuretic peptide, gastrin, secretin, cholecystokinin (CCK), neuropeptide Y, ghrelin, PYY₃₋₃₆, angiotensinogen, thrombopoietin, and leptin. By using appropriate polynucleotide sequences, species-specific hormones may be made by transgenic animals.

In one embodiment of the present invention, the gene of interest is a proinsulin gene and the desired molecule is insulin. Proinsulin consists of three parts: a C-peptide and two long strands of amino acids (called the alpha and beta chains) that later become linked together to form the insulin molecule. Figures 2 and 3 are schematics of transposon-based vector constructs containing a proinsulin gene operably-linked to an ovalbumin promoter and ovalbumin protein or an ovomucoid promoter and ovomucoid signal sequence, respectively. In these embodiments, proinsulin is expressed in the oviduct tubular gland cells and then deposited in the egg white. One example of a proinsulin polynucleotide sequence is shown in SEQ ID NO:21, wherein the C-peptide cleavage site spans from Arg at position 31 to Arg at position 65.

Serum proteins including lipoproteins such as high density lipoprotein (HDL), HDL-Milano and low density lipoprotein, albumin, clotting cascade factors, factor VIII, factor IX, fibrinogen, and globulins are also included in the group of desired proteins of the present invention. Immunoglobulins are one class of desired globulin molecules and include but are not limited to IgG, IgM, IgA, IgD, IgE, IgY, lambda chains, kappa chains and fragments thereof; Fc fragments, and Fab fragments. Desired antibodies include, but are not limited to, naturally occurring antibodies, human antibodies, humanized antibodies, and hybrid antibodies. Genes encoding modified versions of naturally occurring antibodies or fragments thereof and genes encoding artificially designed antibodies or fragments thereof may be incorporated into the transposon-based vectors of the present invention. Desired antibodies also include antibodies with the ability to bind specific ligands, for example, antibodies against proteins associated with cancer-related molecules, such as anti-her 2, or anti-CA125. Accordingly, the present invention encompasses a transposon-based vector containing one or more genes encoding a heavy immunoglobulin (Ig) chain and a light Ig chain. Further, more than one gene encoding for more than one antibody may be administered in one or more transposon-based vectors of the present invention. In this manner, an egg may contain more than one type of antibody in the egg white, the egg yolk or both.

In one embodiment, a transposon-based vector contains a heavy Ig chain and a light Ig chain, both operably linked to a promoter. Figures 5 and 6 schematically depict exemplary constructs of this embodiment. More specifically, Figure 5 shows a

construct containing a cecropin pre-pro sequence and a cecropin pro sequence, wherein the pre sequence functions to direct the resultant protein into the endoplasmic reticulum and the pro sequences are cleaved upon secretion of the protein from a cell into which the construct has been transfected. Figure 6 shows a construct containing an enterokinase cleavage site. In this embodiment, it may be required to further remove several additional amino acids from the light chain following cleavage by enterokinase. In another embodiment, the transposon-based vector comprises a heavy Ig chain operably-linked to one promoter and a light Ig chain operably-linked to another promoter. Figure 7 schematically depicts an exemplary construct of this embodiment. The present invention also encompasses a transposon-based vector containing genes encoding portions of a heavy Ig chain and/or portions of a light Ig chain. The present invention further includes a transposon-based vector containing a gene that encodes a fusion protein comprising a heavy and/or light Ig chain, or portions thereof.

Antibodies used as therapeutic reagents include but are not limited to antibodies for use in cancer immunotherapy against specific antigens, or for providing passive immunity to an animal or a human against an infectious disease or a toxic agent. Antibodies used as diagnostic reagents include, but are not limited to antibodies that may be labeled and detected with a detector, for example antibodies with a fluorescent label attached that may be detected following exposure to specific wavelengths. Such labeled antibodies may be primary antibodies directed to a specific antigen, for example, rhodamine-labeled rabbit anti-growth hormone, or may be labeled secondary antibodies, such as fluorescein-labeled goat-anti chicken IgG. Such labeled antibodies are known to one of ordinary skill in the art. Labels useful for attachment to antibodies are also known to one of ordinary skill in the art. Some of these labels are described in the "Handbook of Fluorescent Probes and Research Products", ninth edition, Richard P. Haugland (ed) Molecular Probes, Inc. Eugene, OR), which is incorporated herein in its entirety.

Antibodies produced with using the present invention may be used as laboratory reagents for numerous applications including radioimmunoassay, western blots, dot blots, ELISA, immunoaffinity columns and other procedures requiring antibodies as known to one of ordinary skill in the art. Such antibodies include

primary antibodies, secondary antibodies and tertiary antibodies, which may be labeled or unlabeled.

Antibodies that may be made with the practice of the present invention include, but are not limited to primary antibodies, secondary antibodies, designer antibodies, anti-protein antibodies, anti-peptide antibodies, anti-DNA antibodies, anti-RNA antibodies, anti-hormone antibodies, anti-hypophysiotropic peptides, antibodies against non-natural antigens, anti-anterior pituitary hormone antibodies, anti-posterior pituitary hormone antibodies, anti-venom antibodies, anti-tumor marker antibodies, antibodies directed against epitopes associated with infectious disease, including, anti-viral, anti-bacterial, anti-protozoal, anti-fungal, anti-parasitic, anti-receptor, anti-lipid, anti-phospholipid, anti-growth factor, anti-cytokine, anti-monokine, anti-idiotypic, and anti-accessory (presentation) protein antibodies. Antibodies made with the present invention, as well as light chains or heavy chains, may also be used to inhibit enzyme activity.

Antibodies that may be produced using the present invention include, but are not limited to, antibodies made against the following proteins: Bovine γ -Globulin, Serum; Bovine IgG, Plasma; Chicken γ -Globulin, Serum; Human γ -Globulin, Serum; Human IgA, Plasma; Human IgA₁, Myeloma; Human IgA₂, Myeloma; Human IgA₂, Plasma; Human IgD, Plasma; Human IgE, Myeloma; Human IgG, Plasma; Human IgG, Fab Fragment, Plasma; Human IgG, F(ab')₂ Fragment, Plasma; Human IgG, Fc Fragment, Plasma; Human IgG₁, Myeloma; Human IgG₂, Myeloma; Human IgG₃, Myeloma; Human IgG₄, Myeloma; Human IgM, Myeloma; Human IgM, Plasma; Human Immunoglobulin, Light Chain κ , Urine; Human Immunoglobulin, Light Chains κ and λ , Plasma; Mouse γ -Globulin, Serum; Mouse IgG, Serum; Mouse IgM, Myeloma; Rabbit γ -Globulin, Serum; Rabbit IgG, Plasma; and Rat γ -Globulin, Serum. In one embodiment, the transposon-based vector comprises the coding sequence of light and heavy chains of a murine monoclonal antibody that shows specificity for human seminoprotein (GenBank Accession numbers AY129006 and AY129304 for the light and heavy chains, respectively).

A further non-limiting list of antibodies that recognize other antibodies is as follows: Anti-Chicken IgG, heavy (H) & light (L) Chain Specific (Sheep); Anti-Goat γ -Globulin (Donkey); Anti-Goat IgG, Fc Fragment Specific (Rabbit); Anti-Guinea Pig γ -Globulin (Goat); Anti-Human Ig, Light Chain, Type κ Specific; Anti-Human Ig,

Light Chain, Type λ Specific; Anti-Human IgA, α -Chain Specific (Goat); Anti-Human IgA, Fab Fragment Specific; Anti-Human IgA, Fc Fragment Specific; Anti-Human IgA, Secretory; Anti-Human IgE, ϵ -Chain Specific (Goat); Anti-Human IgE, Fc Fragment Specific; Anti-Human IgG, Fc Fragment Specific (Goat); Anti-Human IgG, γ -Chain Specific (Goat); Anti-Human IgG, Fc Fragment Specific; Anti-Human IgG, Fd Fragment Specific; Anti-Human IgG, H & L Chain Specific (Goat); Anti-Human IgG₁, Fc Fragment Specific; Anti-Human IgG₂, Fc Fragment Specific; Anti-Human IgG₂, Fd Fragment Specific; Anti-Human IgG₃, Hinge Specific; Anti-Human IgG₄, Fc Fragment Specific; Anti-Human IgM, Fc Fragment Specific; Anti-Human IgM, μ -Chain Specific; Anti-Mouse IgE, ϵ -Chain Specific; Anti-Mouse γ -Globulin (Goat); Anti-Mouse IgG, γ -Chain Specific (Goat); Anti-Mouse IgG, γ -Chain Specific (Goat) F(ab')₂ Fragment; Anti-Mouse IgG, H & L Chain Specific (Goat); Anti-Mouse IgM, μ -Chain Specific (Goat); Anti-Mouse IgM, H & L Chain Specific (Goat); Anti-Rabbit γ -Globulin (Goat); Anti-Rabbit IgG, Fc Fragment Specific (Goat); Anti-Rabbit IgG, H & L Chain Specific (Goat); Anti-Rat γ -Globulin (Goat); Anti-Rat IgG, H & L Chain Specific; Anti-Rhesus Monkey γ -Globulin (Goat); and, Anti-Sheep IgG, H & L Chain Specific.

Another non-limiting list of the antibodies that may be produced using the present invention is provided in product catalogs of companies such as Phoenix Pharmaceuticals, Inc. (www.phoenixpeptide.com; 530 Harbor Boulevard, Belmont, CA), Peninsula Labs San Carlos CA, SIGMA, St.Louis, MO www.sigma-aldrich.com, Cappel ICN, Irvine, California, www.icnbiomed.com, and Calbiochem, La Jolla, California, www.calbiochem.com, which are all incorporated herein by reference in their entirety. The polynucleotide sequences encoding these antibodies may be obtained from the scientific literature, from patents, and from databases such as GenBank. Alternatively, one of ordinary skill in the art may design the polynucleotide sequence to be incorporated into the genome by choosing the codons that encode for each amino acid in the desired antibody. Antibodies made by the transgenic animals of the present invention include antibodies that may be used as therapeutic reagents, for example in cancer immunotherapy against specific antigens, as diagnostic reagents and as laboratory reagents for numerous applications including immunoneutralization, radioimmunoassay, western blots, dot blots, ELISA, immunoprecipitation and immunoaffinity columns. Some of these antibodies include,

but are not limited to, antibodies which bind the following ligands: adrenomedulin, amylin, calcitonin, amyloid, calcitonin gene-related peptide, cholecystokinin, gastrin, gastric inhibitory peptide, gastrin releasing peptide, interleukin, interferon, cortistatin, somatostatin, endothelin, sarafotoxin, glucagon, glucagon-like peptide, insulin, atrial natriuretic peptide, BNP, CNP, neurokinin, substance P, leptin, neuropeptide Y, melanin concentrating hormone, melanocyte stimulating hormone, orphanin, endorphin, dynorphin, enkephalin, leumorphin, peptide F, PACAP, PACAP-related peptide, parathyroid hormone, urocortin, corticotrophin releasing hormone, PHM, PHI, vasoactive intestinal polypeptide, secretin, ACTH, angiotensin, angiotensin, bombesin, endostatin, bradykinin, FMRF amide, galanin, gonadotropin releasing hormone (GnRH) associated peptide, GnRH, growth hormone releasing hormone, inhibin, granulocyte-macrophage colony stimulating factor (GM-CSF), motilin, neurotensin, oxytocin, vasopressin, osteocalcin, pancreastatin, pancreatic polypeptide, peptide YY, proopiomelanocortin, transforming growth factor, vascular endothelial growth factor, vesicular monoamine transporter, vesicular acetylcholine transporter, ghrelin, NPW, NPB, C3d, prokineticin, thyroid stimulating hormone, luteinizing hormone, follicle stimulating hormone, prolactin, growth hormone, beta-lipotropin, melatonin, kallikriens, kinins, prostaglandins, erythropoietin, p146 (SEQ ID NO:18 amino acid sequence, SEQ ID NO:19, nucleotide sequence), estrogen, testosterone, corticosteroids, mineralocorticoids, thyroid hormone, thymic hormones, connective tissue proteins, nuclear proteins, actin, avidin, activin, agrin, albumin, and prohormones, propeptides, splice variants, fragments and analogs thereof.

The following is yet another non-limiting of antibodies that can be produced by the methods of present invention: abciximab (ReoPro), abciximab anti-platelet aggregation monoclonal antibody, anti-CD11a (hu1124), anti-CD18 antibody, anti-CD20 antibody, anti-cytomegalovirus (CMV) antibody, anti-digoxin antibody, anti-hepatitis B antibody, anti-HER-2 antibody, anti-idiotypic antibody to GD3 glycolipid, anti-IgE antibody, anti-IL-2R antibody, antimetastatic cancer antibody (mAb 17-1A), anti-rabies antibody, anti-respiratory syncytial virus (RSV) antibody, anti-Rh antibody, anti-TCR, anti-TNF antibody, anti-VEGF antibody and fab fragment thereof, rattlesnake venom antibody, black widow spider venom antibody, coral snake venom antibody, antibody against very late antigen-4 (VLA-4), C225 humanized antibody to EGF receptor, chimeric (human & mouse) antibody against TNF α ,

antibody directed against GPIIb/IIIa receptor on human platelets, gamma globulin, anti-hepatitis B immunoglobulin, human anti-D immunoglobulin, human antibodies against *S. aureus*, human tetanus immunoglobulin, humanized antibody against the epidermal growth receptor-2, humanized antibody against the α subunit of the interleukin-2 receptor, humanized antibody CTLA4IG, humanized antibody to the IL-2 R α -chain, humanized anti-CD40-ligand monoclonal antibody (5c8), humanized mAb against the epidermal growth receptor-2, humanized mAb to rous sarcoma virus, humanized recombinant antibody (IgG1k) against respiratory syncytial virus (RSV), lymphocyte immunoglobulin (anti-thymocyte antibody), lymphocyte immunoglobulin, mAb against factor VII, MDX-210 bi-specific antibody against HER-2, MDX-22, MDX-220 bi-specific antibody against TAG-72 on tumors, MDX-33 antibody to Fc γ R1 receptor, MDX-447 bi-specific antibody against EGF receptor, MDX-447 bispecific humanized antibody to EGF receptor, MDX-RA immunotoxin (ricin A linked) antibody, Medi-507 antibody (humanized form of BTI-322) against CD2 receptor on T-cells, monoclonal antibody LDP-02, muromonab-CD3 (OKT3) antibody, OKT3 ("muromonab-CD3") antibody, PRO 542 antibody, ReoPro ("abciximab") antibody, and TNF-IgG fusion protein.

The antibodies prepared using the methods of the present invention may also be designed to possess specific labels that may be detected through means known to one of ordinary skill in the art. The antibodies may also be designed to possess specific sequences useful for purification through means known to one of ordinary skill in the art. Specialty antibodies designed for binding specific antigens may also be made in transgenic animals using the transposon-based vectors of the present invention.

Production of a monoclonal antibody using the transposon-based vectors of the present invention can be accomplished in a variety of ways. In one embodiment, two vectors may be constructed: one that encodes the light chain, and a second vector that encodes the heavy chain of the monoclonal antibody. These vectors may then be incorporated into the genome of the target animal by methods disclosed herein. In an alternative embodiment, the sequences encoding light and heavy chains of a monoclonal antibody may be included on a single DNA construct. For example, the coding sequence of light and heavy chains of a murine monoclonal antibody that show specificity for human seminoprotein can be expressed using transposon-based

constructs of the present invention (GenBank Accession numbers AY129006 and AY129304 for the light and heavy chains, respectively).

Further included in the present invention are proteins and peptides synthesized by the immune system including those synthesized by the thymus, lymph nodes, spleen, and the gastrointestinal associated lymph tissues (GALT) system. The immune system proteins and peptides proteins that can be made in transgenic animals using the transposon-based vectors of the present invention include, but are not limited to, alpha-interferon, beta-interferon, gamma-interferon, alpha-interferon A, alpha-interferon 1, G-CSF, GM-CSF, interleukin-1 (IL-1), IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, TNF- α , and TNF- β . Other cytokines included in the present invention include cardiotrophin, stromal cell derived factor, macrophage derived chemokine (MDC), melanoma growth stimulatory activity (MGSA), macrophage inflammatory proteins 1 alpha (MIP-1 alpha), 2, 3 alpha, 3 beta, 4 and 5.

Lytic peptides such as p146 are also included in the desired molecules of the present invention. In one embodiment, the p146 peptide comprises an amino acid sequence of SEQ ID NO:19. The present invention also encompasses a transposon-based vector comprising a p146 nucleic acid comprising a polynucleotide sequence of SEQ ID NO:20.

Enzymes are another class of proteins that may be made through the use of the transposon-based vectors of the present invention. Such enzymes include but are not limited to adenosine deaminase, alpha-galactosidase, cellulase, collagenase, dnaseI, hyaluronidase, lactase, L-asparaginase, pancreatin, papain, streptokinase B, subtilisin, superoxide dismutase, thrombin, trypsin, urokinase, fibrinolysin, glucocerebrosidase and plasminogen activator. In some embodiments wherein the enzyme could have deleterious effects, additional amino acids and a protease cleavage site are added to the carboxy end of the enzyme of interest in order to prevent expression of a functional enzyme. Subsequent digestion of the enzyme with a protease results in activation of the enzyme.

Extracellular matrix proteins are one class of desired proteins that may be made through the use of the present invention. Examples include but are not limited to collagen, fibrin, elastin, laminin, and fibronectin and subtypes thereof. Intracellular

proteins and structural proteins are other classes of desired proteins in the present invention.

Growth factors are another desired class of proteins that may be made through the use of the present invention and include, but are not limited to, transforming growth factor- α ("TGF- α "), transforming growth factor- β (TGF- β), platelet-derived growth factors (PDGF), fibroblast growth factors (FGF), including FGF acidic isoforms 1 and 2, FGF basic form 2 and FGF 4, 8, 9 and 10, nerve growth factors (NGF) including NGF 2.5s, NGF 7.0s and beta NGF and neurotrophins, brain derived neurotrophic factor, cartilage derived factor, growth factors for stimulation of the production of red blood cells, growth factors for stimulation of the production of white blood cells, bone growth factors (BGF), basic fibroblast growth factor, vascular endothelial growth factor (VEGF), granulocyte colony stimulating factor (G-CSF), insulin like growth factor (IGF) I and II, hepatocyte growth factor, glial neurotrophic growth factor (GDNF), stem cell factor (SCF), keratinocyte growth factor (KGF), transforming growth factors (TGF), including TGFs alpha, beta, beta1, beta2, beta3, skeletal growth factor, bone matrix derived growth factors, bone derived growth factors, erythropoietin (EPO) and mixtures thereof.

Another desired class of proteins that may be made may be made through the use of the present invention include but are not limited to leptin, leukemia inhibitory factor (LIF), tumor necrosis factor alpha and beta, ENBREL, angiostatin, endostatin, thrombospondin, osteogenic protein-1, bone morphogenetic proteins 2 and 7, osteonectin, somatomedin-like peptide, and osteocalcin.

A non-limiting list of the peptides and proteins that may be made may be made through the use of the present invention is provided in product catalogs of companies such as Phoenix Pharmaceuticals, Inc. (www.phoenixpeptide.com; 530 Harbor Boulevard • Belmont, CA), Peninsula Labs San Carlos CA, SIGMA, St.Louis, MO www.sigma-aldrich.com, Cappel ICN, Irvine, California, www.icnbiomed.com, and Calbiochem, La Jolla, California, www.calbiochem.com. The polynucleotide sequences encoding these proteins and peptides of interest may be obtained from the scientific literature, from patents, and from databases such as GenBank. Alternatively, one of ordinary skill in the art may design the polynucleotide sequence to be incorporated into the genome by choosing the codons that encode for each amino acid in the desired protein or peptide.

Some of these desired proteins or peptides that may be made through the use of the present invention include but are not limited to the following: adrenomedulin, amylin, calcitonin, amyloid, calcitonin gene-related peptide, cholecystokinin, gastrin, gastric inhibitory peptide, gastrin releasing peptide, interleukin, interferon, cortistatin, 5 somatostatin, endothelin, sarafotoxin, glucagon, glucagon-like peptide, insulin, atrial natriuretic peptide, BNP, CNP, neurokinin, substance P, leptin, neuropeptide Y, melanin concentrating hormone, melanocyte stimulating hormone, orphanin, endorphin, dynorphin, enkephalin, leuromorphin, peptide F, PACAP, PACAP-related peptide, parathyroid hormone, urocortin, corticotrophin releasing hormone, PHM, 10 PHI, vasoactive intestinal polypeptide, secretin, ACTH, angiotensin, angiotensin, bombesin, endostatin, bradykinin, FMRF amide, galanin, gonadotropin releasing hormone (GnRH) associated peptide, GnRH, growth hormone releasing hormone, inhibin, granulocyte-macrophage colony stimulating factor (GM-CSF), motilin, neurotensin, oxytocin, vasopressin, osteocalcin, pancreastatin, pancreatic polypeptide, 15 peptide YY, proopiomelanocortin, transforming growth factor, vascular endothelial growth factor, vesicular monoamine transporter, vesicular acetylcholine transporter, ghrelin, NPW, NPB, C3d, prokineticin, thyroid stimulating hormone, luteinizing hormone, follicle stimulating hormone, prolactin, growth hormone, beta-lipotropin, melatonin, kallikriens, kinins, prostaglandins, erythropoietin, p146 (SEQ ID NO:19, 20 amino acid sequence, SEQ ID NO:20, nucleotide sequence), thymic hormones, connective tissue proteins, nuclear proteins, actin, avidin, activin, agrin, albumin, and prohormones, propeptides, splice variants, fragments and analogs thereof.

Other desired proteins that may be made by the transgenic animals of the present invention include bacitracin, polymixin b, vancomycin, cyclosporine, anti- 25 RSV antibody, alpha-1 antitrypsin (AAT), anti-cytomegalovirus antibody, anti-hepatitis antibody, anti-inhibitor coagulant complex, anti-rabies antibody, anti-Rh(D) antibody, adenosine deaminase, anti-digoxin antibody, antivenin crotalidae (rattlesnake venom antibody), antivenin latrodectus (black widow spider venom antibody), antivenin micrurus (coral snake venom antibody), aprotinin, corticotropin 30 (ACTH), diphtheria antitoxin, lymphocyte immune globulin (anti-thymocyte antibody), protamine, thyrotropin, capreomycin, α -galactosidase, gramicidin, streptokinase, tetanus toxoid, tyrothricin, IGF-1, proteins of varicella vaccine, anti-TNF antibody, anti-IL-2r antibody, anti-HER-2 antibody, OKT3 ("muromonab-

CD3") antibody, TNF-IgG fusion protein, ReoPro ("abciximab") antibody, ACTH fragment 1-24, desmopressin, gonadotropin-releasing hormone, histrelin, leuprolide, lypressin, nafarelin, peptide that binds GPIIb/GPIIIa on platelets (integrilin), goserelin, capreomycin, colistin, anti-respiratory syncytial virus, lymphocyte immune globulin (Thymoglovin, Atgam), panorex, alpha-antitrypsin, botulinin, lung surfactant protein, tumor necrosis receptor-IgG fusion protein (enbrel), gonadorelin, proteins of influenza vaccine, proteins of rotavirus vaccine, proteins of haemophilus b conjugate vaccine, proteins of poliovirus vaccine, proteins of pneumococcal conjugate vaccine, proteins of meningococcal C vaccine, proteins of influenza vaccine, megakaryocyte growth and development factor (MGDF), neuroimmunophilin ligand-A (NIL-A), brain-derived neurotrophic factor (BDNF), glial cell line-derived neurotrophic factor (GDNF), leptin (native), leptin B, leptin C, IL-1RA (interleukin-1RA), R-568, novel erythropoiesis-stimulating protein (NESP), humanized mAb to rous sarcoma virus (MEDI-493), glutamyl-tryptophan dipeptide IM862, LFA-3TIP immunosuppressive, humanized anti-CD40-ligand monoclonal antibody (5c8), gelsolin enzyme, tissue factor pathway inhibitor (TFPI), proteins of meningitis B vaccine, antimetastatic cancer antibody (mAb 17-1A), chimeric (human & mouse) mAb against TNF α , mAb against factor VII, relaxin, capreomycin, glycopeptide (LY333328), recombinant human activated protein C (rhAPC), humanized mAb against the epidermal growth receptor-2, altepase, anti-CD20 antigen, C2B8 antibody, insulin-like growth factor-1, atrial natriuretic peptide (anaritide), tenectapase, anti-CD11a antibody (hu 1124), anti-CD18 antibody, mAb LDP-02, anti-VEGF antibody, fab fragment of anti-VEGF Ab, APO2 ligand (tumor necrosis factor-related apoptosis-inducing ligand), rTGF- β (transforming growth factor- β), alpha-antitrypsin, ananain (a pineapple enzyme), humanized mAb CTLA4IG, PRO 542 (mAb), D2E7 (mAb), calf intestine alkaline phosphatase, α -L-iduronidase, α -L-galactosidase (humanglutamic acid decarboxylase, acid sphingomyelinase, bone morphogenetic protein-2 (rhBMP-2), proteins of HIV vaccine, T cell receptor (TCR) peptide vaccine, TCR peptides, V beta 3 and V beta 13.1. (IR502), (IR501), BI 1050/1272 mAb against very late antigen-4 (VLA-4), C225 humanized mAb to EGF receptor, anti-idiotypic antibody to GD3 glycolipid, antibacterial peptide against *H. pylori*, MDX-447 bispecific humanized mAb to EGF receptor, anti-cytomegalovirus (CMV), Medi-491 B19 parvovirus vaccine, humanized recombinant mAb (IgG1k) against respiratory syncytial virus (RSV), urinary tract

infection vaccine (against "pili" on *Escherichia coli* strains), proteins of lyme disease vaccine against *B. burgdorferi* protein (DbpA), proteins of Medi-501 human papilloma virus-11 vaccine (HPV), *Streptococcus pneumoniae* vaccine, Medi-507 mAb (humanized form of BTI-322) against CD2 receptor on T-cells, MDX-33 mAb to FcγR1 receptor, MDX-RA immunotoxin (ricin A linked) mAb, MDX-210 bi-specific mAb against HER-2, MDX-447 bi-specific mAb against EGF receptor, MDX-22, MDX-220 bi-specific mAb against TAG-72 on tumors, colony-stimulating factor (CSF) (molgramostim), humanized mAb to the IL-2 R α-chain (basiliximab), mAb to IgE (IGE 025A), myelin basic protein-altered peptide (MSP771A), humanized mAb against the epidermal growth receptor-2, humanized mAb against the α subunit of the interleukin-2 receptor, low molecular weight heparin, anti-hemophilic factor, and bactericidal/permeability-increasing protein (r-BPI).

The peptides and proteins made using the present invention may be labeled using labels and techniques known to one of ordinary skill in the art. Some of these labels are described in the "Handbook of Fluorescent Probes and Research Products", ninth edition, Richard P. Haugland (ed) Molecular Probes, Inc. Eugene, OR), which is incorporated herein in its entirety. Some of these labels may be genetically engineered into the polynucleotide sequence for the expression of the selected protein or peptide. The peptides and proteins may also have label-incorporation "handles" incorporated to allow labeling of an otherwise difficult or impossible to label protein.

It is to be understood that the various classes of desired peptides and proteins, as well as specific peptides and proteins described in this section may be modified as described below by inserting selected codons for desired amino acid substitutions into the gene incorporated into the transgenic animal.

The present invention may also be used to produce desired molecules other than proteins and peptides including, but not limited to, lipoproteins such as high density lipoprotein (HDL), HDL-Milano, and low density lipoprotein, lipids, carbohydrates, siRNA and ribozymes. In these embodiments, a gene of interest encodes a nucleic acid molecule or a protein that directs production of the desired molecule.

The present invention further encompasses the use of inhibitory molecules to inhibit endogenous (i.e., non-vector) protein production. These inhibitory molecules include antisense nucleic acids, siRNA and inhibitory proteins. In one embodiment, a

transposon-based vector containing an ovalbumin DNA sequence, that upon transcription forms a double stranded RNA molecule, is transfected into an animal such as a bird and the bird's production of endogenous ovalbumin protein is reduced by the interference RNA mechanism (RNAi). Additionally, inducible knockouts or
 5 knockdowns of the endogenous protein may be created to achieve a reduction or inhibition of endogenous protein production.

Modified Desired Proteins and Peptides

"Proteins", "peptides", "polypeptides" and "oligopeptides" are chains of amino acids (typically L-amino acids) whose alpha carbons are linked through peptide bonds
 10 formed by a condensation reaction between the carboxyl group of the alpha carbon of one amino acid and the amino group of the alpha carbon of another amino acid. The terminal amino acid at one end of the chain (i.e., the amino terminal) has a free amino group, while the terminal amino acid at the other end of the chain (i.e., the carboxy terminal) has a free carboxyl group. As such, the term "amino terminus" (abbreviated
 15 N-terminus) refers to the free alpha-amino group on the amino acid at the amino terminal of the protein, or to the alpha-amino group (imino group when participating in a peptide bond) of an amino acid at any other location within the protein. Similarly, the term "carboxy terminus" (abbreviated C-terminus) refers to the free carboxyl group on the amino acid at the carboxy terminus of a protein, or to the
 20 carboxyl group of an amino acid at any other location within the protein.

Typically, the amino acids making up a protein are numbered in order, starting at the amino terminal and increasing in the direction toward the carboxy terminal of the protein. Thus, when one amino acid is said to "follow" another, that amino acid is positioned closer to the carboxy terminal of the protein than the preceding amino acid.

25 The term "residue" is used herein to refer to an amino acid (D or L) or an amino acid mimetic that is incorporated into a protein by an amide bond. As such, the amino acid may be a naturally occurring amino acid or, unless otherwise limited, may encompass known analogs of natural amino acids that function in a manner similar to the naturally occurring amino acids (i.e., amino acid mimetics). Moreover, an amide
 30 bond mimetic includes peptide backbone modifications well known to those skilled in the art.

Furthermore, one of skill will recognize that, as mentioned above, individual substitutions, deletions or additions which alter, add or delete a single amino acid or a

small percentage of amino acids (typically less than about 5%, more typically less than about 1%) in an encoded sequence are conservatively modified variations where the alterations result in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. The following six groups each contain amino acids that are conservative substitutions for one another:

- 1) Alanine (A), Serine (S), Threonine (T);
- 2) Aspartic acid (D), Glutamic acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 4) Arginine (R), Lysine (K);
- 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and
- 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

A conservative substitution is a substitution in which the substituting amino acid (naturally occurring or modified) is structurally related to the amino acid being substituted, i.e., has about the same size and electronic properties as the amino acid being substituted. Thus, the substituting amino acid would have the same or a similar functional group in the side chain as the original amino acid. A "conservative substitution" also refers to utilizing a substituting amino acid which is identical to the amino acid being substituted except that a functional group in the side chain is protected with a suitable protecting group.

Suitable protecting groups are described in Green and Wuts, "Protecting Groups in Organic Synthesis", John Wiley and Sons, Chapters 5 and 7, 1991, the teachings of which are incorporated herein by reference. Preferred protecting groups are those which facilitate transport of the peptide through membranes, for example, by reducing the hydrophilicity and increasing the lipophilicity of the peptide, and which can be cleaved, either by hydrolysis or enzymatically (Ditter et al., 1968. J. Pharm. Sci. 57:783; Ditter et al., 1968. J. Pharm. Sci. 57:828; Ditter et al., 1969. J. Pharm. Sci. 58:557; King et al., 1987. Biochemistry 26:2294; Lindberg et al., 1989. Drug Metabolism and Disposition 17:311; Tunek et al., 1988. Biochem. Pharm. 37:3867; Anderson et al., 1985 Arch. Biochem. Biophys. 239:538; and Singhal et al., 1987. FASEB J. 1:220). Suitable hydroxyl protecting groups include ester, carbonate and carbamate protecting groups. Suitable amine protecting groups include acyl groups and alkoxy or aryloxy carbonyl groups, as described above for N-terminal protecting

groups. Suitable carboxylic acid protecting groups include aliphatic, benzyl and aryl esters, as described below for C-terminal protecting groups. In one embodiment, the carboxylic acid group in the side chain of one or more glutamic acid or aspartic acid residues in a peptide of the present invention is protected, preferably as a methyl, ethyl, benzyl or substituted benzyl ester, more preferably as a benzyl ester.

Provided below are groups of naturally occurring and modified amino acids in which each amino acid in a group has similar electronic and steric properties. Thus, a conservative substitution can be made by substituting an amino acid with another amino acid from the same group. It is to be understood that these groups are non-limiting, i.e. that there are additional modified amino acids which could be included in each group.

Group I includes leucine, isoleucine, valine, methionine and modified amino acids having the following side chains: ethyl, n-propyl n-butyl. Preferably, Group I includes leucine, isoleucine, valine and methionine.

Group II includes glycine, alanine, valine and a modified amino acid having an ethyl side chain. Preferably, Group II includes glycine and alanine.

Group III includes phenylalanine, phenylglycine, tyrosine, tryptophan, cyclohexylmethyl glycine, and modified amino residues having substituted benzyl or phenyl side chains. Preferred substituents include one or more of the following: halogen, methyl, ethyl, nitro, —NH_2 , methoxy, ethoxy and —CN . Preferably, Group III includes phenylalanine, tyrosine and tryptophan.

Group IV includes glutamic acid, aspartic acid, a substituted or unsubstituted aliphatic, aromatic or benzylic ester of glutamic or aspartic acid (e.g., methyl, ethyl, n-propyl iso-propyl, cyclohexyl, benzyl or substituted benzyl), glutamine, asparagine, —CO—NH— alkylated glutamine or asparagines (e.g., methyl, ethyl, n-propyl and iso-propyl) and modified amino acids having the side chain $\text{—(CH}_2\text{)}_3\text{—COOH}$, an ester thereof (substituted or unsubstituted aliphatic, aromatic or benzylic ester), an amide thereof and a substituted or unsubstituted N-alkylated amide thereof. Preferably, Group IV includes glutamic acid, aspartic acid, methyl aspartate, ethyl aspartate, benzyl aspartate and methyl glutamate, ethyl glutamate and benzyl glutamate, glutamine and asparagine.

Group V includes histidine, lysine, ornithine, arginine, N-nitroarginine, β -cycloarginine, γ -hydroxyarginine, N-amidinocitruline and 2-amino-4-guanidinobutanoic acid, homologs of lysine, homologs of arginine and homologs of ornithine. Preferably, Group V includes histidine, lysine, arginine and ornithine. A homolog of an amino acid includes from 1 to about 3 additional or subtracted methylene units in the side chain.

Group VI includes serine, threonine, cysteine and modified amino acids having C1-C5 straight or branched alkyl side chains substituted with $-\text{OH}$ or $-\text{SH}$, for example, $-\text{CH}_2\text{CH}_2\text{OH}$, $-\text{CH}_2\text{CH}_2\text{CH}_2\text{OH}$ or $-\text{CH}_2\text{CH}_2\text{OHCH}_3$. Preferably, Group VI includes serine, cysteine or threonine.

In another aspect, suitable substitutions for amino acid residues include "severe" substitutions. A "severe substitution" is a substitution in which the substituting amino acid (naturally occurring or modified) has significantly different size and/or electronic properties compared with the amino acid being substituted. Thus, the side chain of the substituting amino acid can be significantly larger (or smaller) than the side chain of the amino acid being substituted and/or can have functional groups with significantly different electronic properties than the amino acid being substituted. Examples of severe substitutions of this type include the substitution of phenylalanine or cyclohexylmethyl glycine for alanine, isoleucine for glycine, a D amino acid for the corresponding L amino acid, or $-\text{NH}-\text{CH}[(\text{CH}_2)_5-\text{COOH}]-\text{CO}-$ for aspartic acid. Alternatively, a functional group may be added to the side chain, deleted from the side chain or exchanged with another functional group. Examples of severe substitutions of this type include adding of valine, leucine or isoleucine, exchanging the carboxylic acid in the side chain of aspartic acid or glutamic acid with an amine, or deleting the amine group in the side chain of lysine or ornithine. In yet another alternative, the side chain of the substituting amino acid can have significantly different steric and electronic properties than the functional group of the amino acid being substituted. Examples of such modifications include tryptophan for glycine, lysine for aspartic acid and $(\text{CH}_2)_4\text{COOH}$ for the side chain of serine. These examples are not meant to be limiting.

In another embodiment, for example in the synthesis of a peptide 26 amino acids in length, the individual amino acids may be substituted according in the following manner:

- AA₁ is serine, glycine, alanine, cysteine or threonine;
- 5 AA₂ is alanine, threonine, glycine, cysteine or serine;
- AA₃ is valine, arginine, leucine, isoleucine, methionine, ornithine, lysine, N-nitroarginine, β-cycloarginine, γ-hydroxyarginine, N-amidinocitruline or 2-amino-4-guanidinobutanoic acid;
- AA₄ is proline, leucine, valine, isoleucine or methionine;
- 10 AA₅ is tryptophan, alanine, phenylalanine, tyrosine or glycine;
- AA₆ is serine, glycine, alanine, cysteine or threonine;
- AA₇ is proline, leucine, valine, isoleucine or methionine;
- AA₈ is alanine, threonine, glycine, cysteine or serine;
- AA₉ is alanine, threonine, glycine, cysteine or serine;
- 15 AA₁₀ is leucine, isoleucine, methionine or valine;
- AA₁₁ is serine, glycine, alanine, cysteine or threonine;
- AA₁₂ is leucine, isoleucine, methionine or valine;
- AA₁₃ is leucine, isoleucine, methionine or valine;
- AA₁₄ is glutamine, glutamic acid, aspartic acid, asparagine, or a substituted or
- 20 unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;
- AA₁₅ is arginine, N-nitroarginine, β-cycloarginine, γ-hydroxy-arginine, N-amidinocitruline or 2-amino-4-guanidino-butanoic acid
- AA₁₆ is proline, leucine, valine, isoleucine or methionine;
- AA₁₇ is serine, glycine, alanine, cysteine or threonine;
- 25 AA₁₈ is glutamic acid, aspartic acid, asparagine, glutamine or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;
- AA₁₉ is aspartic acid, asparagine, glutamic acid, glutamine, leucine, valine, isoleucine, methionine or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;
- 30 AA₂₀ is valine, arginine, leucine, isoleucine, methionine, ornithine, lysine, N-nitroarginine, β-cycloarginine, γ-hydroxyarginine, N-amidinocitruline or 2-amino-4-guanidinobutanoic acid;
- AA₂₁ is alanine, threonine, glycine, cysteine or serine;

AA₂₂ is alanine, threonine, glycine, cysteine or serine;

AA₂₃ is histidine, serine, threonine, cysteine, lysine or ornithine;

AA₂₄ is threonine, aspartic acid, serine, glutamic acid or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid;

5 AA₂₅ is asparagine, aspartic acid, glutamic acid, glutamine, leucine, valine, isoleucine, methionine or a substituted or unsubstituted aliphatic or aryl ester of glutamic acid or aspartic acid; and

AA₂₆ is cysteine, histidine, serine, threonine, lysine or ornithine.

10 It is to be understood that these amino acid substitutions may be made for longer or shorter peptides than the 26 mer in the preceding example above, and for proteins.

In one embodiment of the present invention, codons for the first several N-terminal amino acids of the transposase are modified such that the third base of each codon is changed to an A or a T without changing the corresponding amino acid. It is
15 preferable that between approximately 1 and 20, more preferably 3 and 15, and most preferably between 4 and 12 of the first N-terminal codons of the gene of interest are modified such that the third base of each codon is changed to an A or a T without changing the corresponding amino acid. In one embodiment, the first ten N-terminal codons of the gene of interest are modified in this manner.

20 When several desired proteins, protein fragments or peptides are encoded in the gene of interest to be incorporated into the genome, one of skill in the art will appreciate that the proteins, protein fragments or peptides may be separated by a spacer molecule such as, for example, a peptide, consisting of one or more amino acids. Generally, the spacer will have no specific biological activity other than to join
25 the desired proteins, protein fragments or peptides together, or to preserve some minimum distance or other spatial relationship between them. However, the constituent amino acids of the spacer may be selected to influence some property of the molecule such as the folding, net charge, or hydrophobicity. The spacer may also be contained within a nucleotide sequence with a purification handle or be flanked by
30 proteolytic cleavage sites.

Such polypeptide spacers may have from about 5 to about 40 amino acid residues. The spacers in a polypeptide are independently chosen, but are preferably all the same. The spacers should allow for flexibility of movement in space and are

- therefore typically rich in small amino acids, for example, glycine, serine, proline or alanine. Preferably, peptide spacers contain at least 60%, more preferably at least 80% glycine or alanine. In addition, peptide spacers generally have little or no biological and antigenic activity. Preferred spacers are (Gly-Pro-Gly-Gly)_x (SEQ ID NO:5) and (Gly₄-Ser)_y, wherein x is an integer from about 3 to about 9 and y is an integer from about 1 to about 8. Specific examples of suitable spacers include (Gly-Pro-Gly-Gly)₃
- SEQ ID NO:6 Gly Pro Gly Gly Gly Pro Gly Gly Gly Pro Gly Gly
(Gly₄-Ser)₃
- 10 SEQ ID NO:7 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
or (Gly₄-Ser)₄
- SEQ ID NO:8 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
Gly Gly Gly Gly Ser.

- Nucleotide sequences encoding for the production of residues which may be useful in purification of the expressed recombinant protein may also be built into the vector. Such sequences are known in the art and include the glutathione binding domain from glutathione S-transferase, polylysine, hexa-histidine or other cationic amino acids, thioredoxin, hemagglutinin antigen and maltose binding protein.

- Additionally, nucleotide sequences may be inserted into the gene of interest to be incorporated so that the protein or peptide can also include from one to about six amino acids that create signals for proteolytic cleavage. In this manner, if a gene is designed to make one or more peptides or proteins of interest in the transgenic animal, specific nucleotide sequences encoding for amino acids recognized by enzymes may be incorporated into the gene to facilitate cleavage of the large protein or peptide sequence into desired peptides or proteins or both. For example, nucleotides encoding a proteolytic cleavage site can be introduced into the gene of interest so that a signal sequence can be cleaved from a protein or peptide encoded by the gene of interest. Nucleotide sequences encoding other amino acid sequences which display pH sensitivity or chemical sensitivity may also be added to the vector to facilitate separation of the signal sequence from the peptide or protein of interest.

In one embodiment of the present invention, a TAG sequence is linked to the gene of interest. The TAG sequence serves three purposes: 1) it allows free rotation of the peptide or protein to be isolated so there is no interference from the native

protein or signal sequence, i.e. vitellogenin, 2) it provides a "purification handle" to isolate the protein using column purification, and 3) it includes a cleavage site to remove the desired protein from the signal and purification sequences. Accordingly, as used herein, a TAG sequence includes a spacer sequence, a purification handle and a cleavage site. The spacer sequences in the TAG proteins contain one or more repeats shown in SEQ ID NO:25. A preferred spacer sequence comprises the sequence provided in SEQ ID NO:26. One example of a purification handle is the gp41 hairpin loop from HIV I. Exemplary gp41 polynucleotide and polypeptide sequences are provided in SEQ ID NO:24 and SEQ ID NO:23, respectively. However, it should be understood that any antigenic region may be used as a purification handle, including any antigenic region of gp41. Preferred purification handles are those that elicit highly specific antibodies. Additionally, the cleavage site can be any protein cleavage site known to one of ordinary skill in the art and includes an enterokinase cleavage site comprising the Asp Asp Asp Asp Lys sequence (SEQ ID NO:9) and a furin cleavage site. Constructs containing a TAG sequence are shown in Figures 2 and 3. In one embodiment of the present invention, the TAG sequence comprises a polynucleotide sequence of SEQ ID NO:22.

Methods of Administering Transposon-Based Vectors

In addition to the transposon-based vectors described above, the present invention also includes methods of administering the transposon-based vectors to an animal, methods of producing a transgenic animal wherein a gene of interest is incorporated into the germline of the animal and methods of producing a transgenic animal wherein a gene of interest is incorporated into cells other than the germline cells of the animal. The transposon-based vectors of the present invention may be administered to an animal via any method known to those of skill in the art, including, but not limited to, intraembryonic, intratesticular, intraoviduct, intraperitoneal, intraarterial, intravenous, topical, oral, nasal, and pronuclear injection methods of administration, or any combination thereof. The transposon-based vectors may also be administered within the lumen of an organ, into an organ, into a body cavity, into the cerebrospinal fluid, through the urinary system or through any route to reach the desired cells.

The transposon-based vectors may be delivered through the vascular system to be distributed to the cells supplied by that vessel. For example, the compositions may

be placed in the artery supplying the ovary or supplying the fallopian tube to transfect cells in those tissues. In this manner, follicles could be transfected to create a germline transgenic animal. Alternatively, supplying the compositions through the artery leading to the oviduct would preferably transfect the tubular gland and epithelial cells. Such transfected cells could manufacture a desired protein or peptide for deposition in the egg white. Administration of the compositions through the portal vein would target uptake and transformation of hepatic cells. Administration through the urethra and into the bladder would target the transitional epithelium of the bladder. Administration through the vagina and cervix would target the lining of the uterus. Administration through the internal mammary artery would transfect secretory cells of the lactating mammary gland to perform a desired function, such as to synthesize and secrete a desired protein or peptide into the milk.

In a preferred embodiment, the animal is an egg-laying animal, and more preferably, an avian. In one embodiment, between approximately 1 and 50 μg , preferably between 1 and 20 μg , and more preferably between 5 and 10 μg of transposon-based vector DNA is administered to the oviduct of a bird. Optimal ranges depending upon the type of bird and the bird's stage of sexual maturity. Intraoviduct administration of the transposon-based vectors of the present invention result in a PCR positive signal in the oviduct tissue, whereas intravascular administration results in a PCR positive signal in the liver. In other embodiments, the transposon-based vector is administered to an artery that supplies the oviduct or the liver. These methods of administration may also be combined with any methods for facilitating transfection, including without limitation, electroporation, gene guns, injection of naked DNA, and use of dimethyl sulfoxide (DMSO).

The present invention includes a method of intraembryonic administration of a transposon-based vector to an avian embryo comprising the following steps: 1) incubating an egg on its side at room temperature for two hours to allow the embryo contained therein to move to top dead center (TDC); 2) drilling a hole through the shell without penetrating the underlying shell membrane; 3) injecting the embryo with the transposon-based vector in solution; 4) sealing the hole in the egg; and 5) placing the egg in an incubator for hatching. Administration of the transposon-based vector can occur anytime between immediately after egg lay (when the embryo is at Stage X) and hatching. Preferably, the transposon-based vector is administered between 1 and

7 days after egg lay, more preferably between 1 and 2 days after egg lay. The transposon-based vectors may be introduced into the embryo in amounts ranging from about 5.0 μg to 10 μg , preferably 1.0 μg to 100 μg . Additionally, the transposon-based vector solution volume may be between approximately 1 μl to 75 μl in quail and between approximately 1 μl to 500 μl in chicken.

The present invention also includes a method of intratesticular administration of a transposon-based vector including injecting a bird with a composition comprising the transposon-based vector, an appropriate carrier and an appropriate transfection reagent. In one embodiment, the bird is injected before sexual maturity, preferably between approximately 4-14 weeks, more preferably between approximately 6-14 weeks and most preferably between 8-12 weeks old. In another embodiment, a mature bird is injected with a transposon-based vector an appropriate carrier and an appropriate transfection reagent. The mature bird may be any type of bird, but in one example the mature bird is a quail.

A bird is preferably injected prior to the development of the blood-testis barrier, which thereby facilitates entry of the transposon-based vector into the seminiferous tubules and transfection of the spermatogonia or other germline cells. At and between the ages of 4, 6, 8, 10, 12, and 14 weeks, it is believed that the testes of chickens are likely to be most receptive to transfection. In this age range, the blood/testis barrier has not yet formed, and there is a relatively high number of spermatogonia relative to the numbers of other cell types, e.g., spermatids, etc. See J. Kumaran et al., 1949. Poultry Sci., 29:511-520. See also E. Oakberg, 1956. Am. J. Anatomy, 99:507-515; and P. Kluin et al., 1984. Anat. Embryol., 169:73-78.

The transposon-based vectors may be introduced into a testis in an amount ranging from about 0.1 μg to 10 μg , preferably 1 μg to 10 μg , more preferably 3 μg to 10 μg . In a quail, about 5 μg is a preferred amount. In a chicken, about 5 μg to 10 μg per testis is preferred. These amounts of vector DNA may be injected in one dose or multiple doses and at one site or multiple sites in the testis. In a preferred embodiment, the vector DNA is administered at multiple sites in a single testis, both testes being injected in this manner. In one embodiment, injection is spread over three injection sites: one at each end of the testis, and one in the middle. Additionally, the transposon-based vector solution volume may be between approximately 1 μl to 75 μl in quail and between approximately 1 μl to 500 μl in chicken. In a preferred

embodiment, the transposon-based vector solution volume may be between approximately 20 μ l to 60 μ l in quail and between approximately 50 μ l to 250 μ l in chicken. Both the amount of vector DNA and the total volume injected into each testis may be determined based upon the age and size of the bird.

5 According to the present invention, the transposon-based vector is administered in conjunction with an acceptable carrier and/or transfection reagent. Acceptable carriers include, but are not limited to, water, saline, Hanks Balanced Salt Solution (HBSS), Tris-EDTA (TE) and lyotropic liquid crystals. Transfection reagents commonly known to one of ordinary skill in the art that may be employed
10 include, but are not limited to, the following: cationic lipid transfection reagents, cationic lipid mixtures, polyamine reagents, liposomes and combinations thereof; SUPERFECT®, Cytofectene, BioPORTER®, GenePORTER®, NeuroPORTER®, and perfectin from Gene Therapy Systems; lipofectamine, cellfectin, DMRIE-C oligofectamine, and PLUS reagent from InVitrogen; Xtreme gene, fugene, DOSPER
15 and DOTAP from Roche; Lipotaxi and Genejammer from Strategene; and Escort from SIGMA. In one embodiment, the transfection reagent is SUPERFECT®. The ratio of DNA to transfection reagent may vary based upon the method of administration. In one embodiment, the transposon-based vector is administered intratesticularly and the ratio of DNA to transfection reagent can be from 1:1.5 to
20 1:15, preferably 1:2 to 1:10, all expressed as wt/vol. Transfection may also be accomplished using other means known to one of ordinary skill in the art, including without limitation electroporation, gene guns, injection of naked DNA, and use of dimethyl sulfoxide (DMSO).

Depending upon the cell or tissue type targeted for transfection, the form of
25 the transposon-based vector may be important. Plasmids harvested from bacteria are generally closed circular supercoiled molecules, and this is the preferred state of a vector for gene delivery because of the ease of preparation. In some instances, transposase expression and insertion may be more efficient in a relaxed, closed circular configuration or in a linear configuration. In still other instances, a purified
30 transposase protein may be co-injected with a transposon-based vector containing the gene of interest for more immediate insertion. This could be accomplished by using a transfection reagent complexed with both the purified transposase protein and the transposon-based vector.

Testing for and Breeding Animals Carrying the Transgene

Following administration of a transposon-based vector to an animal, DNA is extracted from the animal to confirm integration of the gene of interest. Actual frequencies of integration are estimated both by comparative strength of the PCR signal, and by histological evaluation of the tissues by quantitative PCR. Another method for estimating the rate of transgene insertion is the so-called primed in situ hybridization technique (PRINS). This method determines not only which cells carry a transgene of interest, but also into which chromosome the gene has inserted, and even what portion of the chromosome. Briefly, labeled primers are annealed to chromosome spreads (affixed to glass slides) through one round of PCR, and the slides are then developed through normal in situ hybridization procedures. This technique combines the best features of in situ PCR and fluorescence in situ hybridization (FISH) to provide distinct chromosome location and copy number of the gene in question. The 28s rRNA gene will be used as a positive control for spermatogonia to confirm that the technique is functioning properly. Using different fluorescent labels for the transgene and the 28s gene causes cells containing a transgene to fluoresce with two different colored tags.

Breeding experiments are also conducted to determine if germline transmission of the transgene has occurred. In a general bird breeding experiment performed according to the present invention, each male bird was exposed to 2-3 different adult female birds for 3-4 days each. This procedure was continued with different females for a total period of 6-12 weeks. Eggs were collected daily for up to 14 days after the last exposure to the transgenic male, and each egg was incubated in a standard incubator. In the first series of experiments the resulting embryos were examined for transgene presence at day 3 or 4 using PCR.

Any male producing a transgenic embryo was bred to additional females. Eggs from these females were incubated, hatched, and the chicks tested for the exogenous DNA. Any embryos that died were necropsied and examined directly for the transgene or protein encoded by the transgene, either by fluorescence or PCR. The offspring that hatched and were found to be positive for the exogenous DNA were raised to maturity. These birds were bred to produce further generations of transgenic birds, to verify efficiency of the transgenic procedure and the stable

incorporation of the transgene into the germ line. The resulting embryos were examined for transgene presence at day 3 or 4 using PCR.

It is to be understood that the above procedure can be modified to suit animals other than birds and that selective breeding techniques may be performed to amplify gene copy numbers and protein output.

Production of Desired Proteins or Peptides in Egg White

In one embodiment, the transposon-based vectors of the present invention may be administered to a bird for production of desired proteins or peptides in the egg white. These transposon-based vectors preferably contain one or more of an ovalbumin promoter, an ovomucoid promoter, an ovalbumin signal sequence and an ovomucoid signal sequence. Oviduct-specific ovalbumin promoters are described in B. O'Malley et al., 1987. EMBO J., vol. 6, pp. 2305-12; A. Qiu et al., 1994. Proc. Nat. Acad. Sci. (USA), vol. 91, pp. 4451-4455; D. Monroe et al., 2000. Biochim. Biophys. Acta, 1517 (1):27-32; H. Park et al., 2000. Biochem., 39:8537-8545; and T. Muramatsu et al., 1996. Poul. Avian Biol. Rev., 6:107-123. Examples of transposon-based vectors designed for production of a desired protein in an egg white are shown in Figures 2 and 3.

Production of Desired Proteins or Peptides in Egg Yolk

The present invention is particularly advantageous for production of recombinant peptides and proteins of low solubility in the egg yolk. Such proteins include, but are not limited to, membrane-associated or membrane-bound proteins, lipophilic compounds; attachment factors, receptors, and components of second messenger transduction machinery. Low solubility peptides and proteins are particularly challenging to produce using conventional recombinant protein production techniques (cell and tissue cultures) because they aggregate in water-based, hydrophilic environments. Such aggregation necessitates denaturation and refolding of the recombinantly-produced proteins, which may deleteriously affect their structure and function. Moreover, even highly soluble recombinant peptides and proteins may precipitate and require denaturation and renaturation when produced in sufficiently high amounts in recombinant protein production systems. The present invention provides an advantageous resolution of the problem of protein and peptide solubility during production of large amounts of recombinant proteins.

In one embodiment of the present invention, deposition of a desired protein into the egg yolk is accomplished by attaching a sequence encoding a protein capable of binding to the yolk vitellogenin receptor to a gene of interest that encodes a desired protein. This transposon-based vector can be used for the receptor-mediated uptake of the desired protein by the oocytes. In a preferred embodiment, the sequence ensuring the binding to the vitellogenin receptor is a targeting sequence of a vitellogenin protein. The invention encompasses various vitellogenin proteins and their targeting sequences. In a preferred embodiment, a chicken vitellogenin protein targeting sequence is used, however, due to the high degree of conservation among vitellogenin protein sequences and known cross-species reactivity of vitellogenin targeting sequences with their egg-yolk receptors, other vitellogenin targeting sequences can be substituted. One example of a construct for use in the transposon-based vectors of the present invention and for deposition of an insulin protein in an egg yolk is provided in SEQ ID NO:27. In this embodiment, the transposon-based vector contains a vitellogenin promoter, a vitellogenin targeting sequence, a TAG sequence, a pro-insulin sequence and a synthetic polyA sequence. The present invention includes, but is not limited to, vitellogenin targeting sequences residing in the N-terminal domain of vitellogenin, particularly in lipovitellin I. In one embodiment, the vitellogenin targeting sequence contains the polynucleotide sequence of SEQ ID NO:18.

In a preferred embodiment, the transposon-based vector contains a transposase gene operably-linked to a liver-specific promoter and a gene of interest operably-linked to a liver-specific promoter and a vitellogenin targeting sequence. Figure 4 shows an example of such a construct. In another preferred embodiment, the transposon-based vector contains a transposase gene operably-linked to a constitutive promoter and a gene of interest operably-linked to a liver-specific promoter and a vitellogenin targeting sequence.

Isolation and Purification of Desired Protein or Peptide

For large-scale production of protein, an animal breeding stock that is homozygous for the transgene is preferred. Such homozygous individuals are obtained and identified through, for example, standard animal breeding procedures or PCR protocols.

Once expressed, peptides, polypeptides and proteins can be purified according to standard procedures known to one of ordinary skill in the art, including ammonium sulfate precipitation, affinity columns, column chromatography, gel electrophoresis, high performance liquid chromatography, immunoprecipitation and the like.

- 5 Substantially pure compositions of about 50 to 99% homogeneity are preferred, and 80 to 95% or greater homogeneity are most preferred for use as therapeutic agents.

In one embodiment of the present invention, the animal in which the desired protein is produced is an egg-laying animal. In a preferred embodiment of the present invention, the animal is an avian and a desired peptide, polypeptide or protein is
10 isolated from an egg white. Egg white containing the exogenous protein or peptide is separated from the yolk and other egg constituents on an industrial scale by any of a variety of methods known in the egg industry. See, e.g., W. Stadelman et al. (Eds.), Egg Science & Technology, Haworth Press, Binghamton, NY (1995). Isolation of the exogenous peptide or protein from the other egg white constituents is accomplished
15 by any of a number of polypeptide isolation and purification methods well known to one of ordinary skill in the art. These techniques include, for example, chromatographic methods such as gel permeation, ion exchange, affinity separation, metal chelation, HPLC, and the like, either alone or in combination. Another means that may be used for isolation or purification, either in lieu of or in addition to
20 chromatographic separation methods, includes electrophoresis. Successful isolation and purification is confirmed by standard analytic techniques, including HPLC, mass spectroscopy, and spectrophotometry. These separation methods are often facilitated if the first step in the separation is the removal of the endogenous ovalbumin fraction of egg white, as doing so will reduce the total protein content to be further purified by
25 about 50%.

To facilitate or enable purification of a desired protein or peptide, transposon-based vectors may include one or more additional epitopes or domains. Such epitopes or domains include DNA sequences encoding enzymatic or chemical cleavage sites including, but not limited to, an enterokinase cleavage site; the glutathione binding
30 domain from glutathione S-transferase; polylysine; hexa-histidine or other cationic amino acids; thioredoxin; hemagglutinin antigen; maltose binding protein; a fragment of gp41 from HIV; and other purification epitopes or domains commonly known to one of skill in the art.

In one representative embodiment, purification of desired proteins from egg white utilizes the antigenicity of the ovalbumin carrier protein and particular attributes of a TAG linker sequence that spans ovalbumin and the desired protein. The TAG sequence is particularly useful in this process because it contains 1) a highly antigenic epitope, a fragment of gp41 from HIV, allowing for stringent affinity purification, and, 2) a recognition site for the protease enterokinase immediately juxtaposed to the desired protein. In a preferred embodiment, the TAG sequence comprises approximately 50 amino acids. A representative TAG sequence is provided below.

10 Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp
Ala Pro Ala Asp Asp Ala Pro Ala Asp Asp Ala Thr Thr Cys Ile Leu Lys Gly Ser Cys
Gly Trp Ile Gly Leu Leu Asp Asp Asp Asp Lys (SEQ ID NO:22)

The underlined sequences were taken from the hairpin loop domain of HIV gp-41 (SEQ ID NO:23). Sequences in italics represent the cleavage site for enterokinase (SEQ ID NO:9). The spacer sequence upstream of the loop domain was made from repeats of (Pro Ala Asp Asp Ala) (SEQ ID NO:25) to provide free rotation and promote surface availability of the hairpin loop from the ovalbumin carrier protein.

Isolation and purification of a desired protein is performed as follows:

- 20 1. Enrichment of the egg white protein fraction containing ovalbumin and the transgenic ovalbumin-TAG-desired protein.
2. Size exclusion chromatography to isolate only those proteins within a narrow range of molecular weights (a further enrichment of step 1).
3. Ovalbumin affinity chromatography. Highly specific antibodies to ovalbumin will eliminate virtually all extraneous egg white proteins except ovalbumin and the transgenic ovalbumin-TAG-desired protein.
- 25 4. gp41 affinity chromatography using anti-gp41 antibodies. Stringent application of this step will result in virtually pure transgenic ovalbumin-TAG-desired protein.
- 30 5. Cleavage of the transgene product can be accomplished in at least one of two ways:
 - a. The transgenic ovalbumin-TAG-desired protein is left attached to the gp41 affinity resin (beads) from step 4 and the protease enterokinase is

- added. This liberates the transgene target protein from the gp41 affinity resin while the ovalbumin-TAG sequence is retained. Separation by centrifugation (in a batch process) or flow through (in a column purification), leaves the desired protein together with enterokinase in solution. Enterokinase is recovered and reused.
- 5 b. Alternatively, enterokinase is immobilized on resin (beads) by the addition of poly-lysine moieties to a non-catalytic area of the protease. The transgenic ovalbumin-TAG-desired protein eluted from the affinity column of step 4 is then applied to the protease resin. Protease
- 10 action cleaves the ovalbumin-TAG sequence from the desired protein and leaves both entities in solution. The immobilized enterokinase resin is recharged and reused.
- c. The choice of these alternatives is made depending upon the size and chemical composition of the transgene target protein.
- 15 6. A final separation of either of these two (5a or 5b) protein mixtures is made using size exclusion, or enterokinase affinity chromatography. This step allows for desalting, buffer exchange and/or polishing, as needed.
- Cleavage of the transgene product (ovalbumin-TAG-desired protein) by enterokinase, then, results in two products: ovalbumin-TAG and the desired protein.
- 20 More specific methods for isolation using the TAG label is provided in the Examples. Some desired proteins may require additions or modifications of the above-described approach as known to one of ordinary skill in the art. The method is scaleable from the laboratory bench to pilot and production facility largely because the techniques applied are well documented in each of these settings.
- 25 It is believed that a typical chicken egg produced by a transgenic animal of the present invention will contain at least 0.001 mg, from about 0.001 to 1.0 mg, or from about 0.001 to 100.0 mg of exogenous protein, peptide or polypeptide, in addition to the normal constituents of egg white (or possibly replacing a small fraction of the latter).
- 30 One of skill in the art will recognize that after biological expression or purification, the desired proteins, fragments thereof and peptides may possess a conformation substantially different than the native conformations of the proteins, fragments thereof and peptides. In this case, it is often necessary to denature and

reduce protein and then to cause the protein to re-fold into the preferred conformation. Methods of reducing and denaturing proteins and inducing re-folding are well known to those of skill in the art.

Production of Protein or Peptide in Milk

5 In addition to methods of producing eggs containing transgenic proteins or peptides, the present invention encompasses methods for the production of milk containing transgenic proteins or peptides. These methods include the administration of a transposon-based vector described above to a mammal. In one embodiment, the transposon-based vector contains a transposase operably-linked to a constitutive
10 promoter and a gene of interest operably-linked to mammary specific promoter. Genes of interest can include, but are not limited to antiviral and antibacterial proteins and immunoglobulins.

Treatment of Disease and Animal Improvement

In addition to production and isolation of desired molecules, the transposon-
15 based vectors of the present invention can be used for the treatment of various genetic disorders. For example, one or more transposon-based vectors can be administered to a human or animal for the treatment of a single gene disorder including, but not limited to, Huntington's disease, alpha-1-antitrypsin deficiency Alzheimer's disease, various forms of breast cancer, cystic fibrosis, galactosemia, congenital
20 hypothyroidism, maple syrup urine disease, neurofibromatosis 1, phenylketonuria, sickle cell disease, and Smith-Lemli-Opitz (SLO/RSH) Syndrome. Other diseases caused by single gene disorders that may be treated with the present invention include, autoimmune diseases, shipping fever in cattle, mastitis, bacterial or viral diseases, alteration of skin pigment in animals. In these embodiments, the
25 transposon-based vector contains a non-mutated, or non-disease causing form of the gene known to cause such disorder. Preferably, the transposase contained within the transposase-based vector is operably linked to an inducible promoter such as a tissue-specific promoter such that the non-mutated gene of interest is inserted into a specific tissue wherein the mutated gene is expressed in vivo.

30 In one embodiment of the present invention, a transposon-based vector comprising a gene encoding proinsulin is administered to diabetic animals or humans for incorporation into liver cells in order to treat or cure diabetes. The specific incorporation of the proinsulin gene into the liver is accomplished by placing the

transposase gene under the control of liver-specific promoter, such as G6P. This approach is useful for treatment of both Type I and Type II diabetes. The G6P promoter has been shown to be glucose responsive (Arguad, D., et al. 1996. Diabetes 45:1563-1571), and thus, glucose-regulated insulin production is achieved using DNA constructs of the present invention. Integrating a proinsulin gene into liver cells circumvents the problem of destruction of pancreatic islet cells in the course of Type I diabetes.

In another embodiment, shortly after diagnosis of Type I diabetes, the cells of the immune system destroying pancreatic β -cells are selectively removed using the transposon-based vectors of the present invention, thus allowing normal β -cells to repopulate the pancreas.

For treatment of Type II diabetes, a transposon-based vector containing a proinsulin gene is specifically incorporated into the pancreas by placing the transposase gene under the control of a pancreas-specific promoter, such as an insulin promoter. In this embodiment, the vector is delivered to a diabetic animal or human via injection into an artery feeding the pancreas. For delivery, the vector is complexed with a transfection agent. The artery distributes the complex throughout the pancreas, where individual cells receive the vector DNA. Following uptake into the target cell, the insulin promoter is recognized by transcriptional machinery of the cell, the transposase encoded by the vector is expressed, and stable integration of the proinsulin gene occurs. It is expected that a small percentage of the transposon-based vector is transported to other tissues, and that these tissues are transfected. However, these tissues are not stably transfected and the proinsulin gene is not incorporated into the cells' DNA due to failure of these cells to activate the insulin promoter. The vector DNA is likely lost when the cell dies or degraded over time.

In other embodiments, one or more transposon-based vectors are administered to an avian for the treatment of a viral or bacterial infection/disease including, but not limited to, Colibacillosis (Coliform infections), Mycoplasmosis (CRD, Air sac, Sinusitis), Fowl Cholera, Necrotic Enteritis, Ulcerative Enteritis (Quail disease), Pullorum Disease, Fowl Typhoid, Botulism, Infectious Coryza, Erysipelas, Avian Pox, Newcastle Disease, Infectious Bronchitis, Quail Bronchitis, Lymphoid Leukosis, Marek's Disease (Visceral Leukosis), Infectious Bursal Disease (Gumboro). In these

embodiments, the transposon-based vectors may be used in a manner similar to traditional vaccines.

In still other embodiments, one or more transposon-based vectors are administered to an animal for the production of an animal with enhanced growth characteristics and nutrient utilization.

The transposon-based vectors of the present invention can be used to transform any animal cell, including but not limited to: cells producing hormones, cytokines, growth factors, or any other biologically active substance; cells of the immune system; cells of the nervous system; muscle (striatal, cardiac, smooth) cells; vascular system cells; endothelial cells; skin cells; mammary cells; and lung cells, including bronchial and alveolar cells. Transformation of any endocrine cell by a transposon-based vector is contemplated as a part of a present invention. In one aspect of the present invention, cells of the immune system may be the target for incorporation of a desired gene or genes encoding for production of antibodies. Accordingly, the thymus, bone marrow, beta lymphocytes (or B cells), gastrointestinal associated lymphatic tissue (GALT), Peyer's patches, bursa Fabricius, lymph nodes, spleen, and tonsil, and any other lymphatic tissue, may all be targets for administration of the compositions of the present invention.

The transposon-based vectors of the present invention can be used to modulate (stimulate or inhibit) production of any substance, including but not limited to a hormone, a cytokine, or a growth factor, by an animal or a human cell. Modulation of a regulated signal within a cell or a tissue, such as production of a second messenger, is also contemplated as a part of the present invention. Use of the transposon-based vectors of the present invention is contemplated for treatment of any animal or human disease or condition that results from underproduction (such as diabetes) or overproduction (such as hyperthyroidism) of a hormone or other endogenous biologically active substance. Use of the transposon-based vectors of the present invention to integrate nucleotide sequences encoding RNA molecules, such as anti-sense RNA or short interfering RNA, is also contemplated as a part of the present invention.

Additionally, the transposon-based vectors of the present invention may be used to provide cells or tissues with "beacons", such as receptor molecules, for binding of therapeutic agents in order to provide tissue and cell specificity for the

therapeutic agents. Several promoters and exogenous genes can be combined in one vector to produce progressive, controlled treatments from a single vector delivery.

The following examples will serve to further illustrate the present invention without, at the same time, however, constituting any limitation thereof. On the contrary, it is to be clearly understood that resort may be had to various embodiments, modifications and equivalents thereof which, after reading the description herein, may suggest themselves to those skilled in the art without departing from the spirit of the invention.

EXAMPLE 1

Preparation of Transposon-Based Vector pTnMod

A vector was designed for inserting a desired coding sequence into the genome of eukaryotic cells, given below as SEQ ID NO:1. The vector of SEQ ID NO:1, termed pTnMod, was constructed and its sequence verified.

This vector employed a cytomegalovirus (CMV) promoter. A modified Kozak sequence (ACCATG) (SEQ ID NO:13) was added to the promoter. The nucleotide in the wobble position in nucleotide triplet codons encoding the first 10 amino acids of transposase was changed to an adenine (A) or thymine (T), which did not alter the amino acid encoded by this codon. Two stop codons were added and a synthetic polyA was used to provide a strong termination sequence. This vector uses a promoter designed to be active soon after entering the cell (without any induction) to increase the likelihood of stable integration. The additional stop codons and synthetic polyA insures proper termination without read through to potential genes downstream.

The first step in constructing this vector was to modify the transposase to have the desired changes. Modifications to the transposase were accomplished with the primers High Efficiency forward primer (Hef) Altered transposase (ATS)-Hef 5' ATCTCGAGACCATGTGTGAACTTGATATTTACATGATCTCTTTACC 3' (SEQ ID NO:10) and Altered transposase- High efficiency reverse primer (Her) 5' GATTGATCATTATCATAATTTCCCAAAGCGTAACC 3' (SEQ ID NO:11, a reverse complement primer). In the 5' forward primer ATS-Hef, the sequence CTCGAG (SEQ ID NO:12) is the recognition site for the restriction enzyme Xho I, which permits directional cloning of the amplified gene. The sequence ACCATG

(SEQ ID NO:13) contains the Kozak sequence and start codon for the transposase and the underlined bases represent changes in the wobble position to an A or T of codons for the first 10 amino acids (without changing the amino acid coded by the codon). Primer ATS-Her (SEQ ID NO:11) contains an additional stop codon TAA in addition to native stop codon TGA and adds a Bcl I restriction site, TGATCA (SEQ ID NO:14), to allow directional cloning. These primers were used in a PCR reaction with pTnLac (p defines plasmid, tn defines transposon, and lac defines the beta fragment of the lactose gene, which contains a multiple cloning site) as the template for the transposase and a FailSafe™ PCR System (which includes enzyme, buffers, dNTP's, MgCl₂ and PCR Enhancer; Epicentre Technologies, Madison, WI). Amplified PCR product was electrophoresed on a 1% agarose gel, stained with ethidium bromide, and visualized on an ultraviolet transilluminator. A band corresponding to the expected size was excised from the gel and purified from the agarose using a Zymo Clean Gel Recovery Kit (Zymo Research, Orange, CA). Purified DNA was digested with restriction enzymes Xho I (5') and Bcl I (3') (New England Biolabs, Beverly, MA) according to the manufacturer's protocol. Digested DNA was purified from restriction enzymes using a Zymo DNA Clean and Concentrator kit (Zymo Research).

Plasmid gWhiz (Gene Therapy Systems, San Diego, CA) was digested with restriction enzymes Sal I and BamH I (New England Biolabs), which are compatible with Xho I and Bcl I, but destroy the restriction sites. Digested gWhiz was separated on an agarose gel, the desired band excised and purified as described above. Cutting the vector in this manner facilitated directional cloning of the modified transposase (mATS) between the CMV promoter and synthetic polyA.

To insert the mATS between the CMV promoter and synthetic polyA in gWhiz, a Stratagene T4 Ligase Kit (Stratagene, Inc. La Jolla, CA) was used and the ligation set up according to the manufacturer's protocol. Ligated product was transformed into *E. coli* Top10 competent cells (Invitrogen Life Technologies, Carlsbad, CA) using chemical transformation according to Invitrogen's protocol. Transformed bacteria were incubated in 1 ml of SOC (GIBCO BRL, CAT# 15544-042) medium for 1 hour at 37° C before being spread to LB (Luria-Bertani media (broth or agar)) plates supplemented with 100 µg/ml ampicillin (LB/amp plates). These plates were incubated overnight at 37° C and resulting colonies picked to

LB/amp broth for overnight growth at 37° C. Plasmid DNA was isolated using a modified alkaline lysis protocol (Sambrook et al., 1989), electrophoresed on a 1% agarose gel, and visualized on a U.V. transilluminator after ethidium bromide staining. Colonies producing a plasmid of the expected size (approximately 6.4 kbp) were cultured in at least 250 ml of LB/amp broth and plasmid DNA harvested using a Qiagen Maxi-Prep Kit (column purification) according to the manufacturer's protocol (Qiagen, Inc., Chatsworth, CA). Column purified DNA was used as template for sequencing to verify the changes made in the transposase were the desired changes and no further changes or mutations occurred due to PCR amplification. For sequencing, Perkin-Elmer's Big Dye Sequencing Kit was used. All samples were sent to the Gene Probes and Expression Laboratory (LSU School of Veterinary Medicine) for sequencing on a Perkin-Elmer Model 377 Automated Sequencer.

Once a clone was identified that contained the desired mATS in the correct orientation, primers CMVf-NgoM IV (5' TTGCCGGCATCAGATTGGCTAT (SEQ ID NO:15); underlined bases denote NgoM IV recognition site) and Syn-polyA-BstE II (5' AGAGGTCAACCGGTCAATTCTTCAGCACCTGGTA (SEQ ID NO:16); underlined bases denote BstE II recognition site) were used to PCR amplify the entire CMV promoter, mATS, and synthetic polyA for cloning upstream of the transposon in pTnLac. The PCR was conducted with FailSafe™ as described above, purified using the Zymo Clean and Concentrator kit, the ends digested with NgoM IV and BstE II (New England Biolabs), purified with the Zymo kit again and cloned upstream of the transposon in pTnLac as described below.

Plasmid pTnLac was digested with NgoM IV and BstE II to remove the ptac promoter and transposase and the fragments separated on an agarose gel. The band corresponding to the vector and transposon was excised, purified from the agarose, and dephosphorylated with calf intestinal alkaline phosphatase (New England Biolabs) to prevent self-annealing. The enzyme was removed from the vector using a Zymo DNA Clean and Concentrator-5. The purified vector and CMVp/mATS/polyA were ligated together using a Stratagene T4 Ligase Kit and transformed into *E. coli* as described above.

Colonies resulting from this transformation were screened (mini-preps) as describe above and clones that were the correct size were verified by DNA sequence

analysis as described above. The vector was given the name pTnMod (SEQ ID NO:1) and includes the following components:

Base pairs 1-130 are a remainder of F1(-) on from pBluescriptII sk(-) (Stratagene), corresponding to base pairs 1-130 of pBluescriptII sk(-).

5 Base pairs 131 - 132 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 133 -1777 are the CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems), corresponding to bp 229-1873 of pGWiz. The CMV promoter was modified by the addition of an ACC sequence upstream of ATG.

10 Base pairs 1778-1779 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 1780 - 2987 are the coding sequence for the transposase, modified from Tn10 (GenBank accession J01829) by optimizing codons for stability of the transposase mRNA and for the expression of protein. More specifically, in each of the
15 codons for the first ten amino acids of the transposase, G or C was changed to A or T when such a substitution would not alter the amino acid that was encoded.

Base pairs 2988-2993 are two engineered stop codons.

Base pair 2994 is a residue from ligation of restriction enzyme sites used in constructing the vector.

20 Base pairs 2995 - 3410 are a synthetic polyA sequence taken from the pGWiz vector (Gene Therapy Systems), corresponding to bp 1922-2337 of 10 pGWiz.

Base pairs 3415 - 3718 are non-coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are non-coding λ DNA that is residual from pNK2859.

25 Base pairs 3762 - 3831 are the 70 bp of the left insertion sequence recognized by the transposon Tn10.

Base pairs 3832-3837 are a residue from ligation of restriction enzyme sites used in constructing the vector.

30 Base pairs 3838 - 4527 are the multiple cloning site from pBluescriptII sk(20), corresponding to bp 924-235 of pBluescriptII sk(-). This multiple cloning site may be used to insert any coding sequence of interest into the vector.

Base pairs 4528-4532 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 4533 - 4602 are the 70 bp of the right insertion sequence recognized by the transposon Tn10.

Base pairs 4603 - 4644 are non-coding λ DNA that is residual from pNK2859.

Base pairs 4645 - 5488 are non-coding DNA that is residual from pNK2859.

5 Base pairs 5489 - 7689 are from the pBluescriptII sk(-) base vector - (Stratagene, Inc.), corresponding to bp 761-2961 of pBluescriptII sk(-).

Completing pTnMod is a pBlueScript backbone that contains a *colE1* origin of replication and an antibiotic resistance marker (ampicillin).

10 It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

All plasmid DNA was isolated by standard procedures. Briefly, *Escherichia coli* containing the plasmid was grown in 500 mL aliquots of LB broth (supplemented with an appropriate antibiotic) at 37°C overnight with shaking. Plasmid DNA was 15 recovered from the bacteria using a Qiagen Maxi-Prep kit (Qiagen, Inc., Chatsworth, CA) according to the manufacturer's protocol. Plasmid DNA was resuspended in 500 μ L of PCR-grade water and stored at -20°C until used.

EXAMPLE 2

20 *Preparation of Transposon-Based Vector pTnMod (CMV/Red)*

A vector was designed for inserting a reporter gene (DsRed) under the control of the CMV promoter into the genome of vertebrate cells given below as SEQ ID NO:2. The reporter gene chosen was the DsRed gene, driven by the immediate early cytomegalovirus promoter, to produce a plasmid called pTnCMV/DsRed. The DsRed 25 gene product is a red fluorescent protein from an IndoPacific sea anemone, *Discosoma* sp., which fluoresces bright red at 558 nm. It is to be understood that the reporter gene, i.e., the DsRed gene, is only one embodiment of the present invention and that any gene of interest may be inserted into the plasmid in place of the DsRed reporter gene in any Experiment described herein.

30 The vector of SEQ ID NO:2, named pTnMod (CMV/Red), was constructed, and its sequence verified by re-sequencing. SEQ ID NO:2, pTnMod (CMV/Red), includes the following components:

Base pairs 1-130 are a remainder of F1(-) on from pBluescriptII sk(-) (Stratagene), corresponding to bp 1-130 of pBluescriptII sk(-).

Base pairs 131 - 132 are a residue from ligation of restriction enzyme sites used in constructing the vector.

5 Base pairs 133 -1777 are the CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems, corresponding to bp 229-1873 of pGWiz.

Base pairs 1778-1779 are a residue from ligation of restriction enzyme sites used in constructing the vector.

10 Base pairs 1780 - 2987 are the coding sequence for the transposase, modified from Tn10 (GenBank accession J01829) by optimizing codons as discussed above.

Base pairs 2988-2993 are two engineered stop codons.

Base pair 2994 is a residue from ligation of restriction enzyme sites used in constructing the vector.

15 Base pairs 2995 - 3410 are a synthetic polyA sequence taken from the pGWiz vector (Gene Therapy Systems), corresponding to bp 1922-2337 of pGWiz.

Base pairs 3415 - 3718 are non-coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are non-coding λ DNA that is residual from pNK2859.

20 Base pairs 3762 - 3831 are the 70 bp of the left insertion sequence recognized by the transposon Tn10.

Base pairs 3832-3837 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 3838 - 4044 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 924-718 of pBluescriptII sk(-).

25 Base pairs 4045-4048 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 4049-5693 are the CMV promoter/enhancer, taken from vector pGWiz (Gene Therapy Systems), corresponding to bp 229-1873 of pGWiz.

30 Base pairs 5694-5701 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 5702 - 6617 are the DsRed reporter coding sequence, including polyA sequence, from pDsRed1.1 (Clontech), corresponding to bp 77 - 992 of pDsRed1.1.

Base pairs 6618 - 7101 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 718-235 of pBluescriptII sk(-).

Base pairs 7102-7106 are a residue from ligation of restriction enzyme sites used in constructing the vector.

5 Base pairs 7107 - 7176 are the 70 bp of the right insertion sequence recognized by the transposon Tn10.

Base pairs 7177 - 7218 are non-coding λ DNA that is residual from pNK2859.

Base pairs 7219 - 8062 are non-coding DNA that is residual from pNK2859.

10 Base pairs 8063 - 10263 are from the pBluescriptII sk(-) base vector (Stratagene, Inc.), corresponding to bp 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s).

EXAMPLE 3

15 *Preparation of Transposon-Based Vector pTnMod (Oval/Red) - Chicken*

A vector was designed for inserting a reporter gene (DsRed) under the control of the ovalbumin promoter, and including the ovalbumin signal sequence, into the genome of a bird. One version of this vector is given below as SEQ ID NO:3. The vector of SEQ ID NO:3, named pTnMod (Oval/Red) - Chicken, includes chicken
20 ovalbumin promoter and signal sequences.

SEQ ID NO:3, pTnMod (Oval/Red) - Chicken, includes the following components:

Base pairs 1-130 are a remainder of F1(-) on from pBluescriptII sk(-) (Stratagene), corresponding to bp 1-130 of pBluescriptII sk(-).

25 Base pairs 131 - 132 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 133 -1777 are the CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems, corresponding to bp 229-1873 of pGWiz.

30 Base pairs 1778-1779 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 1780 - 2987 are the coding sequence for the transposase, modified from Tn10 (GenBank accession J01829) by optimizing codons as discussed above.

Base pairs 2988-2993 are two engineered stop codons.

Base pair 2994 is a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 2995 - 3410 are a synthetic polyA sequence taken from the pGWiz vector (Gene Therapy Systems), corresponding to bp 1922-2337 of pGWiz.

5 Base pairs 3415 --3718 are non-coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are non-coding λ DNA that is residual from 10 pNK2859.

Base pairs 3762 - 3831 are the 70 bp of the left insertion sequence recognized
10 by the transposon Tn10.

Base pairs 3832-3837 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 3838 - 4044 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 924-718 of pBluescriptII sk(-).

15 Base pairs 4045-4049 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 4050 - 4951 contain upstream elements of the (including SDRE, steroid-dependent response element). See GenBank accession number J00895 M24999, bp 431-1332. Base pairs 4952-4959 are a residue from ligation of
20 restriction enzyme sites used in constructing the vector.

Base pairs 4960 - 5112 are the chicken ovalbumin signal sequence (GenBank accession number J00895 M24999, bp 2996-3148).

Base pairs 5113-5118 are a residue from ligation of restriction enzyme sites used in constructing the vector.

25 Base pairs 5119 - 6011 are the DsRed reporter coding sequence, including polyA sequence, from pDsRed1.1 (Clontech), corresponding to bp 100 - 992 of pDsRed1.1.

Base pairs 6012-6017 are a residue from ligation of restriction enzyme sites used in constructing the vector.

30 Base pairs 6018 - 6056 are part of the multiple cloning site of the ZeroBlunt Topo cloning vector (Invitrogen), corresponding to bp 337-377 of ZeroBlunt.

Base pairs 6057-6062 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 6063 - 6495 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 667-235 of pBluescriptII sk(-).

Base pairs 6496-6500 are a residue from ligation of restriction enzyme sites used in constructing the vector.

5 Base pairs 6501 - 6570 are the 70 bp of the right insertion sequence recognized by the transposon Tn10.

Base pairs 6571 - 6612 are non-coding λ DNA that is residual from pNK2859.

Base pairs 6613 - 7477 are non-coding DNA that is residual from pNK2859.

10 Base pairs 7478 - 9678 are from the pBluescriptII sk(-) base vector (Stratagene, Inc.), corresponding to bp 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s).

EXAMPLE 4

15 *Preparation of Transposon-Based Vector pTnMod(Oval/Red) - Quail*

A vector was designed for inserting a reporter gene (DsRed) under the control of the ovalbumin promoter, and including the ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:4. The vector of SEQ ID NO:4, named pTnMod (Oval/Red) - Quail, has been constructed, and selected portions of the
20 sequence have been verified by re-sequencing.

SEQ ID NO:4, pTnMod (Oval/Red) - Quail, includes the following components:

Base pairs 1-130 are a remainder of F1(-) on from pBluescriptII sk(-) (Stratagene), corresponding to bp 1-130 of pBluescriptII sk(-).

25 Base pairs 131 - 132 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 133 - 1777 are the CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems), corresponding to bp 229-1873 of pGWiz.

30 Base pairs 1778-1779 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 1780 - 2987 are the coding sequence for the transposase, modified from Tn10 (GenBank accession J01829) by optimizing codons as discussed above.

Base pairs 2988-2993 are two engineered stop codons. Base pair 2994 is a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 2995 - 3410 are a synthetic polyA sequence taken from the pGWiz vector (Gene Therapy Systems), corresponding to bp 1922-2337 of pGWiz.

5 Base pairs 3415 - 3718 are non-coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are non-coding λ DNA that is residual from pNK2859.

Base pairs 3762 - 3831 are the 70 base pairs of the left insertion sequence recognized by the transposon Tn10.

10 Base pairs 3832-3837 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 3838 - 4044 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 924-718 of pBluescriptII sk(-).

15 Base pairs 4045-4049 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 4050 - 4934 are the Japanese quail ovalbumin promoter (including SDRE, steroid-dependent response element). The Japanese quail ovalbumin promoter was isolated by its high degree of homology to the chicken ovalbumin promoter (GenBank accession number J00895 M24999, base pairs 431-1332). Some deletions were noted in the quail sequence, as compared to the chicken sequence.

20 Base pairs 4935-4942 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 4943 - 5092 are the Japanese quail ovalbumin signal sequence. The quail signal sequence was isolated by its high degree of homology to the chicken signal sequence (GenBank accession number J00895 M24999, base pairs 2996-3148). Some deletions were noted in the quail sequence, as compared to the chicken sequence.

Base pairs 5093-5098 are a residue from ligation of restriction enzyme sites used in constructing the vector.

30 Base pairs 5099 - 5991 are the DsRed reporter coding sequence, including polyA sequence, from pDsRed1.1 (Clontech), corresponding to bp 100 - 992 of pDsRed 1.1.

Base pairs 5992-5997 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 5998 - 6036 are part of the multiple cloning site of the ZeroBlunt Topo cloning vector (Invitrogen), corresponding to base pairs 337-377 of ZeroBlunt.

5 Base pairs 6037-6042 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 6043 - 6475 are part of the multiple cloning site from pBluescriptII sk(-), corresponding to bp 667-235 of pBluescriptII sk(-).

10 Base pairs 6476-6480 are a residue from ligation of restriction enzyme sites used in constructing the vector.

Base pairs 6481 - 6550 are the 70 bp of the right insertion sequence recognized by the transposon Tn10.

Base pairs 6551 - 6592 are non-coding λ DNA that is residual from pNK2859.

Base pairs 6593 - 7457 are non-coding DNA that is residual from pNK2859.

15 Base pairs 7458 - 9658 are from the pBluescriptII sk(-) base vector (Stratagene, Inc.), corresponding to base pairs 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s).

20 EXAMPLE 5

Transfection of Stage X Japanese Quail Eggs with pTnMod(Oval/Red) – Quail via embryo injection

Transgenic Japanese quail were produced by transfecting Stage X embryos and the heritability of the transgene delivered by embryo transfection was established. 25 More specifically, fertile eggs were collected in the morning and placed at 15° C until enough were collected for injection, but were held no longer than 7 days. Stage X embryos (eggs) were assigned to one of two treatment groups. Before treatment, each egg was incubated on its side at room temperature for about 2 hours to allow the embryo to move to “top dead center” (TDC). Each egg was transfected by drilling a 1 30 mm hole (directly above the embryo) through the shell without penetrating the underlying shell membrane. A 0.5 ml syringe fitted with a 28 gauge needle was used to deliver DNA complexed to a transfecting reagent, i.e. SUPERFECT®, in a 50 μ l volume. An adhesive disc was used to seal the hole and provide a label for treatment identification. After all eggs were transfected, they were set in an incubator with the 35 adhesive disc pointing upward for hatching.

Each bird that hatched was bled at one week of age, DNA was extracted from blood cells, and PCR was conducted using 28s primers as a positive control and primers specific to DsRed. Any bird that was negative was terminated, while positive birds were monitored to determine maintenance of the transgene. Birds consistently positive were maintained until sexual maturity and bred. Positive male and female birds were mated. The eggs of mated hens were hatched and the resulting chicks, the G1 generation, were evaluated to determine if they were transgenic. All G1s resulting from this mating were bled and PCR conducted as described above.

Egg injection: Two treatment groups and one control group were used for this experiment. Vector pTnMod (Oval/Red) in supercoiled form (Treatment 1) and in linear form (Treatment 2) were used to transfect 15 eggs per treatment. To obtain linear DNA for this experiment, pTnMod (Oval/Red) was digested with NgoM IV, column purified, and resuspended in TE buffer.

Each egg was injected with 0.75 µg of DNA complexed with SUPERFECT® in a 1:3 ratio in a total injection volume of 50 µl Hank's Balanced Salt solution (HBSS) was used to bring the volume to 50 µl. The DNA Superfect mixture must be allowed to incubate (for complex formation) at room temperature for 10 minutes prior to injection and must be used within 40 minutes post initial mixing. Eggs were incubated as described above after injection.

Results: In the supercoiled injection group, 2 females and 1 male were identified as PCR positive using primers specific to the DsRed coding sequence. These birds were mated as described above. Blood was taken from the G1 chicks and PCR was conducted. The results showed that the transgene was incorporated into the gametes of these birds. The G1 chicks from these birds were examined on a weekly basis until it was verified that the gene was not present or enough transgenic G1s were obtained to initiate a breeding flock of fully transgenic birds. Eggs from these G1 chicks expressed DsRed protein in the albumin portion of their eggs.

EXAMPLE 6

Intratesticular Injection of Chickens with pTnMod(CMV/Red) (SEQ ID NO:2)

Immature birds of different ages (4, 6, 8, 10, 12, and 14 weeks) were placed under anesthesia and injected in the testes with the construct pTnMod(CMV/Red). A saline solution containing 1-5 µg of purified DNA vector, mixed with SUPERFECT® transfecting reagent (Qiagen, Valencia, CA) in a 1:6 (wt:vol) ratio. The volume of saline was adjusted so that the total volume injected into each testis was 150-200 µl, depending on the age and size of the bird. For the 4- and 6-week-old chickens, 1 µg DNA in 150 µl was injected in each testis, divided into three doses of 50 µl each. For

the older birds, 200 µl total volume was injected, containing either 3 µg DNA (for 8-week-old birds) or 5 µg DNA (for older birds) per testis. First, one testis was surgically exposed prior to injection. After injection, the incision was sutured, and the sequence was repeated for the alternate testis.

5 From six to nine months post-surgery, weekly sperm samples were taken from each injected bird, as well as from control birds. Each sperm sample was evaluated for uptake and expression of the injected gene. Samples were evaluated by PCR on whole sperm, within one week after collection.

10 Approximately 100 male white leghorn chickens, in groups of 5-26, at ages 4, 6, 8, 10, 12, and 14 weeks, were used as this is the age range in which it is expected that the testes are likely to be most "receptive." In this age range, the blood/testis barrier has not yet formed, and there is a relatively high number of spermatogonia relative to the numbers of other cell types, e.g., spermatids, etc. See J. Kumaran et al., 1949. Poultry Sci., vol. 29, pp. 511-520. See also E. Oakberg, 1956. Am. J. Anatomy, 15 vol. 99, pp. 507-515; and P. Kluin et al., 1984. Anat. Embryol., vol. 169, pp. 73-78.

The experimental and control males were obtained from commercial sources at one day of age, and maintained in brooders until used. The male birds were housed in temperature-controlled spaces in individual standard caging as they approached maturity. They were given water and standard commercial feed ad lib. They were 20 kept initially in a 23:1 hour light/dark cycle, stepped down at approximately weekly intervals to a 15:8 hour light/dark cycle, as this regimen has been reported to optimize sexual maturity and fertility.

Surgical and DNA Injection Procedures

25 At the appropriate ages, groups of individual males were starved overnight and then subjected to transgene delivery by direct intratesticular injection of DNA by experienced animal surgeons. Each male was anesthetized with isoflurane via a simplified gas machine.

Various devices and anesthesia machines have previously been described for administering isoflurane (and other gaseous anesthetics) to birds. See Alsage et al., 30 Poultry Sci., 50:1876-1878 (1971); Greenlees et al., Am. J. Vet. Res., vol. 51, pp. 757-758 (1990). However, these prior techniques are somewhat cumbersome and complex to implement. A novel and much simpler system to administer isoflurane (or other gaseous) anesthesia was developed due to the deficiencies in the prior art, a

system that we found worked well on all ages of chicks. A standard nose cone was placed over the chick's head, similar to the system that has been used for decades to administer ether to mice. A plastic tube approximately 3.5 cm in diameter and 12 cm long was filled with cotton, into which was poured approximately 2 mL isoflurane (Abbott Laboratories, Chicago). The chick's head was placed partially into the cylinder, and was held in place there intermittently throughout the surgery as required to maintain the proper plane of anesthesia, without overdosing.

Each anesthetized bird was positioned on its side on an animal board with cords tractioning the wings and feet to allow access to the testes area. The area was swabbed with 0.5% chlorhexidine, and a 2 cm dorsolateral incision was made in the skin over the testis (similar to the procedure commonly used for caponization). A small-animal retractor was used to spread the last two ribs, exposing the testis. The DNA solution was then mixed with SUPERFECT® (Qiagen) according to the manufacturer's protocol, approximately a 1:6 wt/vol ratio, to a final concentration of 0.01 - 0.05 µg/µl. This resulted in 1 - 5 µg total DNA (in a 150-200 µl volume) being injected into each testis, spread over three injection sites: one at each end of the testis, and one in the middle.

The injection device was a standard 25 gauge, 1/2 inch (1.27 cm) hypodermic needle, attached to a 50, 100, or 200 µl syringe. Approximately 5 mm of the needle tip was bent at a 90 degree angle, to facilitate insertion into the testes. Approximately 50 - 70 µl of the DNA-SUPERFECT® solution was injected into each of three sites per testis. The multiple injections were calculated to suffuse the DNA throughout the whole testis, the idea being to promote contact between DNA and spermatogonia as much as feasible. We estimated that our procedure resulted in the injection of about 100,000 DNA molecules per spermatogonium. The construct used in these tests was a highly potent constitutive modified CMV promoter, operatively linked to the dsRed gene as shown in SEQ ID NO:2.

Following injection, the incision was closed in two layers with 4-0 absorbable suture, and then the contralateral testis was similarly exposed and injected. Following surgery, each bird was returned to its cage to recover. One hundred thirteen males were ultimately used in the experimental regimen to increase the overall likelihood of success, along with 4 control birds (16 weeks 20 old) subjected to sham surgery (with injections containing only the transfection reagent).

Evaluation of Birds

Thus, a total of 113 white leghorn chickens were injected with the DNA vector in groups of 5-26 at varying ages. Fourteen birds were transformed at 4 weeks, 23 birds at 6 weeks; 26 birds at 8 weeks; 23 birds at 10 weeks; 5 birds at 12 weeks; and 22 birds at 14 weeks. Sixteen birds died before they could be sampled, so to date, 97 roosters have been sampled, plus the four controls. Birds were evaluated at 18-24 weeks of age for (a) potential transformation in the sperm, and (b) successful testis transfection. Sperm samples were obtained from each rooster by manual manipulation using standard techniques. The sperm were washed, and their DNA was extracted following the techniques of G. Mann et al., 1993. J. Reprod. Fert., 99:505-12. The samples were then frozen until analyzed. Evaluation was conducted by PCR analysis to detect DNA integration into the sperm, or into any of the testicular cells. Additionally, selected testes were harvested at the end of the sperm sampling period.

Of 97 birds tested, at least 22 showed probable positive results. Positive results were observed at all transformation ages, except for 4 weeks, which was not tested. At least two birds were confirmed positive by PCR of sperm, conducted four months after the initial injection. These results were transient in many cases, however since it was believed that the DsRed gene product used in these initial proof of concept experiments was toxic. Nevertheless, the positive PCR results presumptively demonstrated that the transgene was incorporated into spermatogonia (before puberty), and that it was carried in transgenic sperm. Such sperm could then transmit the gene to subsequent generations, resulting in the production of true, germ-line transgenic "founder" birds.

To further confirm that the DNA had been incorporated into the sperm, and that contaminating vector was not being detected from other sources, it was confirmed through PCR on sperm of experimental birds, and on positive and negative controls that the sperm of the experimental birds lacked DNA encoding the transposase. The design of the preferred transposon-based vector is such that the sequence encoding the transposase is contained in the vector, but is not incorporated into the transformed chromosome. Thus, presence of the exogenous coding sequence, coupled with absence of the transposase gene, is strong evidence for incorporation of the exogenous coding sequence, or transgene.

These results demonstrated proof of concept, as positive PCR results were obtained from the sperm of treated birds. Interpretation of these preliminary results was made more difficult by the fact that the modified CMV promoter used in the experiment was probably too "hot." As the DsRed product is not secreted from the cells, the product built up intracellularly to levels that were toxic, frequently killing the cells. Even this result, of course, means that the transformation was successful. The transgene could not have killed the cells otherwise.

In order to resolve to the problem with toxicity of the DsRed gene product, experiments were conducted using a different reporter gene operably linked to the ovalbumin promoter, so that the transgene was expressed in the egg white. These experiments are provided in Examples 12-15 below.

EXAMPLE 7

Transfection of Male White Leghorn Chickens Using the Vector pTnMod(Oval/Red) – Quail (SEQ ID NO:4) via Testicular Injections

In further experiments conducted on leghorn chickens, it was demonstrated that chickens injected intratesticularly at 8, 10, 12, or 14 weeks of age, had, on average, approximately 40% positive sperm between 6 and 8 months after injection. In other experiments, successful transfection was achieved with chickens injected at 13 weeks of age.

Forty-nine white leghorn roosters approximately 8, 10, 12, or 14 weeks of age were obtained and housed. Birds were identified, wing banded, and assigned to a treatment group. If appropriate (based on testes size and vascularization), one testis was caponized and the entire DNA injection volume was delivered to the remaining testis. Thirty-two males received DNA injections of 5µg DNA/testis at a 1:3 ratio of DNA to SUPERFECT®. The remaining birds were used as controls. After injection, all birds were mated with at least 5 females and observed until sexual maturity and egg-laying began. All eggs collected prior to peak egg production (approximately 24 weeks of age for the hens) were incubated and candled to determine embryo presence. Any embryos identified were incubated to hatch to extract DNA, PCR was conducted, and transgene presence was determined.

Roosters positive for the pTnMod(Oval/Red) – Quail construct were kept to produce F1 offspring (eggs collected at peak production). Offspring from this hatch

were bled, DNA extracted from the blood, and PCR conducted using primers specific for the DsRed gene. It was determined that 77% of the offspring were transgenic.

EXAMPLE 8

5 *Transfection of Mature Male Japanese Quail using the vector pTnMod(Oval/Red) – Quail (SEQ ID NO:4) via Testicular Injections*

Twelve sexually mature males (at approximately 13 weeks of age) underwent surgery for testicular injection as described above for chickens. At 21-28 days of age, the birds were identified, leg banded, debeaked, and separated based on sex. Injections comprised 5 µg/testes of the vector in concentrations 1:3 or 1:10 for SUPERFECT® or a 1:1 ratio with Mirrus. The study consisted of 3 treatment groups with 5 males in the 1:3 DNA:SUPERFECT® group, 3 males in the 1:10 DNA:SUPERFECT® group, and 4 males in the 1:1 Mirrus group. All surgeries were conducted in one day.

15 Any unincorporated DNA was allowed to clear from the testes by holding the birds for 19 days before mating with females. At 15 weeks of age, 2 age-matched females were housed with each treated male. The presence of the transfected DNA was determined in the fertilized eggs during the second week of egg lay. The subsequent eggs collected from parents producing positively identified transgenic eggs were collected and stored until taken to hatch.

PCR performed on the sperm of quail injected at three months of age indicated successful incorporation of the DsRed transgene into the quail sperm.

EXAMPLE 9

25 *Transfection of Immature Male Japanese Quail using the vector pTnMod(Oval/Red) – Quail (SEQ ID NO:4) via Testicular Injections*

Approximately 450 quail eggs were set and hatched. At 21-28 days of age, the birds were identified, wingbanded, debeaked, and separated based on sex. At 4 weeks of age, 65 male birds underwent surgery and testicular injections as described above. Injections comprised a control and 2 µg/testes of the vector in varying concentrations (0, 1/3, 1/5, and 1/10) of three different transfection reagents: 1) SUPERFECT®, 2) Mirus/Panvera and 3) Dosper. The study comprised 13 treatment groups with 5 males per group. One transfection reagent was administered per day.

At 7 weeks of age, 2 age-matched females were housed with each treated male. The presence of the transfected DNA was determined in the fertilized eggs during the second week of egg lay. The subsequent eggs collected from parents producing positively identified transgenic eggs were collected and stored until taken to hatch. PCR performed on the sperm of quail injected at four and five weeks of age indicated successful incorporation of the DsRed transgene into the quail sperm.

EXAMPLE 10

Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/p146/PA) – Chicken

A vector is designed for inserting a p146 gene under the control of a chicken ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:29.

Base pairs 1 - 130 are a remainder of F1(-) ori of pBluescriptII sk(-) (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

Base pairs 133 - 1777 are a CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

Base pairs 1780 - 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

Base pairs 2988-2993 are an engineered stop codon.

Base pairs 2995 - 3410 are a synthetic polyA from pGWiz (Gene Therapy Systems) corresponding to base pairs 1922- 2337 of pGWiz.

Base pairs 3415 - 3718 are non coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 - 3831 are the 70 base pairs of the left insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 3838 - 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

Base pairs 4050 - 4951 are a chicken ovalbumin promoter (including SDRE) that corresponds to base pairs 431-1332 of the chicken ovalbumin promoter in GenBank Accession Number J00895 M24999.

Base pairs 4958 - 6115 are a chicken ovalbumin signal sequence and Ovalbumin gene that correspond to base pairs 66-1223 of GenBank Accession Number V00383.1 (The STOP codon being omitted).

5 Base pairs 6122 - 6271 are a TAG sequence containing a gp41 hairpin loop from HIV I, an enterokinase cleavage site and a spacer (synthetic).

Base pairs 6272 - 6316 are a p146 sequence (synthetic) with 2 added stop codons.

Base pairs 6324 - 6676 are a synthetic polyadenylation sequence from pGWiz (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

10 Base pairs 6682 - 7114 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

Base pairs 7120- 7189 are the 70 base pairs of the right insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 7190 - 7231 are λ DNA that is residual from pNK2859.

15 Base pairs 7232 - 8096 are non coding DNA that is residual from pNK2859.

Base pairs 8097 - 10297 are pBlueScript sk(-) base vector (Stratagene, Inc.) corresponding to base pairs 761-2961 of pBluescriptII sk(-).

20 It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

EXAMPLE 11

Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/p146/PA) - Quail

25 A vector is designed for inserting a p146 gene under the control of a quail ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:30.

Base pairs 1 - 130 are a remainder of F1(-) ori of pBluescriptII sk(-) (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

30 Base pairs 133 - 1777 are a CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

Base pairs 1780 - 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

Base pairs 2988-2993 are an engineered stop codon.

Base pairs 2995 – 3410 are a synthetic polyA from pGWiz (Gene Therapy Systems) corresponding to base pairs 1922-2337 of pGWiz.

Base pairs 3415 – 3718 are non coding DNA that is residual from vector pNK2859.

5 Base pairs 3719 – 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 – 3831 are the 70 base pairs of the left insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 3838 – 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

10 Base pairs 4050 - 4938 are the Japanese quail ovalbumin promoter (including SDRE, steroid-dependent response element). The Japanese quail ovalbumin promoter was isolated by its high degree of homology to the chicken ovalbumin promoter (GenBank accession number J00895 M24999, base pairs 431-1332).

15 Bp 4945 - 6092 are a quail ovalbumin signal sequence and ovalbumin gene that corresponds to base pairs 54 – 1201 of GenBank accession number X53964.1. (The STOP codon being omitted).

Base pairs 6097 - 6246 are a TAG sequence containing a gp41 hairpin loop from HIV I, an enterokinase cleavage site and a spacer (synthetic).

20 Base pairs 6247 – 6291 are a p146 sequence (synthetic) with 2 added stop codons.

Base pairs 6299 – 6651 are a synthetic polyadenylation sequence from pGWiz (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

Base pairs 6657 - 7089 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

25 Base pairs 7095- 7164 are the 70 base pairs of the right insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 7165 - 7206 are λ DNA that is residual from pNK2859.

Base pairs 7207 – 8071 are non coding DNA that is residual from pNK2859.

30 Base pairs 8072 - 10272 are pBlueScript sk(-) base vector (Stratagene, Inc.) corresponding to base pairs 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

5 EXAMPLE 12

Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/ProIns/PA) – Chicken

A vector is designed for inserting a proinsulin gene under the control of a chicken ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal
10 sequence, into the genome of a bird given below as SEQ ID NO:31.

Base pairs 1 - 130 are a remainder of F1(-) ori of pBluescriptII sk(-) (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

Base pairs 133 - 1777 are a CMV promoter/enhancer taken from vector
15 pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

Base pairs 1780 - 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

Base pairs 2988-2993 are an engineered stop codon.

Base pairs 2995 - 3410 are a synthetic polyA from pGWiz (Gene Therapy
20 Systems) corresponding to base pairs 1922- 2337 of pGWiz.

Base pairs 3415 - 3718 are non coding DNA that is residual from vector pNK2859.

Base pairs 3719 - 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 - 3831 are the 70 base pairs of the left insertion sequence
25 (IS10) recognized by the transposon Tn10.

Base pairs 3838 - 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

Base pairs 4050 - 4951 are a chicken ovalbumin promoter (including SDRE) that corresponds to base pairs 431-1332 of the chicken ovalbumin promoter in
30 GenBank Accession Number J00895 M24999.

Base pairs 4958 - 6115 are a chicken ovalbumin signal sequence and ovalbumin gene that correspond to base pairs 66-1223 of GenBank Accession Number V00383.1. (The STOP codon being omitted).

Base pairs 6122 - 6271 are a TAG sequence containing a gp41 hairpin loop from HIV I, an enterokinase cleavage site and a spacer (synthetic).

Base pairs 6272 – 6531 are a proinsulin gene.

Base pairs 6539 – 6891 are a synthetic polyadenylation sequence from pGWiz (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

Base pairs 6897 - 7329 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

Base pairs 7335- 7404 are the 70 base pairs of the right insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 7405 - 7446 are λ DNA that is residual from pNK2859.

Base pairs 7447 – 8311 are non coding DNA that is residual from pNK2859.

Base pairs 8312 - 10512 are pBlueScript sk(-) base vector (Stratagene, Inc.) corresponding to base pairs 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

EXAMPLE 13

Preparation of Transposon-Based Vector pTnMod(Oval/ENT TAG/ProIns/PA) – Quail

A vector is designed for inserting a proinsulin gene under the control of a chicken ovalbumin promoter, and a ovalbumin gene including an ovalbumin signal sequence, into the genome of a bird given below as SEQ ID NO:32.

Base pairs 1 -130 are a remainder of F1(-) ori of pBluescriptII sk(-) (Stratagene) corresponding to base pairs 1-130 of pBluescriptII sk(-).

Base pairs 133 – 1777 are a CMV promoter/enhancer taken from vector pGWiz (Gene Therapy Systems) corresponding to base pairs 229-1873 of pGWiz.

Base pairs 1780 – 2987 are a transposase, modified from Tn10 (GenBank accession number J01829).

Base pairs 2988-2993 are an engineered stop codon.

Base pairs 2995 – 3410 are a synthetic polyA from pGWiz (Gene Therapy Systems) corresponding to base pairs 1922- 2337 of pGWiz.

Base pairs 3415 – 3718 are non coding DNA that is residual from vector pNK2859.

5 Base pairs 3719 – 3761 are λ DNA that is residual from pNK2859.

Base pairs 3762 – 3831 are the 70 base pairs of the left insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 3838 – 4044 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 924-718 of pBluescriptII sk(-).

10 Base pairs 4050 - 4938 are the Japanese quail ovalbumin promoter (including SDRE, steroid-dependent response element). The Japanese quail ovalbumin promoter was isolated by its high degree of homology to the chicken ovalbumin promoter (GenBank accession number J00895 M24999, base pairs 431-1332). Some deletions were noted in the quail sequence, as compared to the chicken sequence.

15 Base pairs 4945 - 6092 are a quail ovalbumin signal sequence and ovalbumin gene that corresponds to base pairs 54 – 1201 of GenBank accession number X53964.1. (The STOP codon being omitted).

Base pairs 6093 - 6246 are a TAG sequence containing a gp41 hairpin loop from HIV I an enterokinase cleavage site and a spacer (synthetic).

20 Base pairs 6247 – 6507 are a proinsulin gene.

Base pairs 6514 – 6866 are a synthetic polyadenylation sequence from pGWiz (Gene Therapy Systems) corresponding to base pairs 1920 - 2272 of pGWiz.

Base pairs 6867 - 7303 are a multiple cloning site from pBlueScriptII sk(-) corresponding to base pairs 667-235 of pBluescriptII sk(-).

25 Base pairs 7304- 7379 are the 70 base pairs of the right insertion sequence (IS10) recognized by the transposon Tn10.

Base pairs 7380 - 7421 are λ DNA that is residual from pNK2859.

Base pairs 7422 – 8286 are non coding DNA that is residual from pNK2859.

Base pairs 8287 - 10487 are pBlueScript sk(-) base vector (Stratagene, Inc.)

30 corresponding to base pairs 761-2961 of pBluescriptII sk(-).

It should be noted that all non-coding DNA sequences described above can be replaced with any other non-coding DNA sequence(s). Missing nucleotide sequences in the above construct represent restriction site remnants.

5 EXAMPLE 14

Transfection of Immature Leghorn Roosters using a Transposon-based Vector containing a Proinsulin Gene via Testicular Injections

Vectors containing the clements Oval promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A (SEQ ID NO:31) and CMV promoter/Oval gene/GP41
10 Enterokinase TAG/Proinsulin/Poly A (SEQ ID NO:42) were each injected into the testes of 11 week old white leghorn roosters. These birds were held under normal conditions until sexual maturity was reached.

At the time of sexual maturity, each bird was handled and manipulated to obtain sperm. Sperm samples were collected in Hank's Buffered Salt Solution
15 (HBSS) and stored at either -20° C or 4° C until needed. DNA was extracted from sperm using a MoBio Ultra Clean DNA Bloodspin Kit (MoBio laboratories, Solana Beach CA). Fifty microliters of sperm was used in the DNA extraction protocol and the purified genomic DNA eluted in 100 µl of water. In each PCR reaction, approximately 0.5 – 0.75 µg of genomic DNA was used with primers anchored in the
20 entag-1 (5') and the synthetic polyA-2 (3'), which amplify a 685 bp fragment. Five of nine birds gave positive reactions for the presence of the appropriate vector construct. These birds were then mated with normal females.

Birds that did not yield positive results with PCR on the sperm were sacrificed, their testes removed, and DNA extracted using an approximately 25 mg
25 piece of tissue in a Qiagen DNEasy Tissue Kit; purified DNA was eluted in 200 µl water and PCR conducted as described above. Two of these birds gave a very strong, positive PCR reaction.

EXAMPLE 15

30 *Transfection of Japanese Quail using a Transposon-based Vector containing a Proinsulin Gene via Oviduct Injections*

Two experiments were conducted in Japanese quail using transposon-based vectors containing either Oval promoter/Oval gene/GP41 Enterokinase

TAG/Proinsulin/Poly A (SEQ ID NO:31) or CMV promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A (SEQ ID NO:42).

5 In the first experiment, the Oval promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A containing construct was injected into the oviduct of sexually mature quail; three hens received 5 µg at a 1:3 Superfect ratio and three received 10 µg at a 1:3 Superfect ratio. As of the writing of the present application, at least one bird that received 10 µg of DNA was producing human proinsulin in egg white (other birds remain to be tested). This experiment indicates that 1) the DNA has been stable for at least 3 months; 2) protein levels are comparable to those observed with a constitutive promoter such as the CMV promoter; and 3) sexually mature birds can be injected and results obtained without the need for cell culture.

10 In the second experiment, the transposon-based vector containing CMV promoter/Oval gene/GP41 Enterokinase TAG/Proinsulin/Poly A was injected into the oviduct of sexually immature Japanese quail. A total of 9 birds were injected. Of the 8 survivors, 3 produced human proinsulin in the white of their eggs for over 6 weeks. An ELISA assay described in detail below was developed to detect GP41 in the fusion peptide (Oval gene/GP41 Enterokinase TAG/Proinsulin) since the GP41 peptide sequence is unique and not found as part of normal egg white protein. In all ELISA assays, the same birds produced positive results and all controls worked as expected.

20 ELISA Procedure: Individual egg white samples were diluted in sodium carbonate buffer, pH 9.6, and added to individual wells of 96 well microtiter ELISA plates at a total volume of 0.1 ml. These plates were then allowed to coat overnight at 4°C. Prior to ELISA development, the plates were allowed warm to room temperature. Upon decanting the coating solutions and blotting away any excess, non-specific binding of antibodies was blocked by adding a solution of phosphate buffered saline (PBS), 1% (w/v) BSA, and 0.05% (v/v) Tween 20 and allowing it to incubate with shaking for a minimum of 45 minutes. This blocking solution was subsequently decanted and replaced with a solution of the primary antibody (Goat Anti-GP41 TAG) diluted in fresh PBS/BSA/Tween 20. After a two hour period of incubation with the primary antibody, each plate was washed with a solution of PBS and 0.05% Tween 20 in an automated plate washer to remove unbound antibody. Next, the secondary antibody, Rabbit anti-Goat Alkaline Phosphatase-conjugated, was diluted in PBS/BSA/Tween 20 and allowed to incubate 1 hour. The plates were then

subjected to a second wash with PBS/Tween 20. Antigen was detected using a solution of *p*-Nitrophenyl Phosphate in Diethanolamine Substrate Buffer for Alkaline Phosphatase and measuring the absorbance at 30 minutes and 1 hour.

5 EXAMPLE 16

Optimization of Intra-oviduct and Intra-ovarian Arterial Injections

Overall transfection rates of oviduct cells in a flock of chicken or quail hens are enhanced by synchronizing the development of the oviduct and ovary within the flock. When the development of the oviducts and ovaries are uniform across a group
10 of hens and when the stage of oviduct and ovarian development can be determined or predicted, timing of injections is optimized to transfect the greatest number of cells. Accordingly, oviduct development is synchronized as described below to ensure that a large and uniform proportion of oviduct secretory cells are transfected with the gene of interest.

15 Hens are treated with estradiol to stimulate oviduct maturation as described in Oka and Schimke (T. Oka and RT Schimke, J. Cell Biol., 41, 816 (1969)), Palmiter, Christensen and Schimke (J Biol. Chem. 245(4):833-845, 1970). Specifically, repeated daily injections of 1 mg estradiol benzoate are performed sometime before the onset of sexual maturation, a period ranging from 1 – 14 weeks of age. After a
20 stimulation period sufficient to maximize development of the oviduct, hormone treatment is withdrawn thereby causing regression in oviduct secretory cell size but not cell number. At an optimum time after hormone withdrawal, the oviducts of treated hens are injected with the transposon-based vector. Hens are subjected to additional estrogen stimulation after an optimized time during which the transposon-
25 based vector is taken up into oviduct secretory cells. Re-stimulation by estrogen activates the transposon mechanism of the transposon-based vector, causing the integration of the gene of interest into the host genome. Estrogen stimulation is then withdrawn and hens continue normal sexual development. If a developmentally regulated promoter such as the ovalbumin promoter is used, expression of the
30 transposon-based vector initiates in the oviduct at the time of sexual maturation. Intra-ovarian artery injection during this window allows for high and uniform transfection efficiencies of ovarian follicles to produce germ-line transfections and possibly oviduct expression.

Other means are also used to synchronize the development, or regression, of the oviduct and ovary to allow high and uniform transfection efficiencies. Alterations of lighting and/or feed regimens, for example, cause hens to 'molt' during which time the oviduct and ovary regress. Molting is used to synchronize hens for transfection, and may be used in conjunction with other hormonal methods to control regression and/or development of the oviduct and ovary.

EXAMPLE 17

Isolation of Human Proinsulin Using Anti-TAG Column Chromatography

A HiTrap NHS-activated 1 mL column (Amersham) was charged with a 30 amino acid peptide that contained the gp-41 epitope containing gp-41's native disulfide bond that stabilizes the formation of the gp-41 hairpin loop. The 30 amino acid gp41 peptide is provided as SEQ ID NO:23. Approximately 10 mg of the peptide was dissolved in coupling buffer (0.2 M NaHCO₃, 0.5 M NaCl, pH 8.3 and the ligand was circulated on the column for 2 hours at room temperature at 0.5 mL/minute. Excess active groups were then deactivated using 6 column volumes of 0.5 M ethanolamine, 0.5 M NaCl, pH 8.3 and the column was washed alternately with 6 column volumes of acetate buffer (0.1 M acetate, 0.5 M NaCl, pH 4.0) and ethanolamine (above). The column was neutralized using 1 X PBS. The column was then washed with buffers to be used in affinity purification: 75 mM Tris, pH 8.0 and elution buffer, 100 mM glycine-HCl, 0.5 M NaCl, pH 2.7. Finally, the column was equilibrated in 75 mM Tris buffer, pH 8.0.

Antibodies to gp-41 were raised in goats by inoculation with the gp-41 peptide described above. More specifically, goats were inoculated, given a booster injection of the gp-41 peptide and then bled. Serum was harvested by centrifugation. Approximately 30 mL of goat serum was filtered to 0.45 µM and passed over a TAG column at a rate of 0.5 mL/min. The column was washed with 75 mM Tris, pH 8.0 until absorbance at 280 nm reached a baseline. Three column volumes (3 mL) of elution buffer (100 mM glycine, 0.5 M NaCl, pH 2.7) was applied, followed by 75 mM Tris buffer, pH 8.0, all at a rate of 0.5 mL/min. One milliliter fractions were collected. Fractions were collected into 200 µL 1 M Tris, pH 9.0 to neutralize acidic fractions as rapidly as possible. A large peak eluted from the column, coincident with the application the elution buffer. Fractions were pooled. Analysis by SDS-PAGE

showed a high molecular weight species that separated into two fragments under reducing condition, in keeping with the heavy and light chain structure of IgG.

Pooled antibody fractions were used to charge two 1 mL HiTrap NHS-activated columns, attached in series. Coupling was carried out in the same manner as that used for charging the TAG column.

Isolation of Ovalbumin-TAG-Proinsulin from Egg White

Egg white from quail and chickens treated by intra-oviduct injection of the CMV-ovalbumin-TAG-proinsulin construct were pooled. Viscosity was lowered by subjecting the allantoic fluid to successively finer pore sizes using negative pressure filtration, finishing with a 0.22 μ M pore size. Through the process, egg white was diluted approximately 1:16. The clarified sample was loaded on the Anti-TAG column and eluted in the same manner as described for the purification of the anti-TAG antibodies. A peak of absorbance at 280 nm, coincident with the application of the elution buffer, indicated that protein had been specifically eluted from the Anti-TAG column. Fractions containing the eluted peak were pooled for analysis.

The pooled fractions from the Anti-TAG affinity column were characterized by SDS-PAGE and western blot analysis. SDS-PAGE of the pooled fractions revealed a 60 kDal molecular weight band not present in control egg white fluid, consistent with the predicted molecular weight of the transgenic protein. Although some contaminating bands were observed, the 60 kDal species was greatly enriched compared to the other proteins. An aliquot of the pooled fractions was cleaved overnight at room temperature with the protease, enterokinase. SDS-PAGE analysis of the cleavage product, revealed a band not present in the uncut material that comigrated with a commercial human proinsulin positive control. Western blot analysis showed specific binding to the 60 kDal species under non-reducing condition (which preserve the hairpin epitope of gp-41 by retaining the disulfide bond). Western analysis of the low molecular weight species that appeared upon cleavage with an anti-human proinsulin antibody, conclusively identified the cleaved fragment as human proinsulin.

EXAMPLE 18

Construction of a Transposon-based Transgene for the Expression of a Monoclonal Antibody

5 Production of a monoclonal antibody using transposon-based transgenic methodology is accomplished in a variety of ways.

1) two vectors are constructed: one that encodes the light chain and a second vector that encodes the heavy chain of the monoclonal antibody. These vectors are then incorporated into the genome of the target animal by at least one of two methods: a) direct transfection of a single animal with both vectors (simultaneously or as separate
10 events); or, b) a male and a female of the species carry in their germline one of the vectors and then they are mated to produce progeny that inherit a copy of each.

2) the light and heavy chains are included on a single DNA construct, either separated by insulators and expression is governed by the same (or different) promoters, or by using a single promoter governing expression of both transgenes with the inclusion of
15 elements that permit separate transcription of both transgenes, such as an internal ribosome entry site.

The following example describes the production of a transposon-based DNA construct that contains both the coding region for a monoclonal light chain and a heavy chain on a single construct. Beginning with the vector pTnMod, the coding
20 sequences for the heavy and light chains are added, each preceded by an appropriate promoter and signal sequence. Using methods known to one skilled in the art, approximately 1 Kb of the proximal elements of the ovalbumin promoter are linked to the signal sequence of ovalbumin or some other protein secreted from the target tissue. Two copies of the promoter and signal sequence are added to the multiple
25 cloning site of pTnMod, leaving space and key restriction sites between them to allow the subsequent addition of the coding sequences of the light and heavy chains of the monoclonal antibody. Methods known to one skilled in the art allow the coding sequences of the light and heavy chains to be inserted in-frame for appropriate expression. For example, the coding sequence of light and heavy chains of a murine
30 monoclonal antibody that show specificity for human seminoprotein have recently been disclosed (GenBank Accession numbers AY129006 and AY129304 for the light and heavy chains, respectively). The light chain cDNA sequence is provided in SEQ

ID NO:34, whereas the cDNA of the heavy chain is reported as provided in SEQ ID NO:35.

Thus one skilled in the art can produce both the heavy and light chains of a monoclonal antibody in a single cell within a target tissue and species. If the modified cell contained normal posttranslational modification capabilities, the two chains would form their native configuration and disulfide attachments and be substrates for glycosylation. Upon secretion, then, the monoclonal antibody is accumulated, for example, in the egg white of a chicken egg, if the transgenes are expressed in the magnum of the oviduct.

It should also be noted that, although this example details production of a full-length murine monoclonal antibody, the method is quite capable of producing hybrid antibodies (e.g. a combination of human and murine sequences; 'humanized' monoclonal antibodies), as well as useful antibody fragments, known to one skilled in the art, such as Fab, Fc, F(ab) and Fv fragments. This method can be used to produce molecules containing the specific areas thought to be the antigen recognition sequences of antibodies (complementarity determining regions), linked, modified or incorporated into other proteins as desired.

EXAMPLE 19

Treatment of rats with a transposon-based vector for tissue-specific insulin gene incorporation

Rats are made diabetic by administering the drug streptozotocin (Zanosar; Upjohn, Kalamazoo, MI) at approximately 200 mg/kg. The rats are bred and maintained according to standard procedures. A transposon-based vector containing a proinsulin gene, an appropriate carrier, and, optionally, a transfection agent, are injected into rats' singhepatic (if using G6P) artery with the purpose of stable transformation. Incorporation of the insulin gene into the rat genome and levels of insulin expression are ascertained by a variety of methods known in the art. Blood and tissue samples from live or sacrificed animals are tested. A combination of PCR, Southern and Northern blots, *in-situ* hybridization and related nucleic acid analysis methods are used to determine incorporation of the vector-derived proinsulin DNA and levels of transcription of the corresponding mRNA in various organs and tissues of the rats. A combination of SDS-PAGE gels, Western Blot analysis, radioimmunoassay, and ELISA and other methods known to one of ordinary skill in

the art are used to determine the presence of insulin and the amount produced. Additional transfections of the vector are used to increase protein expression if the initial amounts of the expressed insulin are not satisfactory, or if the level of expression tapers off. The physiological condition of the rats is closely examined

5 post-transfection to register positive or any negative effects of the gene therapy. Animals are examined over extended periods of time post-transfection in order to monitor the stability of gene incorporation and protein expression.

EXAMPLE 20

10

Exemplary Transposon-Based Vectors

The following example provides a description of various transposon-based vectors of the present invention and several constructs for insertion into the transposon-based vectors of the present invention. These examples are not meant to

15 be limiting in any way. The constructs for insertion into a transposon-based vector are provided in a cloning vector labeled pTnMCS.

pTnMCS (base vector)

- Bp 1 – 130 Remainder of F1 (-) ori of pBluescriptII sk(-) (Stratagene) bp1-130
- 20 Bp 133 – 1777 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems) bp2 29-1873
- Bp 1783 – 2991 Transposase, from Tn10 (GenBank accession #J01829) bp 108-1316
- Bp 2992 – 3344 Non coding DNA from vector pNK2859
- Bp 3345 – 3387 Lambda DNA from pNK2859
- 25 Bp 3388 – 3457 70 bp of IS10 left from Tn10
- Bp 3464 – 3670 Multiple cloning site from pBluescriptII sk(-), thru the XmaI site bp924-718
- Bp 3671 - 3715 Multiple cloning site from pBluescriptII sk(-), from the XmaI site thru the XhoI site. These base pairs are usually lost when cloning into pTnMCS.bp
- 30 717-673
- Bp 3716 – 4153 Multiple cloning site from pBluescriptII sk(-), from the XhoI site bp672-235
- Bp 4159 - 4228 70 bp of IS10 left from Tn10
- Bp 4229 - 4270 Lambda DNA from pNK2859

Bp 4271 – 5114 Non-coding DNA from pNK2859

Bp 5115 - 7315 pBluescript sk (-) base vector (Stratagene, Inc.) bp 761-2961

pTnMCS (CMV-prepro-ent-hGH-CPA)

- 5 Bp 1 – 3670 from vector PTnMCS, bp 1 - 3670
- Bp 3676 – 5320 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems), bp 230-1864
- Bp 5326 - 5496 Capsite/Prepro taken from GenBank accession # X07404, bp 563 – 733
- 10 Bp 5504 - 5652 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 5653 – 6306 Human growth hormone taken from GenBank accession # V00519, bp 1-654
- Bp 6313 – 6720 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651-11058
- 15 Bp 6722 – 10321 from cloning vector pTnMCS, bp 3716-7315

pTnMCS (CMV-CHOVg-ent-ProInsulin-synPA) (SEQ ID NO:41)

- Bp 1 – 3670 from vector PTnMCS, bp 1 - 3670
- 20 Bp 3676 – 5320 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems), bp 230-1864
- Bp 5327 – 6480 Chicken ovalbumin gene taken from GenBank accession # V00383, bp 66-1219
- Bp 6487 - 6636 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- 25 Bp 6637 – 6897 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377
- Bp 6898 – 6942 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)
- 30 Bp 6943 - 7295 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271
- Bp 7296 – 10895 from cloning vector pTnMCS, bp 3716-7315

pTnMCS (CMV-prepro-ent-ProInsulin-synPA)

Bp 1 – 3670 from vector pTnMCS, bp 1 - 3670

Bp 3676 – 5320 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems), bp 230-1864

5 Bp 5326 - 5496 Capsite/Prepro taken from GenBank accession # X07404, bp 563 – 733

Bp 5504 - 5652 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

10 Bp 5653 – 5913 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377

Bp 5914 – 5958 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)

Bp 5959 - 6310 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271

15 Bp 6313 – 9912 from cloning vector pTnMCS, bp 3716-7315

pTnMCS(Chicken OVEp+OVg'+ENT+proinst+syn polyA)

Bp 1 – 3670 from vector pTnMCS, bp 1 - 3670

20 Bp 3676 – 4350 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1 bp 1-675

Bp 4357 – 5692 Chicken Ovalbumin promoter taken from GenBank accession # J00895M24999 bp 1-1336

25 Bp 5699 – 6917 Chicken Ovalbumin gene from GenBank Accession # V00383.1 bp 2-1220. (This sequence includes the 5'UTR, containing putative cap site, bp 5699-5762.)

Bp 6924 - 7073 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 7074 - 7334 Human proinsulin GenBank Accession # NM000207 bp 117-377

30 Bp 7335 - 7379 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

Bp 7380 - 7731 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 7733 – 11332 from vector pTnMCS, bp 3716 - 7315

pTnMCS(Chicken OVep+prepro+ENT+proins+syn polyA)

- Bp 1 – 3670 from cloning vector pTnMCS, bp 1 - 3670
- Bp 3676 – 4350 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1 bp 1-675
- 5 Bp 4357 – 5692 Chicken Ovalbumin promoter taken from GenBank accession # J00895-M24999 bp 1-1336
- Bp 5699 – 5869 Cecropin cap site and Prcpro, Genbank accession # X07404 bp 563-733
- 10 Bp 5876 - 6025 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 6026 - 6286 Human proinsulin GenBank Accession # NM000207 bp 117-377
- Bp 6287 - 6331 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- 15 Bp 6332 - 6683 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- Bp 6685 – 10284 from cloning vector pTnMCS, bp 3716 - 7315

pTnMCS(Quail OVep+OVg'+ENT+proins+syn polyA)

- 20 Bp 1 – 3670 from cloning vector pTnMCS, bp 1 - 3670
- Bp 3676 – 4333 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house from quail genomic DNA, roughly equivalent to the far-upstream chicken ovalbumin enhancer, GenBank accession # S82527.1, bp 1-675. (There are multiple base pair substitutions and deletions in the quail sequence, relative to chicken, so the number of
- 25 bases does not correspond exactly.)
- Bp 4340 – 5705 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter, GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair substitutions and deletions between the quail and chicken sequences, so the number of
- 30 bases does not correspond exactly.)
- Bp 5712 – 6910 Quail Ovalbumin gene, EMBL accession # X53964, bp 1-1199. (This sequence includes the 5'UTR, containing putative cap site bp 5712-5764.)

- Bp 6917 - 7066 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 7067 - 7327 Human proinsulin GenBank Accession # NM000207 bp 117-377
- Bp 7328 - 7372 Spacer DNA, derived as an artifact from the cloning vectors pTOPO
- 5 Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- Bp 7373 - 7724 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- Bp 7726 - 11325 from cloning vector pTnMCS, bp 3716 - 7315
- 10 pTnMCS (CHOVep-prepro-ent-hGH-CPA)
- Bp 1 - 3670 from vector PTnMCS, bp 1-3670
- Bp 3676 - 4350 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1, bp 1-675
- Bp 4357 - 5692 Chicken Ovalbumin promoter taken from GenBank accession # J00899-M24999, bp 1-1336
- 15 Bp 5699 - 5869 Capsite/Prepro taken from GenBank accession # X07404, bp 563-733
- Bp 5877 - 6025 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 6026 - 6679 Human growth hormone taken from GenBank accession # V00519,
- 20 bp 1-654
- Bp 6686 - 7093 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651-11058
- Bp 7095 - 10694 from cloning vector pTnMCS, bp 3716-7315
- 25 pTnMCS(Quail OVep+prepro+ENT+proins+syn polyA)
- Bp 1 - 3670 from cloning vector pTnMCS, bp 1 - 3670
- Bp 3676 - 4333 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house from quail genomic DNA, roughly equivalent to the far- upstream chicken ovalbumin enhancer, GenBank accession #S82527.1, bp 1-675. (There are multiple base pair
- 30 substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)
- Bp 4340 - 5705 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter,

- GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)
- Bp 5712 – 5882 Cecropin cap site and Prepro, Genbank accession # X07404 bp 563-733
- Bp 5889 - 6038 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 6039 - 6299 Human proinsulin GenBank Accession # NM000207 bp 117-377
- Bp 6300 - 6344 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- Bp 6345 - 6696 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- Bp 6698 – 10297 from cloning vector pTnMCS, bp 3716 - 7315
- 15 PTnMOD
- Bp 1 - 130 remainder of F1 (-) ori of pBluescriptII sk(-) (Stratagene) bp1-130
- Bp 133 - 1777 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems) bp229-1873
- Bp 1783 - 2991 Transposase, modified from Tn10 (GenBank accession #J01829) bp 108-1316
- 20 Bp 2992 - 2994 Engineered stop codon
- Bp 2996 - 3411 Synthetic polyA from gWIZ (Gene Therapy Systems) bp 1922 - 2337
- Bp 3412 - 3719 Non-coding DNA from vector pNK2859
- Bp 3720 - 3762 Lambda DNA from pNK2859
- 25 Bp 3763 - 3832 70 bp of IS10 left from Tn10
- Bp 3839 – 4045 Multiple cloning site from pBluescriptII sk(-), thru the XmaI site bp 924-718
- Bp 4046 - 4090 Multiple cloning site from pBluescriptII sk(-), from the XmaI site thru the XhoI site. These base pairs are usually lost when cloning into pTnMCS. bp 717-673
- 30 Bp 4091 – 4528 Multiple cloning site from pBluescriptII sk(-), from the XhoI site bp 672-235

- Bp 4534 - 4603 70 bp of IS10 left from Tn10
 Bp 4604 - 4645 Lambda DNA from pNK2859
 Bp 4646 - 5489 Non-coding DNA from pNK2859
 Bp 5490 - 7690 pBluescript sk (-) base vector (Stratagene, INC) bp 761-2961
- 5 pTnMOD (CHOVep-prepro-ent-hGH-CPA)
 Bp 1 - 4045 from vector PTnMCS, bp 1 - 4045
 Bp 4051 - 4725 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1, bp 1 - 675
- 10 Bp 4732 - 6067 Chicken Ovalbumin promoter taken from GenBank accession # J00899-M24999, bp 1-1336
 Bp 6074 - 6245 Capsite/Prepro taken from GenBank accession # X07404, bp 563 - 733
 Bp 6252 - 6400 Synthetic spacer sequence and hairpin loop of HIV gp41 with an
- 15 added enterokinase cleavage site
 Bp 6401 - 7054 Human growth hormone taken from GenBank accession # V00519, bp 1-654
 Bp 7061 - 7468 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651-11058
- 20 Bp 7470 - 11069 from cloning vector pTnMCS, bp 3716-7315
- pTnMOD (CMV-CHOVg-ent-ProInsulin-synPA) (SEQ ID NO:42)
 Bp 1 - 4045 from vector PTnMCS, bp 1 - 4045
 Bp 4051 - 5695 CMV promoter/enhancer taken from vector pGWIZ (Gene therapy
- 25 systems), bp 230-1864
 Bp 5702 - 6855 Chicken ovalbumin gene taken from GenBank accession # V00383, bp 66-1219
 Bp 6862 - 7011 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- 30 Bp 7012 - 7272 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377
 Bp 7273 - 7317 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)

Bp 7318 - 7670 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271

Bp 7672 - 11271 from cloning vector pTnMCS, bp 3716-7315

5 pTnMOD (CMV-prepro-ent-hGH-CPA)

Bp 1 - 4045 from vector PTnMCS, bp 1 - 4045

Bp 4051 - 5695 CMV promoter/enhancer taken from vector pGWIZ (Gene therapy systems), bp 230-1864

Bp 5701 - 5871 Capsite/Prepro taken from GenBank accession # X07404, bp 563 - 733

Bp 5879 - 6027 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 6028 - 6681 Human growth hormone taken from GenBank accession # V00519, bp 1-654

15 Bp 6688 - 7095 Conalbumin polyA taken from GenBank accession # Y00407, bp 10651-11058

Bp 7097 - 10696 from cloning vector pTnMCS, bp 3716-7315

pTnMOD (CMV-prepro-ent-ProInsulin-synPA)

20 Bp 1 - 4045 from vector PTnMCS, bp 1 - 4045

Bp 4051 - 5695 CMV promoter/enhancer taken from vector pGWIZ (Gene therapy systems), bp 230-1864

Bp 5701 - 5871 Capsite/Prepro taken from GenBank accession # X07404, bp 563 - 733

25 Bp 5879 - 6027 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 6028 - 6288 Human Proinsulin taken from GenBank accession # NM000207, bp 117-377

Bp 6289 - 6333 Spacer DNA, derived as an artifact from the cloning vectors pTOPO

30 Blunt II (Invitrogen) and pGWIZ (Gene Therapy Systems)

Bp 6334 - 6685 Synthetic polyA from the cloning vector pGWIZ (Gene Therapy Systems), bp 1920-2271

Bp 6687 - 10286 from cloning vector pTnMCS, bp 3716-7315

- pTnMOD(Chicken OVep+OVg'+ENT+proins+syn polyA) (SEQ ID NO:43)
 Bp 1 – 4045 from cloning vector pTnMOD, bp 1 - 4045
 Bp 4051 – 4725 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1 bp 1-675
 5 Bp 4732 – 6067 Chicken Ovalbumin promoter taken from GenBank accession # J00895-M24999 bp 1-1336
 Bp 6074 – 7292 Chicken Ovalbumin gene from GenBank Accession # V00383.1 bp 2-1220. (This sequence includes the 5'UTR, containing putative cap site bp 6074-6137.)
 10 Bp 7299 - 7448 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
 Bp 7449 - 7709 Human proinsulin GenBank Accession # NM000207 bp 117-377
 Bp 7710 - 7754 Spacer DNA, derived as an artifact from the cloning vectors pTOPO
 15 Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
 Bp 7755 - 8106 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
 Bp 8108 – 11707 from cloning vector pTnMCS, bp 3716 - 7315
- 20 pTnMOD(Chicken OVep+prepro+ENT+proins+syn polyA)
 Bp 1 – 4045 from cloning vector pTnMCS, bp 1 - 4045
 Bp 4051 – 4725 Chicken Ovalbumin enhancer taken from GenBank accession # S82527.1 bp 1-675
 Bp 4732 – 6067 Chicken Ovalbumin promoter taken from GenBank accession # J00895-M24999 bp 1-1336
 25 Bp 6074 – 6244 Cecropin cap site and Prepro, Genbank accession # X07404 bp 563-733
 Bp 6251 - 6400 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
 30 Bp 6401 - 6661 Human proinsulin GenBank Accession # NM000207 bp 117-377
 Bp 6662 - 6706 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

Bp 6707 - 7058 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 7060 - 10659 from cloning vector pTnMCS, bp 3716 - 7315

5 pTnMOD(Quail OVep+OVg'+ENT+proins+syn polyA)

Bp 1 - 4045 from cloning vector pTnMCS, bp 1 - 4045

Bp 4051 - 4708 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house from quail genomic DNA, roughly equivalent to the far-upstream chicken ovalbumin enhancer, GenBank accession # S82527.1, bp 1-675. (There are multiple base pair

10 substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)

Bp 4715 - 6080 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter, GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair

15 substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)

Bp 6087 - 7285 Quail Ovalbumin gene, EMBL accession # X53964, bp 1-1199. (This sequence includes the 5'UTR, containing putative cap site bp 6087-6139.)

20 Bp 7292 - 7441 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site

Bp 7442 - 7702 Human proinsulin GenBank Accession # NM000207 bp 117-377

Bp 7703 - 7747 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)

25 Bp 7748 - 8099 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271

Bp 8101 - 11700 from cloning vector pTnMCS, bp 3716 - 7315

pTnMOD(Quail OVep+prepro+ENT+proins+syn polyA)

Bp 1 - 4045 from cloning vector pTnMCS, bp 1 - 4045

30 Bp 4051 - 4708 Quail Ovalbumin enhancer: 658 bp sequence, amplified in-house from quail genomic DNA, roughly equivalent to the far-upstream chicken ovalbumin enhancer, GenBank accession #S82527.1, bp 1-675. (There are multiple

- base pair substitutions and deletions in the quail sequence, relative to chicken, so the number of bases does not correspond exactly.)
- Bp 4715 – 6080 Quail Ovalbumin promoter: 1366 bp sequence, amplified in-house from quail genomic DNA, roughly corresponding to chicken ovalbumin promoter,
- 5 GenBank accession # J00895-M24999 bp 1-1336. (There are multiple base pair substitutions and deletions between the quail and chicken sequences, so the number of bases does not correspond exactly.)
- Bp 6087 – 6257 Cecropin cap site and Prepro, Genbank accession # X07404 bp 563-733
- 10 Bp 6264 - 6413 Synthetic spacer sequence and hairpin loop of HIV gp41 with an added enterokinase cleavage site
- Bp 6414 - 6674 Human proinsulin GenBank Accession # NM000207 bp 117-377
- Bp 6675 - 6719 Spacer DNA, derived as an artifact from the cloning vectors pTOPO Blunt II (Invitrogen) and gWIZ (Gene Therapy Systems)
- 15 Bp 6720 - 7071 Synthetic polyA from the cloning vector gWIZ (Gene Therapy Systems) bp 1920 - 2271
- Bp 7073 – 10672 from cloning vector pTnMCS, bp 3716 - 7315
- PtMod(CMV/Transposase/ChickOvep/prepro/ProteinA/ConpolyA)
- 20 BP 1-130 remainder of F1 (-) ori of pBluescriptII sk(-) (Stragagene) bp 1-130.
- BP 133-1777 CMV promoter/enhancer taken from vector pGWIZ (Gene Therapy Systems) bp 229-1873.
- BP 1780-2987 Transposase, modified from Tn10 (GenBank #J01829).
- BP 2988-2990 Engineered stop codon.
- 25 BP 2991-3343 non coding DNA from vector pNK2859.
- BP 3344-3386 Lambda DNA from pNK2859.
- BP 3387-3456 70bp of IS10 left from Tn10.
- BP 3457-3674 multiple cloning site from pBluescriptII sk(-) bp 924-707.
- BP 3675-5691 Chicken Ovalbumin enhancer plus promoter from a Topo Clone 10
- 30 maxi 040303 (5' XmaI, 3' BamHI)
- BP 5698-5865 prepro with Cap site amplified from cecropin of pMON200
- GenBank # X07404 (5'BamHI, 3'KpnI)

BP 5872-7338 Protein A gene from GenBank# J01786, mature peptide bp 292-1755
(5'KpnI, 3'SacII)

BP 7345-7752 ConPolyA from Chicken conalbumin polyA from GenBank # Y00407
bp 10651-11058. (5'SacII, 3'XhoI)

5 BP 7753-8195 multiple cloning site from pBluescriptII sk(-) bp 677-235.

BP 8196-8265 70 bp of IS10 left from Tn10.

BP 8266-8307 Lamda DNA from pNK2859

BP 8308-9151 noncoding DNA from pNK2859

BP 9152-11352 pBluescriptII sk(-) base vector (Stratagene, INC.) bp 761-2961

10

All patents, publications and abstracts cited above are incorporated herein by
reference in their entirety. It should be understood that the foregoing relates only to
preferred embodiments of the present invention and that numerous modifications or
alterations may be made therein without departing from the spirit and the scope of the

15 present invention as defined in the following claims.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A vector comprising:
 - a) a modified transposase gene operably-linked to a first promoter; wherein the nucleic acid sequence 3' to the first promoter comprises the sequence as set forth in SEQ ID NO:13, wherein SEQ ID NO:13 contains the Kozak sequence and a start codon for the transposase, and wherein at least one of the first twenty codons for the transposase gene are modified from the wild-type sequence by changing a nucleotide at a third base position of the codon to an adenine or thymine without modifying the amino acid encoded by the codon, and
 - b) one or more genes of interest operably-linked to one or more additional promoters; and wherein the one or more genes of interest and their operably-linked promoters are flanked by transposase insertion sequences recognized by the transposase encoded by the modified transposase gene.
2. The vector of claim 1, wherein the modified transposase gene comprises an adenine or thymine at the third position in each of codons 2-10 of the modified transposase gene.
3. The vector of claim 1 or 2, comprising the sequence set forth in SEQ ID NO: 1.
4. The vector of any one of claims 1 to 3, wherein the transposase is a Tn10 transposase.
5. The vector of any one of claims 1 to 4, wherein one gene of interest is operably-linked to a second promoter.
6. The vector according to any one of claims 1 to 5, wherein the first promoter and the second promoter are independently selected from the group consisting of a constitutive promoter and an inducible promoter.
7. The vector of claim 6, wherein the inducible promoter is selected from the group consisting of an ovalbumin promoter, a conalbumin promoter, a vitellogenin promoter or an ovomucoid promoter.

8. The vector according to any preceding claim, further comprising a polyA sequence operably-linked to the transposase gene.
9. The vector of claim 8, wherein the polyA sequence is a conalbumin polyA sequence.
10. The vector according to any preceding claim, further comprising two stop codons operably-linked to the transposase gene.
11. The vector of any one of claims 1 to 4, wherein a first gene of interest is operably-linked to a second promoter and a second gene of interest is operably-linked to a third promoter.
12. The vector of any one of claims 1 to 4, wherein a first and a second gene of interest are operably-linked to a second promoter.
13. The vector according to any preceding claim, further comprising an enhancer operably-linked to the one or more genes of interest.
14. The vector of claim 13, wherein the enhancer comprises at least a portion of an ovalbumin enhancer.
15. The vector according to any preceding claim, further comprising an egg directing sequence operably-linked to the one or more genes of interest.
16. The vector of claim 15, wherein the egg directing sequence is an ovalbumin signal sequence, an ovomucoid signal sequence or a vitellogenin targeting sequence.
17. A method of producing a transgenic animal comprising administering to the animal a vector according to any one of claims 1 to 16.
18. The method of claim 17, wherein the vector is administered via an intratesticular, intraarterial, intraoviductal or intraembryonic route.

19. The method of claim 17 or 18, wherein the animal is an avian animal.
20. The method of claim 19, wherein the avian animal is a chicken or a quail.
21. An egg produced by the transgenic avian animal of claim 19 or 20, wherein the egg contains one or more desired proteins encoded by the one or more genes of interest.
22. A transgenic sperm produced by the transgenic animal produced according to any one of claims 17 to 20.
23. A method for producing a desired protein comprising:
 - a) administering to the animal a vector according to any one of claims 1 to 16; and
 - b) isolating the desired protein produced in the animal.
24. The method of claim 23, wherein the animal is an egg-laying animal, and the method of administration is intraoviductal, such that the desired protein produced by the at least one gene of interest is isolated from the egg white of eggs laid by the egg-laying animal.
25. The method of claim 23 or 24, wherein the vector further comprises a TAG sequence and wherein the desired protein is purified using the TAG sequence.
26. The method of claim 25, wherein the TAG sequence comprises: (i) a sequence that encodes a polypeptide that functions as a purification handle; (ii) a cleavage site; and (iii) a polynucleotide spacer.
27. The method of claim 25 or 26, wherein the TAG sequence comprises a polynucleotide sequence shown in SEQ ID NO: 22.
28. The method of any one of claims 23 to 25, wherein the desired protein is a lytic protein, proinsulin, or a human growth hormone.

29. The method according to any one of claims 23 to 28, wherein the vector further comprises a second gene of interest operably-linked to a third promoter and wherein the genes of interest encode antibody polypeptides.

30. The vector of claim 7, wherein the inducible promoter comprises the sequence as set forth in SEQ ID NO: 17, SEQ ID NO: 40, or nucleic acids 4050-4938 of SEQ ID NO: 30.

31. The vector of claim 15 or 16, wherein the egg directing sequence comprises at least one of the sequences as set forth in SEQ ID NO: 18, nucleic acids 4960-5112 of SEQ ID NO: 3, nucleic acids 4943-5092 of SEQ ID NO: 4, nucleic acids 4958-6115 of SEQ ID NO: 29, or nucleic acids 4945-6092 of SEQ ID NO: 30.

32. The vector of claim 8, wherein the polyA sequence comprises at least one of the sequences as set forth in SEQ ID NO: 28, SEQ ID NO: 33, or nucleic acids 2995-3410 of SEQ ID NO: 1.

33. The vector according to any one of claims 1 to 16, wherein the modified transposase gene comprises an A or a T at the third position in each of codons 2-10 of the modified transposase gene.

34. The method of claim 23, wherein the vector is administered via an intratesticular, intraarterial, intraperitoneal, intravenous, intraoviductal, intraembryonic, nasal, or pronuclear route.

FIGURE 1



FIGURE 2

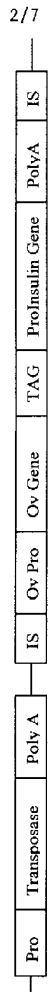


FIGURE 3



FIGURE 4



FIGURE 5

IS	Oval Pro	prepro	Heavy chain	pro	Light chain	polyA	IS
----	----------	--------	-------------	-----	-------------	-------	----

FIGURE 6

IS	Oval Pro	prepro	Light chain	ent	Heavy chain	polyA	IS
----	----------	--------	-------------	-----	-------------	-------	----

FIGURE 7

A.

Tail-to-Tail

IS	Oval Pro	Oval SS	Light chain	Poly A	Spacer DNA	Poly A	Heavy chain	Oval SS	Oval Pro	IS
----	----------	---------	-------------	--------	------------	--------	-------------	---------	----------	----

B.

Tail-to-Head

IS	Oval Pro	Oval SS	Light chain	Poly A	Spacer DNA	Oval Pro	Oval SS	Heavy chain	Poly A	IS
----	----------	---------	-------------	--------	------------	----------	---------	-------------	--------	----

Appendix A

SEQ ID NO:1 (pTnMod)

```

5  CTGACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACS 50
   CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCCG 100
   TTTCTTCCCT TCCTTTCTCG CCACGTTTCG CGGCATCAGA TTGGCTATTG 150
   GCCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200
   CATGTCCAAC ATTACCGCCA TGTGACATT GATTATTGAC TAGTTAITAA 250
10 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCG 300
   CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC 350
   CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA 400
   GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACGCGCCA 450
   CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG 500
15 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA 550
   TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC 600
   CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGG TAGCGGTTTG 650
   ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG 700
   TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ACAAACCTCCG 750
20 CCCATTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA 800
   GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC 850
   TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCCGGCGCG 900
   GGAACGGTGC ATTGGAACGC GGATTCCCGG TGCCAAGAGT GACGTAAGTA 950
   CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA 1000
25 CTGTTTTTGG CTTGGGGCCT ATACACCCCT GCTTCCTTAT GCTATAGGTG 1050
   ATGGTATAGC TTAGCCTATA GGTGTGGGTT ATTGACCAT ATTGACCACT 1100
   CCCCTATTGG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTTT 1150
   GCCACAATA TCTCTATTGG CTATATGCCA ATACTCTGTC CTTCAGAGAC 1200
   TGACACGGAC TCTGTATTTT TACAGGATGG GGTCCCATTT ATTATTTACA 1250
30 AATTACATA TACAACAACG CCGTCCCGCG TGCCCGCAGT TTTTATTAAA 1300
   CATAGCGTGG GATCTCCACG CGAATCTCGG GTACGTGTTT CGGACATGGG 1350
   CTCTTCTCCG GTAGCGGCGG AGCTTCCACA TCCGAGCCCT GGTCCCATGC 1400
   CTCACGCGGC TCATGGTCCG TCGGCAGCTC CTTGCTCCTA ACAGTGGAGG 1450
   CCAGACTTAG GCACAGCACA ATGCCACCA CCACCATGTG GCCGCACAAG 1500
35 GCCGTGGCGG TAGGGTATGT GTCTGAAAT GAGCGTGGAG ATTGGGCTCG 1550
   CACGGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAAGATGCAG 1600
   GCAGCTGAGT TGTGTATTTC TGATAAGAGT CAGAGGTAAC TCCCGTTGCG 1650
   GTGCTGTAA CCGTGGAGGG CAGTGTAGTC TGAGCAGTAC TCGTTGCTGC 1700
   CGCGCGCGCC ACCAGACATA ATAGCTGACA GACTAACAGA CTGTTCTCTT 1750
40 CCATGGGTCT TTTCTGCAGT CACCGTCGGA CCATGTGTA ACTTGATATT 1800
   TTACATGATT CTCCTTACCA ATTCTGCCCG GAATTACACT TAAACGAGCT 1850
   CAACAGCTTA ACGTTGGCTT GCCACGCATT ACTTGACTGT AAACTCTCA 1900
   CTCTTACCGA ACTTGGCCGT AACCTGCCAA CCAAAGCGAG AACAAAACAT 1950
   AACATCAAA CAAATCGACG ATTGTAGGT AATCGTCACC TCCACAAAGA 2000
45 GCGACTCGCT GTATACCGTT GGCATGCTAG CTTTATCTGT TCGGGAATAC 2050
   GATGCCCAT TACTTGTGTT ACTGGTCTGA TATTCGTGAG CAAAAACGAC 2100
   TTATGGTATT GCGAGCTTCA GTGCACTAC ACGGTCGTTT TGTTACTCTT 2150
   TATGAGAAAG CGTTCCCGCT TTCAGAGCAA TGTTCAAAGA AAGCTCATGA 2200
   CCAATTTCTA GCCGACCTTG CGAGCATTCT ACCGAGTAAC ACCACACCGC 2250
50 TCATTTGTGAG TGATGCTGGC TTTAAAGTGC CATGGTATAA ATCCGTTGAG 2300
   AAGCTGGGTT GGTACTGGTT AAGTCGAGTA AGAGGAAAAG TACAATATGC 2350
   AGACCTAGGA GCGGAAAAC GGAACCTAT CAGCAACTTA CATGATATGT 2400
   CATCTAGTCA CTCAAAGACT TTAGGCTATA AGAGGCTGAC TAAAGCAAT 2450
   CCAATCTCAT GCCAAATTCT ATTGTATAAA TCTCGCTCTA AAGGCGGAAA 2500
55 AAATCAGCGC TCGACACGGA CTCATTGTCA CCACCCGTCA CCTAAATCT 2550
   ACTCAGCGTC GGCAAAAGGAG CCATGGGTTT TAGCAACTAA CTTACCTGTT 2600
   GAAATTCGAA CACCCAAACA ACTTGTTAAT ATCTATTGCA AGCGAATGCA 2650
   GATTGAAGAA ACCTTCCGAG ACTTGAAAAG TCCTGCCTAC GGACTAGGCC 2700
   TACGCCATAG CCGAACGAGC AGCTCAGAGC GTTTTGATAT CATGCTGCTA 2750
60 ATCGCCCTGA TCGTTCAACT AACATGTTGG CTTGCGGGCG TTCATGCTCA 2800
   GAAACAAGGT TGGGACAAGC ACTTCAGGC TAACACAGTC AGAATCGAA 2850

```

103

ATLJ02 113492.1

5 ACGTACTCTC AACAGTTCGC TTAGGCATGG AAGTTTTGCG GCAITTCGGC 2900
 TACACAATAA CAAGGGAAGA CTTACTCGTG GCTGCAACCC TACTAGCTCA 2950
 AAATTTATTC ACACATGGTT ACGCTTTGGG GAAATTATGA TAATGATCCA 3000
 GATCACCITCT GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTTGG 3050
 10 TTTTTTGTGG ATCTGCTGTG CTTCTAGTT GCCAGCCATC TGTGTTTGC 3100
 CCCGCCCGG TGCTTCCTT GACCTGGAA GGTGCCACTC CCACGTCTCT 3150
 TTCCTAATAA AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCTAT 3200
 CTATTCTGGG GGGTGGGGTG GGGCAGCACA GCAAGGGGGA GGATTGGGAA 3250
 GACAAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 15 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCGGTAC CTCTCTCTCT 3350
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCAGG TGCTGAAGAA 3400
 TTGACCCGGT GACCAAGGT GCCTTTTATC ATCACTTTAA AAATAAAAAA 3450
 CAATTACTCA GTGCTGTGTA TAAGCAGCAA TTAATTATGA TTGATGCTTA 3500
 CATCACAAAC AAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 20 TATTTGAACA TTATCTTGAT TATATTATG ATAATAATAA AAACCTTAATC 3600
 CCTATCCAAG AAGTGATGCC TATCATTGGT TGGAAATGAC TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GGTAAAGTAA ACGCCATTGT CAGCAAAATTG 3700
 ATCCAAGAGA ACCAACTTAA AGCTTTCCTG ACGGAATGTT AATCTCGTIT 3750
 GACCTGAGC ACTGATGAAT CCCCTAATGA TTTGGSTAAA AATCATTAAG 3800
 25 TTAAGGIGGA TACACATCTT GTCATATGAT CCCGGTAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCGAGG CTTTACACTT TATGCTFCCG GCTCGTATGT 3900
 TGTGTGGAAT TGTGAGCGGA TAACAATTTT ACACAGGAAA CAGCTATGAC 3950
 CATGATTACG CCAAGCGCGC AATTAAACCCT CACTAAAGGG AACAAAAGCT 4000
 GGAGCTCCAC CGCGGTGGCG GCCGCTCTAG AACTAGTGGG TCCCCCGGCG 4050
 30 TGCAGGAATT CGATATCAAG CTTATCGATA CCGCTGACCT CGAGGGGGGG 4100
 CCGGTACCC AATTGCGCCT ATAGTGAGTC GTATTACGCG CGCTCACTGG 4150
 CCGTCTGTTT ACAACGTCGT GACTGGGAAA ACCCTGGCGT TACCCAACTT 4200
 AATCGCCTTG CAGCACATCC CCCTTTCGCC AGCTGGCGTA ATAGCGAAGA 4250
 GGCCCGCACC GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT 4300
 35 GGAAATGTGA AGCGTTAATA TTTTGTAAAA ATTCGCGTTA AATTTTTGIT 4350
 AAATCAGCTC ATTTTAAAC CAATAGGCGG AAATCGGCAA AATCCCTTAT 4400
 AAATCAAAAG AATAGACCGA GATAGGGTTG AGTGTGTTC CAGTTGGGAA 4450
 CRAAGATCCA CTATTAAGA ACGTGGACTC CAACGTCAAA GGGCGAAAAA 4500
 CCGTCTATCA GGGCGATGGC CCACTACTCC GGGATCATAT GACAAGATGT 4550
 40 GTATCCACCT TAACTTAATG ATTTTACCA AAATCATTAG GGGATTCAATC 4600
 AGTGCTCAGG GTCAACGAGA ATTAACATTC CGTCAGGAAA GCTTATGATG 4650
 ATGATGTGCT TAAAACTTA CTCATGGCTT GGTATGTCAT ATCGCAATAC 4700
 ATGCGAAAAA CCTAAAAGAG CTTGCCGATA AAAAAGGCCA ATTTATTGCT 4750
 ATTTACCGCG GCTTTTATTT GAGCTTGAAA GATAAATAAA ATAGATAGGT 4800
 45 TTTATTGAA GCTAAATCTT CTTTATCGTA AAAAATGCCC TCTTGGGTAA 4850
 TCAAGAGGGT CATTATATTT CGCGGAATAA CATCATTGG TGACGAAATA 4900
 ACTAAGCACT TGTCTCTGT TTAATCCCTT GAGCTTGAGG GGTAAACATG 4950
 AAGGTCATCG ATAGCAGGAT AATAATACAG TAAAACGCTA AACCAATAAT 5000
 CCAATCCAG CCATCCCAAA TTGGTAGTGA ATGATTATAA ATAACAGCAA 5050
 50 ACAGTAATGG GCCAATAACA CCGGTGTCAT TGGTAAGGCT CACCAATAAT 5100
 CCCTGTAAG CACCTTGCTG ATGACTCTTT GTTGGATAG ACATCACTCC 5150
 CTGTAATGCA GGTAAAGCGA TCCCACCACC AGCCAATAAA ATTAACACAG 5200
 GGAAACTTAA CCAACCTTCA GATATAAAGC CTAAAAAGGC AAATGCACTA 5250
 CTATCTGCAA TAAATCCGAG CAGTACTGCC GTTTTTTCG CCAATTAGTG 5300
 55 GCTATCTTTC CTGCCACAAA GGCTTGGAAT ACTGAGTGTA AAAGACCAAG 5350
 ACCCOTAATG AAAAGCCAAAC CATCATGCTA TTCATCATCA CGAATTCTGT 5400
 AATAGCACCA CACCGTGTCT GATTGGCTAT CAATGCGCTG AAATAATAAT 5450
 CAACAAATGG CATCGTTAAA TAAGTGATGT ATACCGATCA GCTTTTGTTC 5500
 CTTTATGTA GGGTTAAATG CGCGCTTGCC GTAATCATGG TCATAGCTGT 5550
 60 TTCCTGTGTG AAATTTGTAT CCGTCAACAA TTCCACACAA CATACGAGCC 5600
 GGAAGCATAA AGTGTAAAGC CTGGGGTCCC TAATGAGTGA GCTAACTCAC 5650
 ATTAATTGCG TTGCGCTCAC TGCCCGCTTT CCAGTCGGGA AACCTGTCGT 5700
 GCCAGCTGCA TTAATGAATC GGCCAACGCG CGGGGAGAGG CGGTTGCGGT 5750
 ATTGGGCGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCTG 5800
 65 TCGGCTCGCG CGACCGGTAT CAGCTCACTC AAAGCGGTA ATACGGTTAT 5850
 CCACAGAAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG 5900

CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG 5950
 GCTCCGCCCC CCTGACGAGC ATCAAAAAA TCGACGCTCA AGTCAGAGGT 6000
 GCGGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCTCGAAGC 6050
 TCCTCTCGTG GCTCTCCTGT TCCGACCTTG CCGCTTACCG GATACCTGTC 6100
 CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA 6150
 5 GGTATCTCAG TTCGGTGTAG GTCGTTGCGT CCAAGCTGGG CTGTGTGCAC 6200
 GAACCCCCCG TTCAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT 6250
 TGATCCCAAC CCGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG 6300
 GTAACAGGAT TAGCAGAGCG AGGTATGTAG CCGGTGCTAC AGAGTTCTTG 6350
 10 AAGTGGTGGC CTAACCTACG CTACACTAGA AGGACAGTAT TTGGTATCTG 6400
 CGCTCTGCTG AAGCCAGTTA CCTTCGGAAG AAGAGTTGGT AGCTCTTGAT 6450
 CCGGCAAAAC AACCCACGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG 6500
 CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC 6550
 TACGGGCTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG 6600
 15 TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCTTTTT AAATTAAGAA 6650
 TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG 6700
 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC 6750
 GTTCACTCAT AGTTGCGTGA CTCGCCGTGG TGTAAGTAAC TACGATACGG 6800
 GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG 6850
 20 CTCACCGGCT CCAGATTATAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG 6900
 AGCGCAGAAG TGGTCTGCA ACTTTATCCG CCTCCATCCA GTCTATTAA 6950
 TGTTCGCGGG AAGCTAGAGT AAGTAGTTCC CCAGTTAATA GTTTGCGCAA 7000
 CGTTGTTGCC ATTGCTACAG GCATCGTGGT GTCACGCTCG TCGTTTGGTA 7050
 TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC 7100
 25 CCCATGTTGT GCAAAAAAGC GGTAGCTCC TTCCGTCCTC CGATCGTTGT 7150
 CAGAACTAAG TTGGCCGAG TGTATCACT CATGGTTATG GCAGCACTGC 7200
 ATAATCTCTT TACTGTCAAT CCATCCGTAA GATGCTTTTC TGTGACTGGT 7250
 CAGTACTCAA CCAAGTCATT CTGAGAATAG TGTATGGGCG GACCGAGTTG 7300
 CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT AGCAGAACTT 7350
 30 TAAAGTGCT CATCATTTGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG 7400
 ATCTTACGCG TGTGAGATC CAGTTGATG TAACCCATC GTGCAACCAA 7450
 CTGATCTTCA GCATCTTTTA CTTTACCAG CGTTTCTGGG TGAGCAAAAA 7500
 CAGGAAGGCA AAATGCCGCA AAAAAGGGA TAAGGGGAGC ACGGAAATGT 7550
 TGAATACTCA TACTCTTCCT TTTCAATAT TATTGAAGCA TTTATCAGGG 7600
 35 TTATTCTCTC ATGAGCGGAT ACATATTGTA ATGTATTAG AAAAATAAAC 7650
 AATAGGGGT TCCGCGACA TTTCCCGAA AAGTGCCAC 7689

SEQ ID NO:2 (PtnMod (CMV/Red))
 40 CTGACGCGCC CTGTAGCGGC GCATTAAAGCG CGGCGGGTGT GGTGGTTACG 50
 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTTCG 100
 TTTCTTCCCT TCCTTTCTCG CCACGTTTCG CGGCATCAGA TTGGCTATTG 150
 GGCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200
 45 CATGTCCAAC ATTACCGCCA TGTGACATT GATTATTGAC TAGTTATTAA 250
 TAGTAATCAA TTACGGGGTC ATTAGTTCAI AGCCCATATA TGGAGTTCCG 300
 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC 350
 CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA 400
 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAC TGCCCA 450
 50 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG 500
 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA 550
 TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC 600
 CATGGTGTAG CGGTTTGGC AGTACATCAA TGGCGGTGGA TAGCGGTTTC 650
 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG 700
 55 TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTGCGTA ACAACTCCGC 750
 CCCATTGACC CAAATGGCGG GTAGGCGTGT ACGGTGGGAG GTCTATATAA 800
 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC 850
 TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG 900
 GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA 950
 60 CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA 1000
 CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTTCCTTAT GCTATAGGTG 1050

ATGSTATAGC TTAGCCTATA GGTGTGGGT ATTGACCATT ATTGACCCT 1100
 CCCTTATTGG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTTT 1150
 GCCACAACIA TCTCTATTGG CTATATGCCA ATACTCTGTC CTTGAGAGAC 1200
 TGACACGGAC TCTGTATTTT TACAGGATGG GGTCCCATTT ATTATTTACA 1250
 5 AATTACATA TACAACAACG CCGTCCCCCG TGCCCGCAGT TTTTATTA 1300
 CATAGCGTGG GATCTCCACG CGAATCTCGG GTACGTGTTC CGGACATGGG 1350
 CTCTTCTCCG GTAGCGGCGG AGCTTCCACA TCCGAGCCCT GGTCCCATGC 1400
 CTCCAGCGGC TCATGGTCGC TCGGCAGCTC CTTGCTCCTA ACAGTGGAGG 1450
 CCAGACTTAG GCACAGCACA ATGCCACCA CCACCACTGT GCCGCACAAG 1500
 10 GCCGTGGCGG TAGGGTATGT GTCTGAAAAT GAGCGTGGAG ATTGGGCTCG 1550
 CACGGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCAG 1600
 GCAGCTGAGT TGTGTATTC TGATAAGAGT CAGAGGTAAC TCCCGTTGCG 1650
 GTGCTGTAA CGGTGGAGGG CAGTGTAGTC TGAGCAGTAC TCGTTGCTGC 1700
 CGCGCCGCCG ACCAGACATA ATAGCTGACA GACTAACAGA CTGTTCCCTT 1750
 15 CCATGGGTCT TTCTGCGAGT CACCGTCGGA CCATGTGTGA ACTTGATATT 1800
 TTACATGATT CTCTTTACCA ATTCTGCCCC GAATTACACT TAAAACGACT 1850
 CAACAGCTTA ACGTTGGCTT GCCACGCATT ACTTGACTGT AAAACTCTCA 1900
 CTCTTACGA ACTTGGCCGT AACCTGCCAA CCAAAGCGAG AACAAAACAT 1950
 AACATCAAA GAATGACCGG ATGTGTAGGT AATCGTCACC TCCACAAAGA 2000
 20 GCGACTCGCT GTATACCGTT GGCATGCTAG CTTTATCTGT TCGGGAATAC 2050
 GATGCCCATT GTACTTGTGG ACTGGTCTGA TATTCGTGAG CAAAACGAC 2100
 TTATGGTATT GCGAGCTTCA GTCCGACTAC ACGGTGCTTC TGTACTCTT 2150
 TATGAGAAAG CGTTCGCGT TTCAGAGCAA TGTTCAAAGA AAGCTCATGA 2200
 CCAATTTCTA GCCGACCTTG CGAGCATCTC ACCGAGTAAC ACCACACCGC 2250
 25 TCATTTGTCG TGATGCTGGC TTTAAAGTGC CATGGTATAA ATCCGTTGAG 2300
 AAGCTGGGTT GGTACTGGTT AAGTCGAGTA AGAGGAAAAG TACAATATGC 2350
 AGACCTAGGA GCGGAAAAC GGAACCTAT CAGCAACTTA CATGATATGT 2400
 CATCTAGTCA CTCAAAGACT TTAGGCTATA AGAGGCTGAC TAAAGCAAT 2450
 CCAATCTCAT GCCAAATCTC ATTGTATAAA TCTCGCTCTA AAGGCCGAAA 2500
 30 AAATCAGCGC TCGACACGGA CTCATTGTCA CCACCCGTCA CCTAAAATCT 2550
 ACTCAGCGTC GGCAAAGGAG CCATGGGTTT TAGCAACTAA CTTACCTGTT 2600
 GAAATTCGAA CACCACAAAC ACTTGTAAAT ATCTATTGCA AGCGAATGCA 2650
 GATIGAAGAA ACCTTCCGAG ACTTGAAAAG TCCTGCCTAC GGACTAGGCC 2700
 TAGCCCATAG CCGAACGAGC AGCTCAGAGC GTTTTGATAT CATGCTGCTA 2750
 35 ATCGCCCTGA TGCTTCAACT AACATGTTGG CTTGCGGGCG TTCATGCTCA 2800
 GAAACAAGGT TGGGACAAGC ACTTCCAGGC TAACACAGTC AGAAATCGAA 2850
 ACGTACTCTC AACAGTTTCG TTAGGCATGG AAGTTTTCGCG GCATTCTGCG 2900
 TACACAATAA CAAGGGAAGA CTTACTCGTG GCTGCAACCC TACTAGCTCA 2950
 AAATTTATT ACACATGGTT ACGCTTTGGG GAAATATGA TAATGATCCA 3000
 40 GATCACTTCT GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTTGG 3050
 TTTTTTGGG ATCTGCTGTG CCTCTAGTT GCCAGCCATC TGTGTTTTGC 3100
 CCTCCGCCG TGCTTCTCT GACCTGGAA GGTGCCACTC CCACTGTCTC 3150
 TTCTAATAA AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCAIT 3200
 CTATTCTGG GGTGGGGTG GGGCAGCACA GCAAGGGGA GGATTGGGAA 3250
 45 GACAATAGCA GGCATGCTGG GATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCGGTAC CTCTCTCTCT 3350
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCAGG TGCTGAAGAA 3400
 TTGACCCGCT GACCAAAGGT GCCTTTTATC ATCACTTTAA AAATAAAAAA 3450
 CAATTACTCA GTGCCTGTTA TAAGCAGCAA TTAATATGA TTGATGCCCTA 3500
 50 CATCACAAAC AAAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 TATTGAACA TTATCTTGAT TATATTATG ATAATAATAA AAACCTTATC 3600
 CCTATCCAAG AAGTGATGCC TATCATTGGT TGAATGAAC TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GGTAAAGTAA ACGCCATTGT CAGCAAAATTG 3700
 ATCCAAGAGA ACCAACTTAA AGCTTTCCTG ACGGAATGTT AATTCTCGTT 3750
 55 GACCTGAGC ACTGATGAAT CCGCTAATGA TTTGGTAAA AATCATTAAG 3800
 TTAAGTGGGA TACACATCTT GTCATATGAT CCGGTAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCAGG CTTTACACTT TATGCTTCG GCTCGTATGT 3900
 TGTGTGAAT TGTGAGCGGA TAACAATTTT ACACAGGAAA CAGCTATGAC 3950
 CATGATTACG CCAAGCGCGC AATTAAACCT CACTAAAGGG AACAAAAGCT 4000
 60 GGAGCTCCAC CGCGTGGCG GCCGCTCTAG AACTAGTGGG TCCCGCGGC 4050
 ATCAGATTGG CTATTGGCCA TTGCATACGT TGTATCCATA TCATAATATG 4100

TACATTTATA TTGGCTCATG TCCAACATTA CCGCCATGTT GACATTGATT 4150
 ATTGACTAGT TATTAATAGT AATCAATTAC GGGGTCATTA GTTCATAGCC 4200
 CATATATGGA GTTCCGCGTT ACATAACTTA CGGTAAATGG CCGCGCTGGC 4250
 TGACCGCCCA ACGACCCCGG CCCATTGACG TCAATAATGA CGTATGTTCC 4300
 5 CATAGTAACG CCAATAGGGA CTTTCCATTG ACGTCAATGG GTGGAGTATT 4350
 TACGGTAAAC TGCCCACTTG GCAGTACATC AAGTCTATCA TATGCCAAGT 4400
 ACGCCCCCTA TTGACGTCAG TGACGGTAAA TGGCCCGCCT GGCATTATGC 4450
 CCAGTACATG ACCTTATGGG ACTTTCCTAC TTGGCAGTAC ATCTACGTAT 4500
 TAGTCATCGC TATTACCATG GTGATGCGGT TTTGGCAGTA CATCAATGGG 4550
 10 CGTGGATAGC GGTTTGACTC ACGGGGATTT CCAAGTCTCC ACCCCATTGA 4600
 CGTCAATGGG AGTTTGTTTT GGCACCAAAA TCAACGGGAC TTTCCAAAAT 4650
 GTCGTAACAA CTCCGCCCCA TTGACGCAAA TGGGCGGTAG GCGTGACGG 4700
 TGGGAGGTCT ATATAAGCAG AGCTCGTTTA CTGAACCGTC AGATCGCCTG 4750
 GAGACGCCAT CCACGCTGTT TTGACCTCCA TAGAAGACAC CGGGACCGAT 4800
 15 CCAGCCTCCG CGCGCGGAAA CGGTGCATTG SAACGCGGAT TCCCGTGCC 4850
 AAGAGTGACG TAAGTACCGC CTATAGACTC TATAGGCACA CCCCTTTGGC 4900
 TCTTATGCTG GCTATCTGT TTTTGGCTTG GGGCCTATAC ACCCCCGCTT 4950
 CTTTATGCTA TAGGTGATGG TATAGCTTAG CCTATAGGTG TGGGTTATTG 5000
 ACCATTATTG ACCACTCCCC TATTGGTGAC GATACTTTCC ATTACTAATC 5050
 20 CATAACATGG CTCTTTGCCA CAACTATCTC TATTGGCTAT ATGCCAATAC 5100
 TCTGTCTTTC AGAGACIGAC ACGGACTCTG TATTTTACA GGATGGGGTC 5150
 CCATTTATTA TTACAAATT CACATATACA ACAACGCCGT CCCCGTGCC 5200
 CGCAGTTTTT ATTAAACATA GCGTGGGATC TCCACGCGAA TCTCGGGTAC 5250
 GTGTTCCGGA CATGGGCTCT TCTCCGCTAG CGGCGGAGCT TCCACATCCG 5300
 25 AGCCCCGTGC CCATGCCTCC AGCGGCTCAT GGTGCTCGG CAGCTCCTTG 5350
 CTCCTAACAG TGGAGGCCAG ACTTAGGCAC AGCACAATGC CCACCAACAC 5400
 CAGTGTGCGC CACAAGGCCG TGGCGGTACG GTATGTGTCT GAAATATGAC 5450
 GTGGAGATTG GGCTCGCACG GCTGACGCG ATGGAAGACT TAAGGCAGCG 5500
 GCAGAAGAAG ATGACGGCAG CTGAGTTGTT GTATTCTGAT AAGAGTCAGA 5550
 30 GGTAACTCCC GTTGGCGTGC TGTTAAACGGT GGAGGGCAGT GTAGTCTGAG 5600
 CAGTACTCGT TGCTGCCGCG CGCGCCACCA GACATAATAG CTGACAGACT 5650
 AACAGACTGT TCCTTTCCAT GGGTCTTTTC TGCACTCACC GTCTCGCGAC 5700
 AAGGATCCAC CGGTGCCAC CATGGTGCGC TCCTCCAAGA ACGTCATCAA 5750
 GGAGTTCATG CGCTTCAAGG TGCGCATGGA GGGCACCCTG AACGCCACCG 5800
 35 AGTTCGAGAT CGAGGGCGAG GCGGAGGCCG GCCCTACGA GGGCCACAC 5850
 ACCGTGAAGC TGAAGGTGAC CAAGGGCGCG CCCCTGCCCT TCGCTGGGA 5900
 CATCCTGTCC CCCAGTTCC AGTACGGCTC CAAGGTGTAT GTGAAGCACC 5950
 CCGCCGACAT CCCGACTAC AAGAAGCTGT CCTTCCCGA GGGCTTCAAG 6000
 TGGGAGCGCG TGATGAACCT CGAGGACGGC GGGCTGGTGA CCGTGACCCA 6050
 40 GGACTCCTCC CTGACGGACG GCTGCTTCAT CTACAAGGTG AAGTTCATCG 6100
 CGGTGAACCT CCCCTCCGAC GCGCCCGTAA TGCAAGAAGAA GACCATGGGC 6150
 TGGGAGGCCT CACCGAGCG CCGTACCCC CGCGACGGCG TGCTGAAGGG 6200
 CGAGATCCAC AAGGCCCTGA AGCTGAAGGA CGGCGGCCAC TACCTGTGG 6250
 AGTICAAAGTC CATCTACATG GCCAAGAAGC CCGTGACGCT GCCCGCTAC 6300
 45 TACTACGTGG ACTCCAAGCT GGCATCACC TCCCACAACG AGGACTACAC 6350
 CATCGTGGAG CAGTACGAGC GCACCGAGGG CCGCCACCAC CTGTTCTGT 6400
 AGCGGCCGCG ACTCTAGATC ATAATCAGCC ATACCACATT TGTAGAGGTT 6450
 TTACTTGCTT TAAAAAACCCT CCCACACCTC CCCCTGAACC TGAACATAA 6500
 AATGAATGCA ATTGTTGTTG TTAACITGTT TATTGACGCT TATAATGGTT 6550
 50 ACAAAATAAG CAATAGCATC ACAAAITTC AATAAAGC ATTTTITTC 6600
 CTGCATTCTA GTTGTGGCCC GGGCTGCAGG AATTGATAT CAAGCTTATC 6650
 GATACCGCTG ACCTCGAGGG GGGGCCCGGT ACCCAATTCC CCCTATAGTG 6700
 AGTCGTATTA OGGCGCTCA CTGGCGCTCG TTTTACAACG TCGTACTGG 6750
 55 GAAAAACCTG GCGTTACCCA ACTTAATCGC CTTCGAGCAC ATCCCCCTT 6800
 CGCCAGCTGG CGTAATAGCG AAGAGGCCCG CACCGATCGC CCTTCCCAAC 6850
 AGTTGCGCAG CCGAATGGC GAATGGAAT TGTAAGCGTT AATATTTTGT 6900
 TAAAAATCGC GTTAAATTTT TGTAAATCA GCTCATTTT TAACCAATAG 6950
 GCCGAAATCG GCAAAATCCC TTATAATCA AAGAATAGA CCGAGATAGG 7000
 GTTGAAGTGT GTTCCAGTTT GGAACAAGAG TCCACTATTA AAGAACGTGG 7050
 60 ACTCCAACGT CAAAGGCGA AAAACCGTCT ATCAGGCGCA TGGCCACTA 7100
 CTCGGGATC ATATGACAAG ATGTGTATCC ACCTTAAGTT AATGATTTTT 7150

ACCAAAATCA TTAGGGGATT CATCAGTCT CAGGGTCAAC GAGAATTAAC 7200
 ATTCCGTCAG GAAAGCTTAT GATGATGATG TGCTTAAAAA CTTACTCAAT 7250
 GGCTGGTTAT GCATATCGCA ATACATGCGA AAAACCTAAA AGAGCTTGCC 7300
 GATAAAAAAG GCCAATTTAT TGCTATTTAC CGCGGCTTTT TATTGAGCTT 7350
 5 GAAAGATAAA TAAATAGAT AGTTTTTATT TGAAGCTAAA TCTTCTTTAT 7400
 CGTAAAAAAT GCCCTCTTGG GTTATCAAGA GGGTCATTAT ATTTCCGGGA 7450
 ATAACATCAT TTGGTGACGA AATAACTAAG CACTTGTCTC CTGTTTACTC 7500
 CCTGAGCTT GAGGGGTTAA CATGAAGGTC ATCGATAGCA GGATAATAAT 7550
 ACAGTAAAAA GCTAAACCAA TAATCCAAAT CCAGCCATCC CAAATTGGTA 7600
 10 GTGAATGATT ATAAATAACA GCAAACAGTA ATGGGCCAAT AACACCGGTT 7650
 GCATTGGTAA CGCTCACCAA TAATCCCTGT AAAGCACCTT GCTGATGACT 7700
 CTTTGTTTGG ATAGACATCA CTCCTGTAA TGCAAGTAAA GCGATCCCAC 7750
 CACCAGCCAA TAAATTTAAA ACAGGGAAAA CTAACCAACC TTCACATATA 7800
 AACGCTAAAA AGGCAAATGC ACTACTATCT GCAATAAATC CGAGCAGTAC 7850
 15 TGCCGTTTTT TCGCCCATTT AGTGGCTATT CTTCTGCCA CAAAGGCTTG 7900
 GAATACTGAG TGTAAAAGAC CAAGACCCGT AATGAAAAGC CAACCATCAT 7950
 GCTATTATC ATCAGGATTT CTGTAATAGC ACCACACCGT GCTGGATTGG 8000
 CTATCAATGC GCTGAAATAA TAATCAACAA ATGGCATCGT TAAATAAGTG 8050
 ATGTATACCG ATCAGCTTTT GTTCCCTTTA GTGAGGGTAA ATTGCGCGCT 8100
 20 TGGCGTAATC ATGGTCATAG CTGTTTCCTG TGTGAAATTG TTATCCGCTC 8150
 ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGA AAGCCTGGGG 8200
 TGCCATATGA GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACTGCCCG 8250
 CTTTCAGTC GGGAAACCTG TCGTGCCAGC TGCAATTAATG AATCGGCCAA 8300
 CGCCCGGGGA GAGGCGGTTT GCGTATTGGG CGCTCTTCG CTTCTCGCTG 8350
 25 CACTGACTCG CTGCGCTCGG TCGTTCGGCT CGCGCGAGCG GTATCAGCTC 8400
 ACTCAAAGGC GGTAAATACG TTATCCACAG AATCAGGGGA TAACGCAGGA 8450
 AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC 8500
 CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG CCCCCTGAC GAGCATCACA 8550
 30 AAAATCGACG CTCAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAGA 8600
 TACCAGGCGT TTCCCTCTGG AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC 8650
 CCTGCCGCTT ACCGGATACC TGTCCGCTT TCTCCCTTCG GGAAGCGTGG 8700
 CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCCGT GTAGGTCGTT 8750
 CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCGGTTGAGC CCGACCGCTG 8800
 35 CGCCTTATCC GGTAACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT 8850
 TATCGCCACT GGCAGCAGCC ACTGTAACA GGATTAGCAG AGCGAGGTAT 8900
 GTAGGCGGTG CTACAGAGTT CTTGAAGTGG TGGCCTAACT ACGGCTACAC 8950
 TAGAAGGACA GTATTGGTA TCTGCGCTCT GCTGAAGCCA GTTACCTTCG 9000
 GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC 9050
 40 GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC 9100
 TCAAGAGAT CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGAAGAC 9150
 AAAAATCAGG TTAAGGGATT TTGGTCATGA GATTATCAAA AAGGATCTTC 9200
 ACCTAGATCC TTTTAAATTA AAAATGAAGT TTTAAATCAA TCTAAAGTAT 9250
 ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC AGTGAGGCAC 9300
 45 CTATCTCAGC GATCTGTCTA TTTGCTTCAT CCATAGTTGC CTGACTCCCC 9350
 GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC 9400
 TGCAATGATA CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA 9450
 TAAACCAAGC AGCCGGAAGG GCCGAGCGCA GAAGTGGTCC TGCAACTTTA 9500
 TCCGCTCCA TCCAGTCTAT TAATTGTTGC CGGGAAGCTA GAGTAAGTAG 9550
 TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT ACAGGCATCG 9600
 50 TGGTGTACAG CTCGTCGTTT GGTATGGCTT CATTAGCTC CGGTTCCTCA 9650
 CGATCAAGGC GAGTTACATG ATCCCCATG TTGTGCAAAA AAGCGGTTAG 9700
 CTCCTTCGGT CCTCCGATCG TTGTGAGAAG TAAGTTGGCC GCAGTGTAT 9750
 CACTCATGCT TATGGCAGCA CTGCATAATT CTCTTACTGT CATGCCATCC 9800
 GTAAGATGCT TTTCTGTGAC TGGTGAGTAC TCAACCAAGT CATTCTGAGA 9850
 55 ATAGTGTATG CCGCGACCGA GTTGCTCTTG CCGGCGTCA ATACGGGATA 9900
 ATACCGGCCC ACATAGCAGA ACTTTAAAG TGCTCATCAT TGGAAAACGT 9950
 TCTTCGGGGC GAAAATCTCT AAGGATCTTA CCGCTGTGTA GATCCAGTTC 10000
 GATGTAACCC ACTCGTGAC CCAACTGATC TTCAGCATCT TTTACTTTCA 10050
 CCAGCGTTTC TGGGTGAGCA AAAACAGGAA GSCAAAATGC CGCAAAAAG 10100
 60 GGAATAAGGG CGACACGGAA ATGTTGAATA CTCATACTCT TCCTTTTCA 10150
 ATATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC GCATACATAT 10200

TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTCCCGG CACATTTCCC 10250
CGAAAAGTGC CAC 10263

5 SEQ ID NO:3 (PtnMod (Oval/Red) Chicken)

CTGACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG 50
CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCCG 100
TTTCTCCCT TCCTTTCTCG CCACGTTCCG CGGCATCAGA TTGGCTATTG 150
10 GCCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200
CATGTCCACG ATTACCGCCA TGTGTACATT GATTATTGAC TAGTATTAA 250
TAGTAATCAA TTACGGGGTC ATTAGTTTAT AGCCCATATA TGGAGTTCCG 300
CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC 350
CCCGCCCAT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA 400
15 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACGCGCA 450
CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG 500
TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA 550
TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGTATTAC 600
CATGGTGATG CGGTTTGGC AGTACATCAA TGGGCGTGA TAGCGGTTG 650
20 ACTCACGGGG ATTTCCAAGT CTCACCCCA TTGACGTCAA TGGGAGTTG 700
TTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGA ACAACTCCG 750
CCCATTCAGC CAAATGGGCG GTAGGCGTGT ACGTGGGAG GTCTATATAA 800
GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACG 850
TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCG 900
25 GGAACGGTGC ATTTGGAACG GATTCGCCG TGCCAAGAGT GACGTAAGTA 950
CGGCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA 1000
CTGTTTTTGG CTTGGGGCCT ATACACCCCG GCTTCCTTAT GCTATAGTGT 1050
ATGGTATAGC TTAGCCTATA GGTGTGGGTT ATTGACCATT ATTGACCACT 1100
CCCCATTGG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTTT 1150
30 GCCACAATA TCTCTATTGG CTATATGCCA ATACTCTGTC CTTCAGAGAC 1200
TGACACGGAC TCTGTATTTT TACAGGATGG GGTCCCATTT ATTATTTACA 1250
AATTACATA TACAACAACG CCGTCCCGCG TGCCCGCAGT TTTTATTAA 1300
CATAGCGTGG GAITCTCACG CGAATCTCGG GTACGTGTTT CGGACATGGG 1350
CTCTTCTCCG GTAGCGGCGG AGCTTCCACA TCCGAGCCCT GGTCCCATGC 1400
35 CTCCAGCGGC TCATGGTCCG TCGGCAGCTC CTGTCTCTTA ACAGTGGAGG 1450
CCAGACTTAG GCACAGCACA ATGCCACCA CCACCACTGT GCCGACAAG 1500
GCCGTGGCGG TAGGGTATGT GTCTGAAAT GAGCGTGGAG ATTGGGCTCG 1550
CCGGCTGAC GCAGATGGA GACTTAAGCG AGCGGCAGAA GAAGATGCAG 1600
CAGCTGAGT TGTGTATTC TGATAAGAGT CAGAGGTAAC TCCGTTGCG 1650
40 GTGCTGTAA CGGTGGAGGG CAGTGTAGTC TGAGCAGTAC TCGTTGCTGC 1700
CGCGCGCGCC ACCAGACATA ATAGCTGACA GACTAACAGA CTGTTCTTT 1750
CCATGGGTCT TTTCTGAGT CACCGTCGGA CCATGTGTGA ACTGTATATT 1800
TTACATGATT CTCTTTACCA ATCTGCCCC GAATTACACT TAAACGACT 1850
CAACAGCTTA ACGTTGGCTT GCCACGCAAT ACTTGACTGT AAAACTCTCA 1900
45 CTCTTACCGA ACTTGGCCGT AACCTGCCAA CCAAGCGAG AACAAACAT 1950
AACATCAAC GAATCGACCG ATTGTTAGGT AATCGTCACC TCCACAAAGA 2000
GCGACTCGCT GTATACCGTT GGCATGCTAG CTTTATCTGT TCGGGAATAC 2050
GATGCCCATT GTACTGTGT ACTGGTCTGA TATTCGTGAG CAAAAACGAC 2100
TTATGGTATT GCGAGCTTCA GTCGCACTAC ACGGTCGTTT TGTACTCTT 2150
50 TATGAGAAAG CGTTCCCGCT TTCAGAGCAA TGTTCAAAGA AAGCTCATGA 2200
CCAATTTCTA GCCGACCTTG CGAGCATCTT ACCGAGTAAC ACCACACCG 2250
TCATTGTGAG TGATGTGGC TTTAAAGTGC CATGGTATAA ATCCGTTGAG 2300
AAGCTGGGT GGTACTGGTT AAGTCGAGTA AGAGGAAAAG TACAATATGC 2350
AGACCTAGGA GCGGAAACT GGAACCTAT CAGCAACTTA CATGATATGT 2400
55 CATCTAGTCA CTCAAAGACT TTAGGCTATA AGAGGCTGAC TAAAGCAAT 2450
CCAATCTCAT GCCAAATCTT ATTGTATAA TCTCGCTCTA AAGGCCGAAA 2500
AAATCAGCGC TCGACACGGA CTCATTGTCA CCACCCGTCA CCTAAATCT 2550
ACTCAGCGTC GGCAAAGGAG CCATGGGTTT TAGCAACTAA CTTACCTGTT 2600
GAAATTCGAA CACCCAAACA ACTTGTTAAT ATCTATTCGA AGCGAATGCA 2650
60 GATTGAAGAA ACCTTCCGAG ACTTGAAGAG TCCTGCCTAC GGACTAGGCC 2700
TACGCCATAG CCGAACGAGC AGCTCAGAGC GTTTTGATAT CATGCTGCTA 2750

109

ATLLIB02 135492.1

ATCGCCCTGA TGCTTCAACT AACATGTGG CTTCGCGGCG TTCTATGCTCA 2600
 GAAACAAGGT TGGGACAAGC ACTTCCAGGC TAACACAGTC AGAAATCGAA 2650
 ACGTACTCTC AACAGTTGCG TTAGGCATGG AAGTTTTCGG GCATTCTGGC 2900
 TACACAATAA CAAGGGAAGA CTACTCGTG GCTGCAACCC TACTAGCTCA 2950
 5 AAATTATTC ACACATGGTT ACGCTTTGGG GAAATTATGA TAATGATCCA 3000
 GATCACTTCT GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTGG 3050
 TTTTGTGTGG ATCTGCTGTG CTTCTAGTT GCCAGCCATC TGTGTGTGG 3100
 CCCTCCCGG TGCCITCCTT GACCCTGGAA GGTGCCACTC CCCTGTCTCT 3150
 TTCCTAATAA AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCTAT 3200
 10 CTATTCCTGG GGGTGGGGTG GGGCAGCACA GCAAGGGGGA GGATTGGGAA 3250
 GACAATAGCA GGCATGCTGG CGATCGGTC GCCTCTATGG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCGGTAC CTCTCTCTCT 3350
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCAGG TGCTGAAGAA 3400
 TTGACCCGGT GACCAAAGGT GCCTTTTATC ATCACTTTAA AAATAAAAAA 3450
 15 CAATTACTCA GTGCCGTGTA TAAGCAGCAA TTAATTATGA TTGATGCTTA 3500
 CATCACAAAC AAAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 TATTTGAACA TTATCTTGAT TATATTATTG ATAATAATAA AAACCTTATC 3600
 CCTATCCAAG AAGTGATGCC TATCATTGGT TGGAAATGAAC TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GGTAAGGTAA ACGCCATTGT CAGCAAATTG 3700
 20 ATCCAAGAGA ACCAACTTAA AGCTTTCTCTG ACGGAATGTT AATTCCTGTT 3750
 GACCTTGAGC ACTGATGAAT CCCCTAATGA TTTTGGTAAA AATCATTAA 3800
 TTAAGGTGGA TACACATCTT GTCATATGAT CCCGGTAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCAGG CTTTACACTT TATGCTTCGG GCTCGTATGT 3900
 TGTGTGGAAT TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC 3950
 25 CATGATTAGC CCAAGCGCGC AATTAACCTT CACTAAAGGG AACAAAAGCT 4000
 GGAGCTCCAC CGCGGTGGCG GCCGCTCTAG AACTAGTGA TCCCCCGGG 4050
 AGGTGAGAA GGTTCCTTTA CTGTTTGTCA ATTCTATTAT TTCAATACAG 4100
 AACAAATAGT TCTATAACTG AAATATATTT GCTATTGTAT ATTATGATTG 4150
 TCCCTCGAAC CATGAACACT CCTCCAGCTG AATTTACAA TTCTCTGTCT 4200
 30 ATCTGCCAGC CCATTAAAGT ATTCATGGAA GATCTTTGAG GAACACTGCA 4250
 AGTTCAATAT ATAAACACAT TTGAAATGGA GTATTGTTTT GCATTGTATG 4300
 GAGCTATGTT TTGCTGTATC CTCAGAAAAA AAGTTTGTGA TAAAGCATTC 4350
 ACACCCATAA AAAAGATAGT TTAATATTTC CAGCTATAGG AAAGAAAGTG 4400
 CGTCTGCTCT TCACCTTAGT CTCAGTTGGC TCCTTCACAT GCATGCTTCT 4450
 35 TTATTTCTCC TATTTTGTCA AGAAAAATAT AGGTACAGTC TTGTTCTCAC 4500
 TTATGTCTCG CCTAGCATGG CTCAGATGCA CGTTGTAGAT ACAAGAAGGA 4550
 TCAAAATGAA CAGACTTCTG GTCTGTTACT ACAACCATAG TAATAAGCAC 4600
 ACTAACTAAT AATGCTAAT TATGTTTTCC ATCTCTAAGG TTCACACATT 4650
 40 TTTCTGTTTT CTAAAGATC CCATTATCTG GTTGTAAGT AAGCTCAATG 4700
 GAACATGAGC AATATTTCCT AGTCTTCTCT CCCATCCAAC AGTCTGTATG 4750
 GATTAGCAGA ACAGGCAGAA AACACATTGT TACCCAGAAT TAAAAACTAA 4800
 TATTTGCTCT CCATTCAATC CAAATGGAC CTATTGAAAC TAAATCTTAA 4850
 CCCAATCCCA TTAATGATT TCTATGGCGT CAAAGGTCAA ACTTCTGAAG 4900
 45 GGAACCTGTG GGTGGGTAC NATTCAGGCT ATATATTCCC CAGGGCTCAG 4950
 CGGATCTCCA TGGGCTCCAT CGGTGCAGCA AGCATGGAAT TTTGTTTTGA 5000
 TGTATTCAGG GAGCTCAAAG TCCACCATGC CAATGAGAAC ATCTTCTACT 5050
 GCCCATTCGC CATCATGTCA GCTCTAGCCA TGGTATACCT GGGTGCAAAA 5100
 GACAGCACCA GGAATTCGT GCGCTCCTCC AAGAACGTCA TCAAGGAGTT 5150
 50 CATGCGCTTC AAGGTGCGCA TGGAGGCGAC CGTGAACGGC CACGAGTTCT 5200
 AGATCGAGGG CGAGGGCGAG GCGCGCCCTC ACGAGGGCCA CAACCCGCTG 5250
 AAGCTGAAGG TGACCAAGGG CGGCCCTCTG CCCTTCGCTT GGGACATCCT 5300
 GTCCCCCCAG TTCCAGTACG GCTCCAAGGT GTACGTGAAG CACCCGCGCG 5350
 ACATCCCCGA CTACAAAGAG CTGTCTTCTC CCGAGGGCTT CAAGTGGGAG 5400
 CGCGTGATGA ACTTCAGGGA CGGCGGCGTG GTGACCGTGA CCCAGGACTC 5450
 55 CTCCCTGCAG GACGGCTGCT TCACTACAAA GGTGAAGTTC ATCGGCGTGA 5500
 ACTTCCCTCC CGACGGCCCC GTAATGCAGA AGAAGACCAT GGGCTGGGAG 5550
 GCCTCCACCG AGCGCTGTGA CCCCAGCGAC GCGGTGTGTA AGGGCGAGAT 5600
 CCACAAGGCC CTGAAGCTGA AGGACGCGCG CCACTACCTG GTGAGTTTCA 5650
 AGTCCATCTA CATGGCCAAG AAGCCCGTGC AGCTGCCCGG CTACTACTAC 5700
 60 GTGCACTCCA AGCTGCACAT CACCTCCAC AACGAGGACT ACACCATCGT 5750
 GGAGCAGTAC GAGCGCACCG AGGGCCGCCA CCACCTGTTT CTGTAGCGGC 5800

CGCGACTCTA GATCATAATC AGCCATACCA CATTTGTAGA GGTTTTACTT 5850
 GCTTTAAAAA ACCTCCCACA CCTCCCCCTG AACCTGAAAC AAAAAATGAA 5900
 TGCAATTGTT GTTGTTAACT TGTTTATTGC AGCTTATAAT GGTACAAAT 5950
 AAAGCAATAG CATCACAAT TCCACAAATA AAGCATTITI TCACTGCAT 6000
 5 TCTAGTTGTG GCTCGAGAAG GCGAATTCT GCAGATATCC ATCACACTGG 6050
 CGGCCGCTCG AGGGGGGGCC CGGTACCCAA TTCGCCCTAT AGTCAGTCGT 6100
 ATTACGCGCG CTCACTGGCC GTCGTTTAC AACGTCGTGA CTGGGAAAAC 6150
 CCTGGCGTTA CCCAACTTAA TCGCCTTGCA GCACATCCCC CTTTCGCCAG 6200
 CTGGCGTAAT AGCGAAGAGG CCGGCACCGA TCGCCCTTCC CAACAGTTGC 6250
 10 GCAGCCTGAA TGGCGAATGG AATTGTAAAG CGTTAATATT TTGTTAAAA 6300
 TCGCGTTAAA TTTTGTAA ATCAGCTCAT TTTTAAACCA ATAGGCCGAA 6350
 ATCGGCAAAA TCCTTATAA ATCAAAGAA TAGACCGAGA TAGGGTTGAG 6400
 TGTGTTCCTA GTTTGGAACA AGAGTCCACT ATTAAGAAC GTGGACTCCA 6450
 ACGTCAAAGG GCGAAAAACC GTCTATCAGG GCGATGGCCC ACTACTCCGG 6500
 15 GATCATATGA CAAGATGTGT ATCCACCTTA ACTTAATGAT TTTTACCAA 6550
 ATCATTAGGG GATTTCATCAG TGCTCAGGGT CAACGAGAAT TAACATTCGG 6600
 TCAGGAAAGC TTATGATGAT GATGTGCTTA AAACTTACT CAATGGCTGG 6650
 TTATGCATAT CGCAATACAT GCGAAAAACC TAAAGAGCT TGCCGATAAA 6700
 AAAGGCCAAT TTATTGCTAT TTACCGCGGC TTTTATTGA GCTTGAAGA 6750
 20 TAAATAAAAT AGATAGGTTT TATTTGAAGC TAAATCTTCT TTATCGTAAA 6800
 AAATGCCCTC TTGGGTATC AAGAGGGTCA TTATATTTCG CGGAATAACA 6850
 TCATTGCTG ACAGAAATAC TAAGCACTTG TCTCCTGTTT ACTCCCTGA 6900
 GCTTGAGGGG TTAACATGAA GGTATCTGAT AGCAGGATAA TAATACAGTA 6950
 AAACGCTAAA CCAATAATCC AATCCAGGCC ATCCCAATT GGTAGTGAAT 7000
 25 GATTATAAAT AACAGCAAC AGTAATGGGC CAATAACACC GGTGTCATTG 7050
 GTAAGGCTCA CCAATAATCC CTGTAAAGCA CCTTGCTGAT GACTCTTTGT 7100
 TTGATAGAC ATCACTCCCT GTAATGCAAG TAAAGCGATC CCACCCAG 7150
 CCAATAAAT TAAACAGGG AAACTAACCC AACCTTCAGA TATAACGCT 7200
 30 AAAAAAGCAA ATGCACTACT ATCTGCAATA AATCCGAGCA GTACTGCCGT 7250
 TTTTTCGCC CATTTAGTGG CTATCTTCC TGCCACAAAG GCTTGAATA 7300
 CTGAGTGTAA AAGACCAAGA CCGCTAATG AAAAGCCAAC CATCATGCTA 7350
 TTCCATCCAA AACGATTTC GGTAAATAGC ACCACACCG TTGCGGAAT 7400
 TTGCGCTATC AATTGCGCTG AAAAAATAAT AATCAACAA ATGGCATCGT 7450
 35 TTTAAATAAA GTGATGTATA CCGAATTGAG CTTTGTCTCC CTTIAGTGAG 7500
 GGTAAATTGC GCGCTTGGCG TAATCATGGT CATAGCTGTT TCCTGTGTGA 7550
 AATTGTTATC CGCTCACAAT TCCACACAAC ATACGAGCCG GAAGCATAAA 7600
 GTGTAAAGCC TGGGGTGCCT AATGAGTGAG CTAACCTACA TTAATTGCGT 7650
 TGCGCTCACT GCCCGCTTTC CAGTCGGGAA ACCTGTCGTG CCAGCTGCAT 7700
 40 TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTGTGCGTA TTGGGCGCTC 7750
 TTCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGCTCGTT CGGCTGCGGC 7800
 GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA 7850
 GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG 7900
 GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGC CTCGCCCCC 7950
 45 CTGACGAGCA TCACAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG 8000
 ACAGGACTAT AAGATACCA GGCSTTTCCC CCTGGAAGCT CCCTCGTGGC 8050
 CTCCTCTGTT CCGACCTGCG CGCTTACCGG ATACCTGTCC GCCTTTCTCC 8100
 CTTCCGGAAG CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTAICTCAGT 8150
 TCGGTGTAGT TCGTTCGCTC CAAGCTGGGC TGTGTGCAGC AACCCCCGT 8200
 TCAGCCGAC CGCTGCGCT TATCGGTAA CTATCGTCTT GAGTCCAACC 8250
 50 CCGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT 8300
 AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTCTTGA AGTGGTGGCC 8350
 TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGTGTA 8400
 AGCCAGTTAC CTTCCGAAAA AGAGTTGGTA GCTCTTGATC CGGCAACAA 8450
 55 ACCACCGCTG GTAGCGGTGG TTTTGTGTT TGCAAGCAGC AGATTACGCG 8500
 CAGAAAAAAA GGAICTCAAG AAGATCCTTT GATCTTTTCT ACGGGCTCTG 8550
 ACGCTCAGTG GAACGAAAC TCACGTTAAG GGAITTTGGT CATGAGATTA 8600
 TCAAAAAGGA TCTTACCTA GATCCTTTTA AATTAAAAAT GAAGTTTAA 8650
 ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT 8700
 TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTATTTCG TTCACTCCATA 8750
 60 GTTGCCTGAC TCCCGCTCGT GTAGATAACT ACGATACGGG AGGGCTTACC 8800
 ATCTGGCCCC AGTGTGCAA TGATACCGCG AGACCCAGC TCACCGGCTC 8850

CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT 8900
 GGTCTGCAA CTTTATCCGC CTCATCCAG TCTATTAAAT GTTGCCGGA 8950
 AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTGCCA 9000
 TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATTC 9050
 5 AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG 9100
 CAAAAAAGCG GTTAGCTCCT TCGGTCTCTC GATCGTTGTC AGAAGTAAGT 9150
 TGGCCGCGAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT 9200
 ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC 9250
 CAAGTCATTC TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCGG 9300
 10 CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC 9350
 ATCATTGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA TCTTACCGCT 9400
 GTTGAGATCC AGTTCGATGT AACCCACTCG TGACCCCAAC TGATCTTCAG 9450
 CATCTTTTAC TTTACCACGC GTTCTGGGT GAGCAAAAAC AGGAAGGCAA 9500
 AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAATGTT GAATACTCAT 9550
 15 ACTCTCTCTT TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA 9600
 TGAGCGGATA CATATTGAA TGTATTAGA AAAATAAACA AATAGGGGTT 9650
 CCGCGCACAT TTCCCCGAAA AGTGCCAC 9678

20 SEQ ID NO:4 (PTnMod (Oval/Red) Quail)

CTGACCGCCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGTTACG 50
 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTGCG 100
 TTTCTCCCT TCCTTTCTCG CCACGTTCCG CGGCATCAGA TTGGCTATTG 150
 25 GCCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200
 CATGTCCAAC ATTACGCCCA TGTGACATT GATTATTGAC TAGTATTAA 250
 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCCG 300
 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC 350
 CCGGCCCAT TACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA 400
 30 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAC TGCCCA 450
 CTGCGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC CCTATTGACG 500
 TCAATGACGG TAAATGGCCC GCCTGGCAT TATGCCAGTA CATGACCTTA 550
 TGGGACTTTC CTACTTGCCA GTACATCTAC GTATTAGTCA TCGCTATTAC 600
 CATGCTGATG CGGTTTGGC AGTACATCAA TGGGCGTGA TAGCGGTTG 650
 35 ACTCAGGGG ATTTCCAAGT CTCACCCCA TTGACGTCAA TGGGAGTTG 700
 TTTGGCACC AAAATCAACG GGACTTTCCA AAATGTGTA ACAACTCCG 750
 CCCATTGACG CAAATGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATA 800
 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACG 850
 TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCCGGCCG 900
 40 GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAGAGT GACGTAAGTA 950
 CCGCCTATAG ACTCTATAG CACACCCCTT TGGCTCTTAI GCATGCTATA 1000
 CTGTTTTTGG CTTGGGGCT ATACACCCCT GCTTCTTAT GCTATAGGTG 1050
 ATGGTATAGC TTAGCCTATA GGTGTGGGT ATTGACCAT ATTGACCAT 1100
 CCCCATTGG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTTT 1150
 45 GCCACAAC TA CTCTATTGG CTATATGCCA ATACTCTGTC CTTGAGAGAC 1200
 TGACACGGAC TCTGTATTT TACAGGATGG GGTCCCATTT ATTATTACA 1250
 AATTACATA TACAACAACG CGTCCCCG TGCCCGCAGT TTTTATTA 1300
 CATAGCGTGG GATCTCCACG CGAATCTCG GTACGTGTT CCGACATGG 1350
 CTCTTCTCCG GTAGCGGCG AGCTTCCACA TCCGAGCCCT GGTCCCATGC 1400
 50 CTCCAGCGGC TCATGCTCG TCGGCGAGTC CTGTCTCTTA ACAGTGGAG 1450
 CCAGACTTAG GCACAGCACA ATGCCACCA CCACAGTGT GCGCACAAG 1500
 GCCGTGGCGG TAGGTATGT GTCTGAAAT GAGCGTGGAG ATTGGGCTCG 1550
 CACGGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCAG 1600
 GCAGCTGAGT TGTGTATTC TGATAAGAGT CAGAGGTAAC TCCGTTGCG 1650
 55 GTGCTGTAA CGGTGGAGG CAGTGTAGT TGAGCAGTAC TCGTTGTGTC 1700
 GCGCGCGCC ACCAGACATA ATAGCTGACA GACTACAGA CTGTTCTTT 1750
 CCATGGGTCT TTTCTGAGT CACCGTCGA CCATGTGTGA ACTTGATATT 1800
 TTACATGATT CTCTTTACCA ATTCTGCCCC GAATTACACT TAAACGACT 1850
 CAACAGCTTA ACGTTGGCTT GCCACGCATT ACTTGACTGT AAAACTCTCA 1900
 60 CTCTTACCGA ACTTGGCGT AACCTGCCAA CCAAGCGAG AACAAACAT 1950
 AACATCAAC GAATCGACCG ATTGTTAGT AATGTCACC TCCACAAAGA 2000

GCGACTCGCT GTATACCGTT GGCATGCTAG CTTTATCTGT TCGGGAATAC 2050
 GATGCCCATT GTACTTGTG ACTGGTCTGA TATTCGTGAG CAAAAACGAC 2100
 TTATGGTATT GCGAGCTTCA GTCGCACTAC ACGGTCTGTC TGTTACTCTT 2150
 TATGAGAAAG CGTTCGCCCT TGCAGAGCAA TGTTCAAAGA AAGTCATGA 2200
 CCAATTTCTA GCGGACCTTG CGAGCATTTCT ACCGAGTAAC ACCACACCGC 2250
 TCATTGTCAG TGATGCTGGC TTTAAAGTGC CATGGTATAA ATCCGTGAG 2300
 AAGCTGGGTT GGTACTGGTT AAGTCGAGTA AGAGGAAAAG TACAATATGC 2350
 AGACCTAGGA GCGGAAAAC TGAACCTAT CAGCAACTTA CATGATATGT 2400
 CATCTAGTCA CTCAAAAGACT TTAGGCTATA AGAGGCTGAC TAAAAGCAAT 2450
 CCAATCTCAT GCCAAATTTCT ATTGTATAAA TCTCGTCTA AAGGCCGAAA 2500
 AAATCAGCGC TCGACACGGA CTCATTGTCA CCACCCGTCA CCTAAAATCT 2550
 ACTCAGCGTC GGCAAAAGGAG CCATGGGTTT TAGCAACTAA CTTACCTGTT 2600
 GAAATTCGAA CACCCAAACA ACTTGTTAAT ATCTATTCGA AGCGAATGCA 2650
 GATTGAAGAA ACCTTCCGAG ACTTGAAAAG TCTGCTTAC GGACTAGGCC 2700
 TACGCCATAG CCGAAGCAGC AGCTCAGAGC GTTTTGATAT CATGCTGCTA 2750
 ATCGCCCTGA TGCTTCAACT AACATGTTGG CTTCGGGGCG TTCATGCTCA 2800
 GAAACAAGGT TGGGACAAGC ACTTCCAGGC TAACACAGTC AGAAATCGAA 2850
 ACGTACTCTC AACAGTTCCG TTAGGCATGG AAGTTTTCGG GCATTCGTGC 2900
 TACACAATAA CAAGGGAAGA CTACTCTGTG GCTGCAACCC TACTAGCTCA 2950
 AAATTTATTC ACACATGGTT ACGCTTTGGG GAAATTATGA TAATGATCCA 3000
 GATCACTTCT GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTTGG 3050
 TTTTTTGTGG ATCTGCTGTG CCTCTAGTT GCCAGCCATC TGTGTGTTGC 3100
 CCCTCCCGCG TGCTTCCCTT GACCCGTGAA GGTGCCACTC CCACTGTCTT 3150
 TTCTAATAAA AATGAGGAAA TTGCATCGCA TTGCTCAGT AGGTGTCTAT 3200
 CTATTCTGGG GGTGGGGTGG GGGCAGCACA GCAAGGGGGA GGATTGGGAA 3250
 GACAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT 3350
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCAGG TGCTGAAGAA 3400
 TTGACCCGGT GACCAAAGGT GCCTTTTATC ATCACTTTAA AAATAAAAAA 3450
 CAATTACTCA GTGCCTGTGA TAAGCAGCAA TTAATTATGA TTGATGCTCA 3500
 CATCACAAACA AAAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 TATTTGAACA TTATCTTGAT TATATTATTG ATAATAATAA AAACCTTATC 3600
 CCTATCCAAG AAGTGATGCC TATCATTGGT TGGAAATGAAC TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GGTAAAGTAA ACGCCATTGT CAGCAAAATTG 3700
 ATCCAAAGAA ACCCACTTAA AGCTTTCTCG ACGGAATGTT AATTCTCGTT 3750
 GACCCTGAGC ACTGATGAAT CCCCTAATGA TTTTGGTAAA AATCATTAA 3800
 TTAAGGTGGA TACACATCTT GTCATATGAT CCGGTAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCAGG CTTTACACTT TATGCTTCCG GCTCGTATGT 3900
 TGTGTGGAAT TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC 3950
 CATGATTAAC CCAAGCGCGC AATTAAACCT CACTAAAGGG AACAAAAGCT 4000
 GGAGCTCCAC CGCGGTGGCG GCCGCTCTAG AACTACTGGA TCCCCCGGGG 4050
 AGGTCAAGAT GGTTCCTTTA CTGTTGTCA ATTCTATTAT TTCAATACAG 4100
 AACAAAAGCT TCTATAACTG AAATATATTT GCTATTGTAT ATTATGATTG 4150
 TCCTCGAAC CATGAACACT CCTCCAGCTG AATTTACAAA TCCTCTGTCT 4200
 ATCTGCCAGG CTGGAAGATC ATGGAAGATC TCTGAGGAAC ATGCAAGATT 4250
 CATACCAATA ACTCAATTGG AATTGAGTAT TATTTTGCTT TGAATGGAGC 4300
 TATGTTTTCG AGTTCCCTCA GAAGAAAGC TTGTTATAAA GCGTCTACAC 4350
 CCATCAAAG ATATATTTAA ATATTCCAAC TACAGAAAGA TTTTGTCTGC 4400
 TCTTCACTCT GATCTCAGTT GGTTCCTTCA CGTACATGCT TCTTTATTTG 4450
 CCTATTTTGT CAAGAAAATA ATAGGTCAAG TCCTGTTCTC ACTTATCTCC 4500
 TGCCTAGCAT GGCTTAGATG CACGTTGTAC ATTCAAGAAG GATCAAATGA 4550
 AACAGACTTC TGGTCTGTGA CAACAACCAT AGTAATAAAC AGACTAACTA 4600
 ATAATTGCTA ATTATGTTT CCATCTCTAA GGTTCACACA TTTTCTGTT 4650
 TTAAGATCCC ATTTATCTGT TGTAACTGAA GCTCAATGGA ACATGAACAG 4700
 TATTTCTCAG TCTTTTCTCC AGCAATCTCG ACGGATTAGA AGAACTGGCA 4750
 GAAAACACTT TGTACCCAG AATTAAAAAC TAATATTTGC TCTCCCTTCA 4800
 ATCCAAAATG GACCTATTGA AACTAAAATC TGACCCAATC CCATTAAATT 4850
 ATTTCTATGG CGTCAAAGGT CAAACITTTG AAGGGAACTT GTGGGTGGGT 4900
 CCCAATTCAG GCTATATATT CCCCAGGGCT CAGCGGATCT CCATGGGCTC 4950
 CTCGTGCAGC AAGCATGGAA TTTTGCCTTG ATGTATTCAA GGAGCTCAAA 5000
 GTCCACCATG CCAATGACAA CATGCTCTAC TCCCCTTTG CCATCTGTCA 5050

ACTCTGGCCA TGGTCTCCCT GGGTGCAAAA GACAGCACCA GGGAAATTCGT 5100
 GCGCTCCTCC AAGAAGCTCA TCAAGGAGTT CATGCGCTTC AAGGTGCGCA 5150
 TGGAGGGCAC CGTGAACGGC CACGAGTTCG AGATCGAGGG CGAGGGCGAG 5200
 GGCCGCCCTT ACGAGGGCCA CAACACCGTG AAGCTGAAGG TGACCAAGGG 5250
 5 CGGCCCGCTG CCCTTGGCCT GGGACATCCT GTCCCCCAG TTCCAGTACG 5300
 GCTCCAAGGT GTACGTGAAG CACCCCGCGG ACATCCCCGA CTACAAGAAG 5350
 CTGTCTCTCC CCGAGGGCTT CAAGTGGGAG CCGGTGATGA ACTTCGAGGA 5400
 CGGCGGCGTG GTGACCGTGA CCCAGSACTC CTCCCTGCAG GACGGCTGCT 5450
 TCATCTACAA GGTGAAGTTC ATCGGCGTGA ACTTCCCCTC CGACGGCCCC 5500
 10 GTAATGCAGA AGAAGACCAT GGGCTGGGAG GCCTCCACCG AGCGCTGTA 5550
 CCCCCGCGAC GCGGTGCTGA AGGGCGAGAT CCACAAGGCC CTGAAGCTGA 5600
 AGGACGGCGG CCACTACCTG GTGGAGTTCA AGTCCATCTA CATGGCCAAG 5650
 AAGCCCGTGC AGCTGCCCGG CTACTACTAC GTGGACTCCA AGCTGGACAT 5700
 CACCTCCAC AAGAGGACT ACACCATCGT GGAGCAGTAC GAGCGCACCG 5750
 15 AGGGCCGCCA CCACCTGTTC CTGTAGCGGC CGCGACTCTA GATCATAATC 5800
 AGCCATACCA CATTTGTAGA GGTTTTACTT GCTTTAAAAA ACCTCCGACA 5850
 CCTCCCCCTG AACCTGAAAC ATAAAAATGAA TGCAATTGTT GTTGTAACT 5900
 TGTATTATGC AGCTTATAAT GGTACAAAT AAAGCAATAG CATCACAAT 5950
 TTCACAAATA AAGCAITTTT TTCACCTCAT TCTAGTTGTG GCTCGAGAAG 6000
 20 GGGCAATTCT CGAGATATCC ATCACACTGG CCGCGCTCG AGGGGGGGCC 6050
 CGGTACCCAA TTCGCCCTAT AGTGAGTCGT ATTACGCGCG CTCACTGGCC 6100
 TCGCTTTTAC AACGTCGTGA CTGGGAAAAC CCTGGCGTTA CCAACTTAA 6150
 TCGCCTTGCA GCACATCCCC CTTCGCCAG CTGGCGTAAT AGCGAAGAGG 6200
 CCGCACCGA TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG 6250
 25 AAATTGTAAAG CGTTAATATT TTGTTAAAAA TCGCGTTAAA TTTTGTAA 6300
 ATCAGCTCAT TTTTAAACCA ATAGGCCGAA ATCGGCAAAA TCCCTTATAA 6350
 ATCAAAGAA TAGACCGAGA TAGGGTTGAG TGTGTTTCCA GTTTGGAACA 6400
 AGAGTCCACT ATTAAGAAGC GTGGACTCCA ACGTCAAAGG GCGAAAAACC 6450
 GTCTATCAGG GCGATGGCCC ACTACTCCGG GATCATATGA CAAGATGTGT 6500
 30 ATCCACCTTA ACTTAATGAT TTTTACCAAA ATCATTAGGG GATTTCATCAG 6550
 TGCTCAGGGT CAACGAGAAT TAACATTCG TCAGGAAAGC TTATGATGAT 6600
 GATGTGCTTA AAAACTTACT CAATGGCTGG TTATGCATAT CGCAATACAT 6650
 GCGAAAAACC TAAAGAGCT TSCCGATAAA AAAGGCCAAT TTATTGCTAT 6700
 TTACCGCGGC TTTTATTGA GCTTGAAAGA TAAATAAAT AGATAGGTTT 6750
 35 TATTTGAAGC TAAATCTTCT TTATCGTAAA AAATGCCCTC TTGGGTATC 6800
 AAGAGGGTCA TTATATTTCG CGGAATAACA TCATTGGTG ACGAAATAAC 6850
 TAAGCACTTG TCTCCTGTTT ACTCCCCTGA GCTTGAGGGG TTAACATGAA 6900
 GGTCACTGAT AGCAGGATAA TAATACAGTA AAACGCTAAA CCAATAATCC 6950
 40 AAATCCAGCC ATCCCAAATT GGTAGTGAAT GATTATAAAT AACAGCAAAC 7000
 AGTAATGGGC CAATAACACC GGTGCTATTG GTAAGGCTCA CCAATAATCC 7050
 CTGTAAAGCA CCTTGCTGAT GACTCTTTGT TTGATAGAC ATCACTCCCT 7100
 GTAATGCAGG TAAAGCGATC CCACCACCAG CCAATAAAT TAAACAGGG 7150
 AAAACTAACC AACCTTCAGA TATAAACGCT AAAAAGGCAA ATGCACTACT 7200
 45 ATCTGCAATA AATCCGAGCA GTACTGCCGT TTTTTCGCC CATTTAGTGG 7250
 CTATTCTTCC TGCCACAAG GCTTGGAATA CTGAGTGTA AAGACCAAGA 7300
 CCGCTAATG AAAAGCCAAC CATCATGCTA TTCCATCCAA AACGATTTTC 7350
 GGTAAATAGC ACCCACACCG TTGCGGGAAT TTGGCCTATC AATTGCGCTG 7400
 AAAATAAAT AATCAACAAA ATGGCATCGT TTTAAATAAA GTGATGTATA 7450
 50 CCGAATTCAG CTTTGTGTTCC CTTTGTGAG GGTAAATGC GCGCTTGGCG 7500
 TAATCATGGT CATAGCTGTT TCCTGTGTGA AATTGTTATC CGCTCACAAT 7550
 TCCACACAAC ATACGAGCCG GAAGCATAAA GTGTAAAGCC TGGGTGCTCT 7600
 AATGAGTGAG CTAACCTACA TTAATTGCGT TCGCTCACT GCCCGCTTTC 7650
 CAGTCGGGAA ACCTGTCTGT CCAGCTGCAT TAATGAATCG GCCAACGCGC 7700
 GGGGAGAGGC GGTGTGCGTA TTGGGCGCTC TCCGCTTCC TCGCTCACTG 7750
 55 ACTCGCTGCG CTCGGTCTGT CGGCTGCGGC GAGCGGTATC AGCTCACTCA 7800
 AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA 7850
 CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCGT 7900
 TGCTGGCGTT TTTCCATAGG CTCGCCCTCC CTGACGAGCA TCACAAAAT 7950
 CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATACCA 8000
 60 GCGGTTTCCC CCTGGAAGCT CCTCGTGGC CTCTCCTGTT CCGACCTGCG 8050
 CGCTTACCGG ATACCTGTCC GCCTTTCTCC TTTCGGGAAG CGTGGCGCTT 8100

TCTCATAGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTGCTC 8150
 CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC CGCTGCGCCT 8200
 TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAGACA CGACTTATCG 8250
 CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA GGTATGTAGG 8300
 5 CCGTGCTACA GAGTTCTTGA AGTGGTGGCC TAACTACGGC TACACTAGAA 8350
 GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC CTTCCGAAAA 8400
 AGAGTTGGTA GCTCTTGATC CGGCAAAACA ACCACCGCTG GTAGCGGTGG 8450
 TTTTITTTGT TGCAAGCAGC AGATTACGGC CAGAAAAAAA GGATCTCAAG 8500
 AAGATCCTTT GATCTTTTCT ACGGGGTCTG ACGCTCAGTG GAAACGAAAA 8550
 10 TCACGTTAAG GGATTTTGGT CATGAGATTA TCAAAAAGGA TCTTCACCTA 8600
 GATCCTTTTA AATTAAAAAT GAAGTTTAA ATCAATCTAA AGTATATATG 8650
 AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTGA GGCACCTATC 8700
 TCAGCGATCT GTCTATTTCT TTCAATCCATA GTTGCCCTGAC TCCCCGTCTG 8750
 GTAGATAACT ACGATACGGG AGGGCTTACC ATCTGGCCCC AGTGCTGCAA 8800
 15 TGATACCGCG AGACCCACGC TCACCGGCTC CAGATTATC AGCAATAAAC 8850
 CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT GGTCCCTCAA CTTTATCCGC 8900
 CTCATCCAG TCTATTAAAT GTTGCCGGGA AGCTAGAGTA AGTAGTTCGC 8950
 CAGTTAATAG TTTGCGCAAC GTTGTGCGA TTGCTACAGS CATCGTGGTG 9000
 TCACGCTCGT CGTTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC 9050
 20 AAGCGGAGTT ACATGATCCC CCATGTTCTG CAAAAAGCG GTTAGCTCCT 9100
 TCGTCTCTCC GATCGTTGTC AGAAGTAAAT TGGCCGCGT GTTATCACTC 9150
 ATGGTTATGG CAGCACTGCA TAATTCTCTT ACTGTCATGC CATCCGTAAG 9200
 ATGCTTTTCT GTGACTGGTG AGTACTCAAC CAAGTCATTC TGAGAATAGT 9250
 GTATGCGGCG ACCGAGTTGC TCTTGCCCGG CGTCAATACG GGAATAATAC 9300
 25 GCGCCACATA GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC 9350
 GGGGCGAAAA CTCTCAAGGA TCTTACCGCT GTTGAGATCC AGTTCGATGT 9400
 AACCCACTCG TGCAACCAAC TGATCTTCAG CATCTTTTAC TTTCAACGAC 9450
 GTTCTGGGT GAGCAAAAAC AGGAAGGCAA AATGCCGCAA AAAAGGGAAT 9500
 AAGGGCGACA CGGAATGTT GAATACTCAT ACTCTTCCTT TTTCAATATT 9550
 30 ATTGAAGCAT TTATCAGGGT TATTGTCTCA TGAGCGGATA CATATTGTAA 9600
 TGTATTTAGA AAAATAAACA AATAGGGGTT CCGCGCACAT TTCCCGGAAA 9650
 AGTCCAC 9658

35 SEQ ID NO:5 (spacer)
 (GPGG)_x

40 SEQ ID NO:6 (spacer)
 GPGGGPGGPGG

45 SEQ ID NO:7 (spacer)
 GGGSGGGSGGGGS

50 SEQ ID NO:8 (spacer)
 GGGSGGGSGGGSGGGGS

55 SEQ ID NO:9 (enterokinase cleavage site)
 DDDDK

60 SEQ ID NO:10 (altered transposase Hef forward primer)
 ATCTCGAGACCATGTGTGAAGTGAATTTTACATGATTCTCTTACC

SEQ ID NO:11 (altered transposase Her reverse primer)
 GATTGATCATTATCATAATTTCCCAAAGCGTAACC

SEQ ID NO:12 (Xho I restriction site)
CTCGAG

5 SEQ ID NO:13 (modified Kozak sequence)
ACCATG

SEQ ID NO:14 (Bcl I restriction site)
10 TGATCA

SEQ ID NO:15 (CMVf-NcoM IV primer)
15 TTGCCGGCATCAGATTGGCTAT

SEQ ID NO:16 (Syn-polyAr-BstE II primer)
AGAGGTCACCGGGTCAATTCTTCAGCACCTGGTA

20 SEQ ID NO:17 (vitellogenin promoter)

TGAATGTGTT CTTGTGTTAT CAATATAAAT CACAGTTAGT GATGAAGTTG GCTGCAAGCC
25 TGCATCAGTT CAGCTACTTG GCTGCATTTT GTATTGTTT CTGTAGGAAA TGCAAAAGGT
TCTAGGCTGA CTTGCACITC TATCCCTCTT GCCTTACTGC TGAGAATCTC TGCAGGTTTT
AATTGTTTAC ATTTTGCTCC CATTCTACTTT GGAAGATAAA ATATTTACAG AATGCTTATG
AAACCTTTGT TCATTAAAA ATATTCTCTG TCAGCGTGAC CGGAGCTGAA AGAACACATT
GATCCCGTGA TTCAATAAAA TACATATGTT CCATATATTG TTCTCAGTA GCCTCTTAAA
30 TCATGTGCGT TGGTGCACAT ATGAATACAT GAATAGCAAA GGTTTATCTG GATTACGCTC
TGGCCTGCCG GAATGGCCAT AAACCAAAAGC TGAGGGAAGA GGGAGAGTAT AGTCAATGTA
GATTATACG ATTGCTGATT GGGTTATTAT CAGCTAGATA ACAACTTGGG TCAGGTGCCA
GGTCAACATA ACCTGGGCAA AACCACTCTC ATCTGTGGCA GGACCATGTA CCAGCAGCCA
35 GCCGTGACCC AATCTAGGAA AGCAAGTAGC ACATCAATTT TAAATTTATT GTAAATGCCG
TAGTAGAAGT GTTTTACTGT GATACATTGA AACTTCTGGT CAATCAGAAA AAGGTTTTTT
ATCAGAGATG CCAAGGTATT ATTTGATTTT CTTTATTCGC CGTGAAGAGA ATTTATGATT
GCAAAAAGAG GAGTGTTTAC ATAACTGAT AAAAACTTG AGGAATTCAG CAGAAAAACAG
CCACGTGTC CTGAACATTC TTCCATAAAA GTCTACCAT GCCTGGCAGA GCCCTATTCA
40 CCTTCGCT

SEQ ID NO:18 (vitellogenin targeting sequence)

ATGAGGGGGATCATACTGGCATTAGTGCTCACCCTTGTAGGCAGCCAGAAAGTTTGACATTGGT

45 SEQ ID NO:19 (p146 protein)
KYKKALKKLAKLL

50 SEQ ID NO:20 (p146 coding sequence)
AAATACAAAAAGCACTGAAAAAAGTGGCAAACTGCTG

SEQ ID NO:21 (pro-insulin sequence)

55 TTTGTGAACCAACCTGTGCGGCTCACACCTGGTGGAAAGCTCTCTACCTAGTGTGCGGGAAACGAGGC
TTCTTCTACACACCCCAAGACCCGCGGGAGGCAGAGGACCTGCAGGTGGGGCAGGTGGAGCTGGCGGG
GGCCCTGGTGCAGGCAGCTGCAGCCCTTGGCCCTGGAGGGGTCCCTGCAGAAGCGTGGCATTTGTGGAA
CAATGCTGTACCAGCATCTGCTCCCTCTACAGCTGGAGAACTCTGCAACTAG

60 SEQ ID NO:22 (TAG sequence)

SEQ ID NO:25 (synthetic polyA sequence)
GGCGCTCGATACGACTCTTCTGGCTCAATAAAAGCATCAGAGCTCTAGAGATCTGTCTGTTGGTTTTT
TGTGGATCTTCGTGCTCTCTAGTTGCAGCCACTCTGTGGTTTGCCTCTCCCGTGCCTCTCTTGACC
CTGGAAGAGTGCACATCCCACTCTCTTCTCTCAATAAAATAGGAAATTCATCGCATCTCTGTAGTAGG
TGTCAATCTATTCTGGGGGGTGGGGTGGAGCAGCAGCAGAGGGGAGGATTTGGGAAGCAATAGTAGG
CATCTCGGGATGCGGTGGGCTCTATGGGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
TCTCGGTACTCTCTC

SEQ ID NO:29 (pTnMod(Oval/ENT tag/P146/PA) - Chicken)

```

5  CTGACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG 50
   CGCAGCGTGA CCGCTACACT TGCAGCGGCC CTAGCGCCCG CTCCTTTCGC 100
   TTTCTTCCCT TCCTTCTCTG CCACGTTCCG CGGCATCAGA TTGGCTATTG 150
   GCCATTGCAI ACGTTGTATC CATATCATAA TATGTACATI TATATTGGCT 200
   CATGTCCAAC ATTACCGCCA TGTGACATT GATTATTGAC TAGTATTAA 250
   TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCGG 300
10 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC 350
   CCCGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA 400
   GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT AAACGTGCCA 450
   CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC CCTATTGACG 500
   TCAATGACGG TAAATGGCCC GCTTGGCATT ATGCCAGTA CATGACCTTA 550
15 TGGGACTTTC CTACTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC 600
   CATGGTGATG CGGTTTGGC AGTACATCAA TGGGCGTGA TAGCGGTTG 650
   ACTCACGGGG AITTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTG 700
   CTTGGCACC AAAATCAACG GGAATTTCCA AAATGTCGTA ACAACTCCGC 750
   CCCATTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA 800
20 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC 850
   TCTTTTGACC TCCATAGAAG ACACGGGAC CGATCCAGCC TCCCGGGCG 900
   GGAACGGTGC ATTGGAACGC GGAATCCCGG TGCCAAGAGT GACCTAAGTA 950
   CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA 1000
   CTGTTTTTGG CTGGGGCCCT ATACACCCCG GCTTCCITAT GCTATAGGTG 1050
25 ATGGTATAGC TTAGCCTATA GGTGTGGGTT ATTCAACCAT ATTGACCACT 1100
   CCCCTATTGG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTTT 1150
   GCCACAACATA TCTCTATTGG CTATATGCCA ATACTCTGTC CTTGAGAGAC 1200
   TGACACGGAC TCTGTATTTT TACAGGATGG GGTCCCATTT ATTTATTACA 1250
   AATTACATA TACAACAACG CCGTCCCCCG TGCCCGCAGT TTTTATTAAA 1300
30 CATAGCGTGG GATCTCCACG CGAATCTCGG GTACGTGTTT CGGACATGGG 1350
   TCTTTCTCCG GTAGCGGCGG AGCTTCCACA TCCGAGCCCT GGTCCCATGC 1400
   CTCAGCGGCG TCATGGTCCG TCGGCGAGTC CTGTCTCTTA ACAGTGGAGG 1450
   CCAGACTTAG GCACAGCACA ATGCCACCA CCACCAAGT GCGGCACAAG 1500
   GCGGTGGCGG TAGGGTATGT GTCTGAAAAT GAGCGTGGAG ATTGGGCTCG 1550
35 CACGGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCAG 1600
   GCAGCTGAGT TGTGTATTTC TGATAAGAGT CAGAGGTAAC TCCCGTTGCG 1650
   GTGCTGTAAA CGGTGGAGGG CAGTGTAGTC TGAGCAGTAC TCGTTGCTGC 1700
   CGCGCGCGCC ACCAGACATA ATAGCTGACA GACTAACAGA CTGTTCTTTT 1750
   CCATGGTCTT TTTCTGCAGT CACCGTCGGA CCAIGTGTA ACTTGATATT 1800
40 TTACATGAT TCTTTTACCA ATTCTGCCCC GAATTACACT TAAACGACT 1850
   CAACAGCTTA ACGTTGGCTT GCCACGCATT ACTTGACTGT AAAACTCTCA 1900
   CTCTTACCGA ACTTGGCCGT AACCTGCCAA CCAAGCGAG AACAAACAT 1950
   AACATCAAA GAAATCGACG ATTGTTAGGT AATCGTCACC TCCACAAAGA 2000
   GCGACTCGCT GTATACCGTT GGCATGCTAG CTTTATCTGT TCGGGAATAC 2050
45 GATGCCCAT TGTACTTGTG ACTGGTCTGA TATTCGTGAG CAAAACGAC 2100
   TTATGGTATT GCGAGCTTCA GTCGCACTAC ACGGTCGTTT TGTACTCTT 2150
   TATGAGAAAG CGTTCGCCCT TTCAGAGCAA TGTTCAAAGA AAGCTCATGA 2200
   CCAATTTCTA GCCGACCTTG CGAGCATCTT ACCGAGTAAC ACCACACCG 2250
   TCATTGTGAG TGATGCTGGC TTFAAAGTGC CATGGTATAA ATCCGTTGAG 2300
50 AAGCTGGGTT GGTACTGGTT AAGTCGAGTA AGAGGAAAAG TACAATATGC 2350
   AGACTAGGA GCGGAAAAC TGAACCTAT CAGCAACTTA CATGATATGT 2400
   CATCTAGTCA CTCAAAGACT TTAGGCTATA AGAGGCTGAC TAAAGCAAT 2450
   CCAATCTCAT GCCAATTCT ATTGTATAAA TCTCGCTCTA AAGGCCGAAA 2500
   AAATCAGCGC TCGACACGGA CTCATTGTCA CCACCCGTCA CCTAAAATCT 2550
55 ACTCAGCGTC GGCAAGGAG CCAATGGGTC TAGCAACTAA CTTACCTGTT 2600
   GAAATTCGAA CACCCAAACA ACTTGTAAAT ATCTATTCTGA AGCGAATGCA 2650
   GATTGAAGAA ACCTCCGAG ACTTGAAAAG TCCTGCCTAC GGAAGAGGCC 2700
   TACGCCATAG CCGAACGAGC AGCTCAGAGC GTTTTGATAT CATGCTGCTA 2750
   ATCGCCCTGA TGCTTCAACT AACATGTTGG CTTGCGGGCG TTCATGCTCA 2800
60 GAAACAAGGT TGGGACAACG ACTTCCAGGC TAACACAGTC AGAAATCGAA 2850
   ACGTACTCTC AACAGTTCGC TTAGGCATGG AAGTTTTCGC GCATTCTGGC 2900

```

118

ATLJ002 133492.1

TACACAATAA CAAGGGAAGA CTTACTCGTG GCTGCAACCC TACTAGCTCA 2950
 AAATTTATTC ACACATGGTT ACGCTTTGGG GAAATTTATGA TAATGATCCA 3000
 GATCACTTCT GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTTGG 3050
 TTTTGTGTGG ATCTGTGTGT CTTCTAGTT GCCAGCCATC TGTGTGTTGC 3100
 CCCTCCCCCG TGCCTTCCTT GACCTGGGAA GGTGCCACTC CCACTGTCTCT 3150
 5 TTCTAATAAA AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCAAT 3200
 CTATTTCTGGG GGGTGGGGTG GGGCAGCACA GCAAGGGGGA GGATTTGGGA 3250
 GACAAATGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT 3350
 10 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT 3400
 TTGACCCGGT GACCAAAGGT GCCTTTTATC ATCACTTTAA AAATAAAAAA 3450
 CAATTACTCA GTGCTGTGTA TAAGCAGCAA TTAATTATGA TTGATGCCTA 3500
 CATCACAACA AAAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 TATTGGAACA TTAATCTGAT TATATTATG ATAATAATAA AAACCTTATC 3600
 15 CCTATCCAAG AAGTATGATC TATCATTTGT TGGAAATGAAC TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GGTAAAGTAA ACGCCATTGT CAGCAAAATTG 3700
 ATCCAAAGAGA ACCAACTTAA AGCTTTCCCTG ACGGAATGTT AATTCCTGTT 3750
 GACCTGAGC ACTGATGAAT CCCCTAATGA TTTTGGTAAA AATCATTAAAG 3800
 TTAAGGTGGA TACACATCTT GTCATATGAT CCCGTAATG TGAGTTAGCT 3850
 20 CACTCATTAG GCACCCGAGG CTTTACACTT TATGCTTCCG GCTCGTATGT 3900
 TGTGTGGAAT TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC 3950
 CATGATTACG CCAAGGCGCG AATTAACCTT CACTAAAGGG AACCAAGCT 4000
 GGAGCTCCAC GCGGTGGCG GCGCTCTAG AACTAGTGA TCCCCGGGG 4050
 AGGTCAGAAAT GGTTCCTTTA CTGTTGTGCA ATTCTATTAT TTCATACAG 4100
 25 AACAAATAGT TCTATAACTG AAATATATTT GCTATTGTAT ATTTGATTTG 4150
 TCCCTCGAAC CATGAACACT CCTCCAGCTG AATTTACAAA TTCCTCTGTC 4200
 ATCTGCCAGG CCATTAAGTT ATTCATGGAA GATCTTTGAG GAACACTGCA 4250
 AGTTCATATC ATAAACACAT TTGAAATGA GTATTGTTT GCATTGTATG 4300
 GAGCTATGTT TTGCTGTATC CTCAGAAAAA AAGTTTGTGA TAAAGCATTC 4350
 30 ACACCCATAA AAAGATAGAT TTAATATTTT CAGCTATAGG AAAGAAAGTG 4400
 CGTGTGCTCT TCACTCTAGT CTCAGTTGGC TCCTTCACAT GCATGCTTCT 4450
 TTATTTCTCC TATTTTGTCA AGAAAATAAT AGGTACAGTC TTGTTCTCAC 4500
 TTATGTCCTG CCTAGCATGG CTCAGATGCA CGTTGTAGAT ACARGAAGGA 4550
 TCAAAATGAA CAGACTTCTG GTCTGTACT ACAACCATAG TAATAAGCAC 4600
 35 ACTAACTAAT AATTTGCTAAT TATGTTTTC ATCTCTAAGG TTCCCATCAT 4650
 TTTCTGTTTT CTTAAAGATC CCATTATCTG GTTGTAACTG AAGCTCAATG 4700
 GAACATGAGC AATATTTCCC AGTCTTCTCT CCCATCCAAC AGTCTGTATG 4750
 GATTAGCAGA ACAGGCGAGAA AACACATTTG TACCCAGAAAT TAAAACTAA 4800
 TATTTGCTCT CCATTCAATC CAAAATGGAC CTATTGAAAC TAAATCTAA 4850
 40 CCCAATCCCA TTAATGATT TCTATGGCGT CAAAGGTCAA ACTTCTGAAG 4900
 GGAACCTGTG GGTGGGTAC AATTCAGGCT ATATATTCCC CAGGGCTCAG 4950
 CGSATCCATG GGCTCCATCG GCGCAGCAAG CATGGAATTT TGTTTTGATG 5000
 TATTCAGGA GCTCAAAGTC CACCATGCCA ATGAGAACAT CTTCTACTGC 5050
 CCCATTGCCA TCATGTCAGC TCTAGCCATG GTATACCTGG GTGCAAAAGA 5100
 45 CAGCACCAGG ACACAGATAA ATAAGGTTGT TCGCTTTGAT AAATCTCCAG 5150
 GATTGCGAGA CAGTATTGAA GCTCAGTGTG GCACATCTGT AAACGTTTCA 5200
 TCTTCACTTA GAGACATCTT CAACCAATC ACCAAACCAA ATGATGTTTA 5250
 TTCGTTGAGC CTTGCCAGTA GACTTTATGC TGAAGAGAGA TACCCATCC 5300
 TGCCAGAATA CTTGCAGTGT GTGAAGGAAC TGTATAGAGG AGGCTTGGAA 5350
 50 CCTATCAACT TTCAAACAGC TGCAAGTCAA GCCAGAGAGC TCATCAATTC 5400
 CTGGGTAGAA AGTCAGACAA ATGGAATTAT CAGAAATGTC CTTGAGCCAA 5450
 GCTCCGTGGA TTCTCAAACT GCAATGGTTC TGGTAAATGC CATTGCTTTC 5500
 AAAGGACTGT GCGAGAAAAC ATTTAAGGAT GAAGACACAC AAGCAATGCC 5550
 TTTGAGAGTG ACTGAGCAAG AAAGCACAAC TGTGCAGATG ATGTACCAGA 5600
 55 TTGTTTATTT TAGAGTGGCA TCAATGGCTT CTGAGAAAAT GAAGATCCTG 5650
 GAGCTTCCAT TTGCCAGTGG GACAAATGAG ATGTTGGTGC TGTTCCTGTA 5700
 TGAAGTCTCA GGCCTTGAGC AGCTTGAGAG TATAATCAAC TTTGAAAAAC 5750
 TGACTGAATG GACCACTTCT AATGTTATGG AAGAGAGGAA GATCAAGATG 5800
 TACTTACCTC GCATGAAGAT GGAGGAAAAA TACAACTCA CATCTGTCTT 5850
 60 AATGGCTATG GGCATTACTG ACGTGTTAG CTCTTCAGCC AATCTGTCTG 5900
 GCATCTCTCT AGCAGAGAGC CTGAAGATAT TCAAGCTGT CCATGACGCA 5950

CATGCAGAAA TCAATGAAGC AGGCAGAGAG GTGGTAGGGT CAGCAGAGGC 6000
 TGGAGTGGAT GCTGCAAGCG TCTCTGAAGA ATTTAGGGCT GACCATCCAT 6050
 TCCTCTTCTG TATCAAGCAC ATCGCAACCA ACGCCGTCTT CTTCTTTGGC 6100
 AGATGTGTTT CCCCTCCGCG GCCAGCAGAT GACGCACCAG CAGATGACGC 6150
 ACCAGCAGAT GACGCACCAG CAGATGACGC ACCAGCAGAT GACGCACCAG 6200
 CAGATGACGC AACCAACATGT ATCCTGAAAG GCTCTTGTGG CTGGATCGGC 6250
 CTGCTGGATG ACGATGACAA AAAATACAAA AAAGCACTGA AAAAAGTGGC 6300
 AAAAAGTGGC TAATGAGGGC GCCTGGATCC AGATCACTTC TGGCTAATAA 6350
 AAGATCAGAG CTCTAGAGAT CTGTGTGTGT GTTTTGTGTG GATCTGCTGT 6400
 GCCTTCTAGT TGCCAGCCAT CTGTTGTTTG CCCCTCCCCC GTGCCCTTCT 6450
 TGACCCCTGA AGGTGCCACT CCCACTGTCC TTTCTAATA AAATGAGGAA 6500
 ATTGATCGC ATTTGTCTGAG TAGGTGTCTAT TCTATTCTGG GGGGTGGGGT 6550
 GGGGCAGCAC AGCAAGGGGG AGGATTGGGA AGACAATAGC AGGCATGCTG 6600
 GGGATGCGGT GGGCTCTATG GGTACCTCTC TCTCTCTCTC TCTCTCTCTC 6650
 TCTCTCTCTC TCTCTCGTA CCTCTCTCGA GGGGGGGCCC GTACCCCAAT 6700
 TCGCCCTATA GTGAGTCGTA TTACGCGCGC TCACTGGCCG TCGTTTACAA 6750
 ACGTGTGAC TGGGAAAACC CTGGCGTTAC CCAACTTAAT CGCCTTGCAG 6800
 CACATCCCCC TTTCCGACG TGGCGTAATA GCGAAGAGGC CCGCACCAGT 6850
 CGCCCTTCCC AACAGTTGCG CAGCCTGAAT GGCGAATGGA AATTGTAAGC 6900
 GTTAATATTT TGTAAAAAT CGCGTTAAAT TTTTGTAAA TCAGCTCATT 6950
 TTTTAACCA TAGGCCGAAA TCGGCAAAAT CCCITATAAA TCAAAAGAAT 7000
 AGACCGAGAT AGGGTTGAGT GTTGTTCAG TTTGGAACAA GAGTCCACTA 7050
 TTAAGAAAG TGGACTCCAA CGTCAAAGGG CGAAAAACCG TCTATCAGGG 7100
 CGATCGCCCA CTACTCCGGG ATCATATGAC AAGTGTGTGA TCCACCTTAA 7150
 CTTAATGATT TTTACCAAAA TCATTAGGGG ATTCATCAGT GCTCAGGGTC 7200
 AACGAGAAAT AACATTCCGT CAGGAAAGCT TATGATGATG ATGTGCTTAA 7250
 AAACCTTACT AATGGCTGGT TATGCATATC GCAATACATG CGAAAAACCT 7300
 AAAAGAGCTT GCCGATAAAA AAGGCCAATT TATTGCTATT TACCGCGGCT 7350
 TTTTATTGAG CTGAAAGAT AAATAAAATA GATAGGTTTT ATTTGAAGCT 7400
 AAATCTTCTT TATCGTAAAA AATGCCCTCT TGGGTTATCA AGAGGGTTCAT 7450
 TATATTTCCG GGAATAACAT CATTGTGTGA CGAAATAACT AAGCACTTGT 7500
 CTCCTGTTTA CTCCCCTGAG CTGAGGGGT TAACATGAAG GTCATCGATA 7550
 GCAGGATAAT AATACAGTAA AACGCTAAAC CAATAATCCA AATCCAGCCA 7600
 TCCCAATTG GTAGTGAATG ATTATAAATA ACAGCAAAACA GTAATGGGCC 7650
 AATAACACCG GTTGCAATTG TAAGGCTCAC CAATAATCCC TGTAAAGCAC 7700
 CTTGCTGATG ACTCTTTGTT TGGATAGACA TCACTCCCTG TAATGCAAGT 7750
 AAAGCGATCC CACCACCAGC CAATAAAAT AAAACAGGGA AAACTAACCA 7800
 ACCCTCAGAT ATAAACGCTA AAAAGSCAAA TGCCTACTA TCTGCAATAA 7850
 ATCCGAGCAG TACTGCCGTT TTTTCGCCCC ATTTAGTGGC TATTCTTCCT 7900
 GCCACAAAGG CTTGGAATAC TGAGTGTAAA AGACCAAGAC CCGCTAATGA 7950
 AAAGCCAAAC ATCATGCTAT TCCATCCAAA ACGATTTTCG GTAAATAGCA 8000
 CCCACACCGT TGGCGGAATT TGGCCTATCA ATTGCGCTGA AAAATAAATA 8050
 ATCAACAAAA TGGCATCGTT TTAATAAAG TGATGTATAC CGAATTCAGC 8100
 TTTTGTTCCT TTTAGTGAGG GTTAATTGCG CGCTTGGCGT AATCATGGTC 8150
 ATAGCTGTTT CCTGTGTGAA ATTGTTATCC GCTCACAAAT CCACACAACA 8200
 TACGAGCCGG AAGCATAAAG TGTAAAGCCT GGGGTGCCTA ATGAGTGAGC 8250
 TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTTCC AGTCGGGAAA 8300
 CCTGTGCTGC CAGCTGCATT AATGAATCGG CCAACGCGCG GGGAGAGGCG 8350
 GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT CGCTCACTGA CTCGCTGCGC 8400
 TCGGTGCTTC GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAAT 8450
 ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA 8500
 AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT 8550
 TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC GACGCTCAAG 8600
 TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC 8650
 CTGGAAGCTC CCTCTGTCGC TCTCTGTTC CGACCCGTGC GCTTACCGGA 8700
 TACCTGTCCG CCTTCTCTCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC 8750
 ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTGCTGCC AAGCTGGGCT 8800
 GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCTTT ATCCGGTAAC 8850
 TATCGTCTTG AGTCCAAACC GGTAAAGACAC GACTTATCGC CACTGGCAGC 8900
 AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG 8950
 AGTTCTTGAA GTGTGTCCT AACTACGGCT ACACAGAAAG GACAGTATTT 9000

```

GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG 9050
CTCTTGATCC GGCAAAACAAA CCACCGCTGG TAGCGGTGGT TTTTGTGTTT 9100
GCAAGCAGCA GATTACGCGC AGAAAAAAG GATCTCAAGA AGATCCTTTG 9150
ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT CACGTTAAGG 9200
5 GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACTAG ATCCTTTTAA 9250
ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG 9300
TCTGACAGTT ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG 9350
TCTATTTCGT TCATCCATAG TTGCTGACT CCCCCTCGTG TAGATAACTA 9400
CGATACGGGA GCGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA 9450
10 GACCCACGCT CACCGGCTCC AGATTATCA GCAATAAACCC AGCCAGCCGG 9500
AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT 9550
CTATTAAATG TTGCCGGGAA GCTAGAGTAA GTAGTTCCGC AGTTAATAGT 9600
TTGCGCAACG TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC 9650
GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA AGGCGAGTTA 9700
15 CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCTCCG 9750
ATCGTTGTCA GAAGTAAGTT GGCCGAGTGT TTATCACTCA TGGTTATGGC 9800
AGCACTGCAT AATTCTCTTA CTGCTATGCC ATCCGTAAGA TGCTTTTCTG 9850
TGACTGTGTA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGCGCA 9900
CCGAGTTGCT CTTGCCCGGG GTCAATACGG GATAATACCG CGCCACATAG 9950
20 CAGAACITTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG GGGCGAAAAC 10000
TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT 10050
GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACGAGCG TTTCTGGGTG 10100
AGCAAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGCGGACAC 10150
GGAAATGTTG AATACTCATA CTCTTCCCTT TTCAATATTA TTGAAGCATT 10200
25 TATCAGGGTT ATGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA 10250
AAATAAACAA ATAGGGGTTC CGGCACATT TCCCGAAAAA GTGCCAC 10297

```

SEQ ID NO:30 (pTnMod(Oval/ENT tag/P146/PA) - QUAIL)

```

30 CTGACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG 50
CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTGCG 100
TTTCTTCCCT TCCCTTCTCG CCACGTTGCG CGGCATCAGA TTGGCTATTG 150
GCCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200
35 CATGTCCAAC ATTACCGCCA TGTGACATT GATTATAGAC TAGTTATTAA 250
TAGTAATCAA TTACGGGGTC ATTAGTTTAT AGCCCATATA TGGAGTTCCG 300
CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC 350
CCCGCCCAAT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA 400
GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAGTGCCCA 450
40 CTGTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC CCTATTGACG 500
TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTGA CATGACCTTA 550
TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC 600
CATGGTGATG CGGTTTTGGC AGTACATCAA TGGCGGTGGA TAGCGGTTTG 650
ACTCACGGGG ATTTCCAAGT CTCACCCCA TTGACGTCAA TGGGAGTTTG 700
45 TTTTGGCACC AAAATCAACG GGAATTCCA AAATGTCGTA ACAAATCCGC 750
CCCATTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA 800
GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC 850
TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCCGCGGCG 900
GGAACGGTGC ATTGGAACGC GGATTCCCG TGCCAAGAGT GACGTAAGTA 950
50 CCGCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA 1000
CTGTTTTTGG CTTGGGGCCT ATACACCCCG GCTTCCTTAT GCTATAGGTG 1050
ATGGTATAGC TTAGCCTATA GGTGTGGGTT ATTGACCAIT ATTGACCACT 1100
CCCTATTG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTTT 1150
GCCACAAC TACTCTATTG CTATATGCCA ATACTCTGTC CTTGAGAGAC 1200
55 TGACACGGAC TCTGTATTTT TACAGSATGG GGTCCCATTT ATTATTTACA 1250
AATTACATA TACAACAACG CCGTCCCCCG TGCCCGCAGT TTTTATTAAA 1300
CATAGCGTGG GATCTCCACG CGAATCTCGG GTACGTGTTT CGGACATGGG 1350
CTCTTCTCCG GTAGCGGCGG AGCTTCCACA TCCGAGCCCT GGTCCCATGC 1400
CTCCAGCGGC TCATGGTTCG TCGGCAGCTC CTTGCTCCTA ACAGTGGAGG 1450
60 CCAGACTTAG GCACAGCACA ATGCCACCA CCACCACTGT GCCGACACAG 1500
GCCGTGGCGG TAGGGTATGT GTCTGAAAAA GAGCGTGGAG ATTGGGCTCG 1550

```

121

ATLLB02 133492.1

	CACGGCTGAC	GCAGATGGAA	GACTTAAGGC	AGCGGCAGAA	GAAGATGCAG	1600
	GCAGCTGAGT	TGTTGTATT	TGATAAGAGT	CAGAGGTAAC	TCCCGTTGGG	1650
	GTGCTGTTAA	CGGTGGAGGG	CAGTGTAGTC	TGAGCAGTAC	TCGTTGCTGC	1700
	CGCGCGCGCC	ACCAGACATA	ATAGCTGACA	GACTAACAGA	CTGTTCTCTT	1750
5	CCATGGGTCT	TTTCTGCAGT	CACCGTCGGA	CCATGTGTGA	ACTTGATATT	1800
	TTACATGATT	CTCTTTACCA	ATTCTGCCCC	GAATTACACT	TAAACCACT	1850
	CAACAGCTTA	ACGTTGGCTT	GCCACGCATT	ACTTGACTGT	AAAACTCTCA	1900
	CTCTTACCGA	ACTTGGCCGT	AACCTGCCAA	CCAAAGCGAG	AACAAAACAT	1950
	AACATCAAAC	GAATCGACCG	ATTGTTAGGT	AATCGTCACC	TCCACAAAGA	2000
10	GCAGCTCGCT	GTATACCGTT	GGCATGCTAG	CTTTATCTGT	TCGGGAATAC	2050
	GATGCCCAT	GTACTTCTTG	ACTGCTCTGA	TATTCGTGAG	CAAAAACGAC	2100
	TTATGGTATT	GCGAGCTTCA	GTGCGACTAC	ACGGTCGTTC	TGTTACTCTT	2150
	TATGAGAAAG	CGTTCGCGCT	TTCAGAGCAA	TGTTCAAAGA	AAGCTCATGA	2200
	CCAATTTCTA	GCCGACCTTG	CGAGCATCT	ACCGAGTAAC	ACCACACCGC	2250
15	TCATTGTICAG	TGATGCTGGC	TTTAAAGTGC	CATGGTATAA	ATCCGTTGAG	2300
	AAGCTGGGTT	GGTACTCGTT	AAGTCGAGTA	AGAGGAAAAG	TACAATATGC	2350
	AGACCTAGGA	GCGGAAAAC	GGAAACCTAT	CAGCAACTTA	CATGATATGT	2400
	CATCTAGTCA	CTCAAAGACT	TTAGGCTATA	AGAGGCTGAC	TAAAGCAAT	2450
	CCAATCTCAT	GCCAAATTCT	ATTGTATAAA	TCTCGCTCTA	AAGGCCGAAA	2500
20	AAATCAGCGC	TCGACACGGA	CTCATTGTCA	CCACCCGTCA	CCTAAAAATCT	2550
	ACTCAGCGTC	GGCAAAGGAG	CCATGGGTC	TAGCAACTAA	CTTACCTGTT	2600
	GAAATTCGAA	CACCCAAACA	ACTTGTAAAT	ATCTATTCTGA	AGCGAATGCA	2650
	GATTGAAGAA	ACCTTCGAG	ACTTGAAAAG	TCTTGCCTAC	GGACTAGGCC	2700
	TACGCCATAG	CCGAACGAGC	AGCTCAGAGC	GTTTTGATAT	CATGCTGCTA	2750
25	ATCGCCCTGA	TGCTTCAAGT	AACATGTTGG	CTTGC GGCGG	TTTATGCTCA	2800
	GAACAAGGT	TGGGACAAGC	ACTTCCAGGC	TAACACAGTC	AGAAATCGAA	2850
	ACGTACTCTC	AACAGTTGCG	TTAGGCATGG	AAGTTTTGCG	GCATTCTGGC	2900
	TACACAATAA	CAAGGGAAGA	CTTACTCGTG	GCTGCAACCC	TACTAGTCA	2950
	AAATTTATTC	ACACATGGTT	ACGCTTTGGG	GAAATTTATGA	TAATGATCCA	3000
30	GATCACTTCT	GGCTAATAAA	AGATCAGAGC	TCTAGAGATC	TGTTGTTTGG	3050
	TTTTTTGGG	ATCTGCTGTG	CCTTCTAGTT	GCCAGCCATC	TGTTGTTTGC	3100
	CCCTCCCGG	TGCCTTCTCT	SACCTCGAAA	GGTGCCACTC	CCACTGTCTC	3150
	TTCTTAATAA	AATGAGGAAA	TTGCATCGCA	TGTCTGAGT	AGGTGTCTAT	3200
	CTATTCTGGG	GGGTGGGGTG	GGGCAGCACA	GCAAGGGGGA	GGATTGGGAA	3250
35	GACAATAGCA	GGCATGCTGG	GGATGCGGTG	GGCTCTATGG	GTACCTCTCT	3300
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCGGTAC	CTCTCTCTCT	3350
	CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CGGTACCAGG	TGCTGAAGAA	3400
	TTGACCCGGT	GACCAAGGT	GCCTTTTATC	ATCACTTTAA	AAATAAAAAA	3450
	CAATTACTCA	GTGCCTGTTA	TAAGCAGCAA	TTAATTATGA	TTGATGCCTA	3500
40	CATCACAACA	AAAACCTGAT	TAACAAATGG	TTGGTCTGCC	TTAGAAAGTA	3550
	TATTTGAACA	TTATCTTGAT	TATATTATTT	ATAATAATAA	AAACCTTATC	3600
	CCATATCCAAG	AAGTGATGCC	TATCATTTGT	TGGAATGAAC	TTGAAAAAAA	3650
	TTAGCCTTGA	ATACATTACT	GGTAAGGTAA	ACGCCATTGT	CAGCAAAATTG	3700
45	ATCCAAGAGA	ACCAACTTAA	AGCTTTCTGT	ACGGAATGTT	AATTCTCGTT	3750
	GACCCCTGAGC	ACTGATGAAT	CCCCTAATGA	TTTTGGTAAA	AATCATTAA	3800
	TTAAGGTGGA	TACACATCTT	GTCAATATGAT	CCCGGTAATG	TGAGTTAGCT	3850
	CACCTCATTAG	GCACCCACAG	CTTACACTT	TATGCTTCCG	GCTCGTATGT	3900
	TGTGTGGAAT	TGTGAGCGGA	TAACAATTTT	ACACAGGAAA	CAGCTATGAC	3950
50	CATGATTACG	CCAAGCGCGC	AATTAACCTT	CACTAAAGGG	AACAAAAGCT	4000
	GGAGCTCCAC	CGCGGTGGCG	GCCGCTCTAG	AACTAGTGGA	TCCCCCGGGG	4050
	AGGTCAGAA	GGTTTCTTTA	CTGTTGTGCA	ATTCTATTAT	TTCAATACAG	4100
	AACAAAAGCT	TCTATAACTG	AAATATATTT	GCTATTGTAT	ATTATGATTG	4150
	TCCCTCGAAC	CATGAACACT	CCTCCAGCTG	AATTTTCAAA	TTCTCTGTGC	4200
55	ATCTGCCAGG	CTGGAAGATC	ATGGAAGATC	TCTGAGGAAC	ATTGCAAGTT	4250
	CATACCATAA	ACTCAITTTG	AATTGAGTAT	TATTTTGCTT	TGAATGGAGC	4300
	TATGTTTTCG	AGTTCCCTCA	GAAGAAAAGC	TGTTTATAAA	GCGTCTACAC	4350
	CCATCAAAAG	ATATATTTAA	ATATTCCAAC	TACAGAAAAG	TTTTGTCTGC	4400
	TCTTCACTCT	GATCTCAGIT	GGTTTCTTCA	CGTACATGCT	TCTTTATTTG	4450
	CCTATTTTGT	CAAGAAAATA	ATAGGTCAAG	TCCTGTTCTC	ACTTATCTCC	4500
60	TGCCTAGCAT	GGCTTAGATG	CACGTTGTAC	ATTCAAGAAG	GATCAAAATGA	4550
	AACAGACTTC	TGGTCTGTTA	CAACAACCAT	AGTAATAAAC	AGACTAACTA	4600

ATAATTGCTA ATTATGTTTT CCATCTCTAA GGTTCACACA TTTTCTGT 4650
 TTAAGATCCC ATTATCTGGT TGTAACTGAA GCTCAATGGA ACATGAACAG 4700
 TATTTCTCAG TCTTTCTCC AGCAATCCTG ACGGATTAGA AGAAGTGGCA 4750
 GAAAACACTT TGTACCCAG AATTAACAAAC TAATATTTGC TCTCCCTTCA 4800
 5 ATCCAAATG GACCTATTGA AACTAAATC TGACCCAATC CCATTAAAT 4850
 ATTTCTATGG CGTCAAGGT CAAACTTTTG AAGGGAACCT GTGGGTGGGT 4900
 CCCATTCAG GCTATATAT CCCAGGGCT CAGCCAGTGG ATCCATGGGC 4950
 TCCATCGTG CAGCAAGCAT GGAATTTGT TTTGATGTAT TCAAGGAGCT 5000
 CAAAGTCCAC CATGCCAATG ACAACATGCT CTACTCCCC TTGCCATCT 5050
 10 TGTCAACTCT GGCCATGGTC TTCCTAGGTG CAAAAGACAG CACCAGGACC 5100
 CAGATAAATA AGGTTGTTCA CTTTGATAAA CTTCCAGGAT TCGGAGACAG 5150
 TATTGAAGCT CAGTGTGGCA CATCTGTAAA TGTTCACCTT TCACCTAGAG 5200
 ACATCTCAA CCAATCACC AAACAAAATG ATGCTTATTC GTTCAGCCTT 5250
 GCCAGTAGAC TTTATGCTCA AGAGACATAC ACAGTCGTGC CGGAATACIT 5300
 15 GCAATGTGTG AAGGAACTGT ATAGAGGAGG CTTAGAATCC GTCAACTTTC 5350
 AAACAGCTGC AGATCAAGCC AGAGGCCCTCA TCAATGCCCTG GTTAGAAAGT 5400
 CAGACAAACG GAATATCAG AAACATCCTT CAGCCAGGCT CCGTGGATTG 5450
 TCAAACTGCA ATGGTCTCG TTAATGCCAT TGCTTCAAG GCACTGTGGG 5500
 AGAAAGCATT TAAGGCTGAA GACACGCAAA CAATACCTTT CAGAGTGACT 5550
 20 GAGCAAGAAA GCAAACTGTG GCAGATGATG TACCAGATTG GTTCATTTAA 5600
 AGTGGCATCA ATGGCTTCTG AGAAATGAA GATCTGGAG CTTCATTG 5650
 CCAGTGGAAC AATGAGCATG TTGGTGTCTG TGCTGTATGA TGTCTCAGGC 5700
 CTGAGCAGC TTGAGGTAT AATCAGCTTT GAAAACTGA CTGAATGGAC 5750
 CAGTTCTAGT ATTATGGAAG AGAGGAAGGT CAAAGTGATC TTACCTCGCA 5800
 25 TGAAGATGGA GGAGAAATAC AACCTCACAT CTCTCTTAAT GGCTATGGGA 5850
 ATTAAGTACC TGTTCAGCTC TTCAGCCAAAT CTGTCTGGCA TCTCCTCAGT 5900
 AGGGAGCCTG AAGATATCTC AAGCTGTCCA TGACGACAT CGAGAAATCA 5950
 ATGAAGCGGG CAGAGATGTG GTAGGCTCAG CAGAGGCTGG AGTGGATGCT 6000
 ACTGAAGAAT TTAGGGCTGA CCATCCATTC CTCTTCTGTG TCAAGCACAT 6050
 30 CGAAACCAAC GCCATTCTCC TCTTTGGCAG ATGTGTTTCT CCGCGGCCAG 6100
 CAGATGACGC ACCAGCAGAT GACGCACCAG CAGATGACGC ACCAGCAGAT 6150
 GACGCACCAG CAGATGACGC ACCAGCAGAT GACGCACAA CATGTATCCT 6200
 GAAAGSCTCT TGTGGCTGGA TCGGCTCTGT GGATGACGAT GACAAAAAAT 6250
 ACAAAAAAGC ACTGAAAAAA CTGGCAAAAC TGCTGTAATG AGGGCGCCTG 6300
 35 GATCCAGATC ACTTCTGGCT AATAAAAGAT CAGAGCTCTA GAGATCTGTG 6350
 TGTGTGTTTT TTGTGGATCT GCTGTGCCTT CTAGTTGCCA GCCATCTGTT 6400
 GTTTGCCCTT CCCCGGTGCC TTCCTTGACC CTGGAAGGTG CCACTCCAC 6450
 TGTCTTTTCC TAATAAATG AGGAAATGCG ATCGCATGCT CTGAGTAGGT 6500
 40 GTCATTCTAT TCTGGGGGT GGGTGGGGC AGCAGACAA GGGGGAGGAT 6550
 TGGGAAGACA ATAGCAGGCA TGTGGGGAT GCGGTGGGCT CTATGGGTAC 6600
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCTCT 6650
 CTGAGGGGGG GGCCCGGTAC CCAATTCGCC CTATAGTGAG TCGTATTACG 6700
 CGCGCTCACT GGCGCTCGTT TTACAACGTC GTGACTGGGA AAACCCCTGGC 6750
 45 GTTACCCAAC TTAATCGCCT TGACGACAT CCCCCTCTCG CCAGCTGGCG 6800
 TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC 6850
 TGAATGGCGA ATGGAAATG TAAGCGTTAA TATTTGTTA AAATTGCGGT 6900
 TAAATTTTTG TAAATCAGC TCATTTTTTA ACCAATAGGC CGAAATCGGC 6950
 AAAATCCCTT ATAAATCAAA AGAATAGACC GAGATAGGGT TGAGTGTGTT 7000
 50 TCCAGTTTGG AACAAAGATC CACTATTAAA GAACGTGGAC TCCAACGTCA 7050
 AAGGGCGAAA AACCGCTAT CAGGCGATG GCCCACTACT CCGGATCAT 7100
 ATGACAGAT GTGTATCCAC CTTAACTTAA TGATTTTAC CAAATCATT 7150
 AGGGGATTCA TCAGTGTCTA GGGTCAACGA GAATTAACAT TCCGTACGGA 7200
 AAGCTTATGA TGATGATGTG CTTAAAACT TACTCAATGG CTGGTTATGC 7250
 55 ATATCGCAAT ACATGCGAAA AACCTAAAG AGCTTGCCGA TAAAAAGGC 7300
 CAATTTATTG CTATTIACCG CGGCTTTTGA TTGAGCTTGA AAGATAAATA 7350
 AAATAGATAG GTTTTATTG AAGCTAAATC TTCTTTATCG TAAAAATGC 7400
 CCTCTGGGT TATCAAGAGG GTCAATATAT TTCGGGAAT AACATCATTT 7450
 GGTGACGAAA TAACTAAGCA CTTGTCTCCT GTTACTCCC CTGAGCTTGA 7500
 60 GGGGTTAACA TGAAGGTCAAT CGATAGCAGG ATAATAATAC AGTAAAAAGC 7550
 TAAACCAATA ATCCAAATCC AGCATCCCA AATTGGTAGT GAATGATTAT 7600
 AAATAACAGC AAACAGTAAT GGGCCAATAA CACCGGTGCT ATTGTAAGG 7650

CTCACCAATA ATCCCTGTAA AGCACCTTGC TGATGACTCT TIGTTTGGAT 7700
 AGACATCACT CCCTGTAAATG CAGGTAAAGC GATCCCACCA CCAGCCAATA 7750
 AAATTAAAC AGGGAACACT AACCAACCTT CAGATATAAA CGCTAAAAAG 7800
 GCAAAATGCAC TACTATCTGC AATAAATCCG AGCAGTACTG CCGTTTTTTC 7850
 5 GCGCCATTTA GTGGCTATTC TTCTGCCAC AAAGGCTTGG AATACTGAGT 7900
 GTAAAGAGCC AAGACCCGCT AATGAAAAGC CAACCATCAT GCTATTCCAT 7950
 CCAAAACGAT TTTCGGTAAA TAGCACCCAC ACCGTTGCGG GAATTTGGCC 8000
 TATCAATTGC GCTGAAAAAT AAATAATCAA CAAAATGGCA TCGTTTAAAA 8050
 TAAAGTGATG TATACCGAAT TCAGCTTTTG TTCCCTTTAG TGAGGGTTAA 8100
 10 TTGCGCGCTT GCGTAATCA TGGTCATAGC TGTTTCCTGT GTGAAATTGT 8150
 TATCCGCTCA CAATTCCACA CAACATACGA GCCGGAAGCA TAAAGTGTAA 8200
 AGCCTGGGGT GCCTAATGAG TGAGCTAACT CACATTAAAT CCGTTGCGCT 8250
 CACTGCCCGC TTTCAGTGC GGAAACCTGT CGTGCCAGCT GCATTAAATGA 8300
 ATCGGCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTCCGCG 8350
 15 TTCTCGCTC ACTGACTCGC TGCCTCGGT CTTTCGGCTG CGGCGAGCGG 8400
 TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT 8450
 AACGCAAGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGC CCAGGAACCG 8500
 TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCTGACG 8550
 AGCATCACAA AAATCGACGC TCAAGTCAGA GTGGCGAAA CCCGACAGGA 8600
 20 CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TCGCTCTCC 8650
 TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCTTT CTCCCTTCGG 8700
 GAAGCGTGGC GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTGCGTG 8750
 TAGGTCTTTC GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTGAGCC 8800
 CGACCGCTGC GCCTTATCCG GTAACATCG TCTTGAGTCC AACCCGGTAA 8850
 25 GACACGACTT ATCGCCACTG GCAGCAGCCA CTGTAACAG GATTAGCAGA 8900
 GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA 8950
 CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG 9000
 TTACCTTCGG AAAAGAGTGT GGTAGCTCTT GATCCGGCAA ACAAAACCACC 9050
 GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA CGCGCAGAAA 9100
 30 AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGG TCTGACGCTC 9150
 AGTGAACGA AAACCTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA 9200
 AGGATCTTCA CCTAGATCCT TTTAAATTA AAATGAAGT TTAATCAAT 9250
 CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA 9300
 GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTTCAT CATAGTTGCC 9350
 35 TGACTCCCCG TCGTGTGAT AACTACGATA CGGAGGGCTT TACCATCTGG 9400
 CCCAGTGTCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT 9450
 TATCAGCAAT AAACCAAGCC GCGGAAGGG CCGAGCGCAG AAGTGGTCTC 9500
 GCAACTTTAT CCGCTTCCAT CCAGTCTATT AATTGTTGCC GGAAGCTAG 9550
 AGTAAGTAGT TCGCCAGTTA ATAGTTTGGC CAACGTTGTI GCCATTGCTA 9600
 40 CAGGCATCGT GGTGTACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC 9650
 GGTTCCTAAC GATCAAGGCG AGTTACATGA TCCCCATGT TGTGCAAAAA 9700
 AGCGGTTAGC TCCTTCGGTC CTCGATCGT TGTCAAGAT AAGTTGGCCG 9750
 CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC 9800
 ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC 9850
 45 ATTCGAGAA TAGTGTATGC GGCAGCCGAG TTGCTCTTGC CCGGCGTCAA 9900
 TACGGGATAA TACCGGCCCA CATAGCAGAA CTTTAAAGT GCTCATCATT 9950
 GGAAACGTT CTTGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG 10000
 ATCCAGTTCT ATGTAACCCA CTCGTGCACC CAACGTATCT TCAGCATCTT 10050
 TTACTTTTAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC 10100
 50 GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTGGAATAC TCACTACTCT 10150
 CTTTTTCAA TATTATTGAA GCATTATCA GGGTTATTGT CTCATGAGCG 10200
 GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCGCGC 10250
 ACATTTCCCG GAAAAGTGCC AC 10272

55
 SEQ ID NO:31 (pTnMod(Oval/BNT tag/Proins/PA) - Chicken)

CTGACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG 50
 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCGC 100
 60 TTTCTTCCCT TCCTTCTCG CACGTTCCG CGGCATCAGA TTGGCTATTG 150
 GCCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200

	CATGTCCAAC	ATTACCGCCA	TGTTGACATT	GATTATTGAC	TAGTATTAA	250
	TAGTAATCAA	TTACGGGGTC	ATTAGTTCAT	AGCCCATATA	TGGAGTTCCG	300
	CGTTACATAA	CTTACGGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC	350
	CCCGCCCAT	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA	400
5	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT	AACTGCCCA	450
	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTACGCCC	CCTATTGACG	500
	TCAATGACGG	TAAATGGCCC	GCCTGGCATT	ATGCCAGTA	CATGACCTTA	550
	TGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC	600
	CATGCTGATC	CGGTTTGGC	AGTACATCAA	TGGGCGTGA	TAGCGGTTTG	650
10	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA	TTGACGTCAA	TGGGAGTTTG	700
	TTTTGGCACC	AAAAACAACG	GGACTTTCCA	AAATGTCGTA	ACAACTCCCG	750
	CCCATTGACG	CAATGGGCG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA	800
	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG	CCATCCACGC	850
	TGTTTTGACC	TCCATAGAAG	ACACGGGAC	CGATCCAGCC	TCCGGGGCGG	900
15	GGAACGGTGC	ATTGGAACGC	GGATTCCCGG	TGCCAAGAGT	GACGTAAGTA	950
	CCGCCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA	1000
	CTGTTTTTGG	CTTGGGGCCT	ATACACCCCC	GCTTCCTTAT	GCTATAGGTG	1050
	ATGGTATAGC	TTAGCCTATA	GGTGTGGGTT	ATTGACCATT	ATTGACCACT	1100
	CCCCIATTGG	TGACGATACT	TTCATTAAT	AATCCATAAC	ATGGCTCTTT	1150
20	GCCACAATA	TCTCTATTGG	CTATATGCCA	ATACTCTGTC	CTTCAGAGAC	1200
	TGACACGGAC	TCTGTATTTT	TACAGGATGG	GSTCCCATTT	ATTATTIACA	1250
	AATTACATA	TACAACAACG	CCGTCCCGCG	TGCCCGCAGT	TTTTATTAAA	1300
	CATAGCGTGG	GATCTCCACG	CGAATCTCGG	GTACGTGTTT	CGACATGGG	1350
	CTCTTCTCCG	GTAGCGGCGG	AGCTTCCACA	TCCGAGCCCT	GGTCCCATGC	1400
25	CTCCAGCGGC	TCATGGTCGC	TCGGCAGCTC	CTTGCTCCTA	ACAGTGGAGG	1450
	CCAGACTTAG	GCACAGCACA	ATGCCACCCA	CCACCACTGT	GCCGCACAAG	1500
	GCCGTGGCGG	TAGGATATGT	GTCTGAAAT	GAGCGTGGAG	ATTGGGCTCG	1550
	CACGGCTGAC	GCAGATGGAA	GACTTAAGGC	AGCGGCAGAA	GAAGATGCAG	1600
	GCAGCTGAGT	TGTTGTATTC	TGATAAGAGT	CAGAGGTAAC	TCCCGTGTGC	1650
30	GTGCTGTTAA	CGGTGGAGGG	CAGGTGAGTC	TGAGCAGTAC	TCGTTGCTGC	1700
	CGCGCGGCGC	ACCAGACATA	ATAGCTGACA	GACTAACAGA	CTGTTCCCTT	1750
	CCATGGGCTC	TTTCTGCAGT	CACCGTCGGA	CCATGTGTGA	ACTTGATATT	1800
	TTACATGATT	CTCTTTACCA	ATTCTGCCCC	GAATTACACT	TAAACGACT	1850
	CAACAGCTTA	ACGTTGGCTT	GCCACGCATT	ACTTGACTGT	AAAACTCTCA	1900
35	CTCTTACCGA	ACTTGGCCGT	AACCTGCCAA	CCAAAGCGAG	AACAAAACAT	1950
	AACATCAAAC	GAATCGACCG	ATTGTTAGGT	AATCGTCACC	TCCACAAAGA	2000
	GCGACTCGCT	GTATACCGTT	GGCATGCTAG	CTTTATCTGT	TCGGGAATAC	2050
	GATGCCCAIT	GTACTGTTG	ACTGGTCTGA	TATTCGTGAG	CAAAAACGAC	2100
	TTATGGTATT	GCGAGCTTCA	GTGCACTAC	ACGGTCTGTC	TGTTACTCTT	2150
40	TATGAGAAAG	CGTTCGCCGT	TTCAGAGCAA	TGTTCAAAGA	AAGCTCATGA	2200
	CCAAATTTCTA	GCCGACCTTG	CGAGCATTCT	ACCGAGTAAC	ACCACACCGC	2250
	TCATGTGCTAG	TGATGCTGGC	TTTAAAGTGC	CATGCTATAA	ATCCGTTGAG	2300
	AAGCTGGGTT	GGTACTGGTT	AAGTCGAGTA	AGAGGAAAAG	TACATATATG	2350
	AGACCTAGGA	CCGGAACAACT	CGAAACCTAT	CAGCAACTTA	CATGATATGT	2400
45	CATCTAGTCA	CTCAAAGACT	TTAGGCTATA	AGAGGCTGAC	TAAAGCAAT	2450
	CCAATCTCAT	GCCAAATCT	ATTGTATAAA	TCTGCTCTTA	AAGGCCGAAA	2500
	AAATCAGCGC	TCGACACGGA	CTCATTGTCA	CCACCCGTCA	CCTAAAATCT	2550
	ACTCAGCGTC	GGCAAAGGAG	CCATGGGTTT	TAGCAACTAA	CTTACCTGTT	2600
	GAAATTCGAA	CACCCAAACA	ACTTGTAAAT	ATCTATTCGA	AGCGAATGCA	2650
50	GATTGAAGAA	ACCTTCGGAG	ACTTGAAAAG	TCCTGCTTAC	GGACTAGGCC	2700
	TACGCCATAG	CCGAACGAGC	AGCTCAGAGC	GTTTGTATAT	CATGCTGCTA	2750
	ATCGCCCTGA	TGCTTCAACT	AACATGTTGG	CTTGCGGGCG	TTTATGCTCA	2800
	GAAACAAGGT	TGGGACAAGC	ACTTCCAGGC	TAAACACAGTC	AGAAATCCAA	2850
	ACGTACTCTC	AACAGTTGCG	TTAGGCATGG	AAGTFTTGGC	GCATTCTGGC	2900
55	TACACAATAA	CAAGGGAGAA	CTTACTCGTG	GCTGCAACCC	TACTAGCTCA	2950
	AAATTTATTC	ACACATGGTT	ACGCTTTGGG	GAAATTATGA	TAATGATCCA	3000
	GATCACTTCT	GGCTAATAAA	AGATCAGAGC	TCTAGAGATC	TGTGTGTTGG	3050
	TTTTTTGTGG	ATCTGCTGTG	CCTTCTAGTT	GCCAGCCATC	TGTTGTTTGC	3100
	CCCTCCCCCG	TGCCTTCCTT	GACCTGGGAA	GGTGCCACTC	CCACTGTCTT	3150
60	TTCTTAATAA	AATGAGGAAA	TTGCATCGCA	TTGCTGAGT	AGGTGTCTAT	3200
	CTATTCTGGG	GGGTGGGGTG	GGGCAGCACA	GCAAGGGGGA	GGATTGGGAA	3250

GACARTAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCGGTAC CTCTCTCTCT 3350
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCAGG TGCTGAAGAA 3400
 TTGACCCGGT GACCAAAGGT GCCTTTTATC ATCACTTTAA AAATAAAAAA 3450
 CAATTACTCA GTGCTGTGTA TAAGCAGCAA TTAATTATGA TTGATGCCTA 3500
 5 CATCACAACA AAAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 TATTTGAACA TTATCTTGAT TATATTATG ATAATAATAA AAACCTTATC 3600
 CCTATCCAAG AAGTGAIGCC TATCATTGGT TGGAAATGAAC TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GGTAAAGTAA ACGCCATTGT CAGCAAATTG 3700
 10 ATCCAAGAGA ACCAACTTAA AGCTTTCTCTG ACGGAATGTT AATTCTCGTT 3750
 GACCCGTAGC ACTGATGAAT CCCCTAATGA TTTTGGTAAA AATCATTAA 3800
 TTAAGGTGGA TACACATCTT GTCATATGAT CCGGTAAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCAGG CTTTACACTT TATGCTTCCG GCTCGTATGT 3900
 TGTGTGGAAT TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC 3950
 15 CATGATTACG CCAAGCGCGC AATTAAACCCT CACTAAAGGG AACAAAAAGCT 4000
 GGAGCTCCAC CGCGGTGGCG GCCGCTCTAG AACTAGTGGG TCCCCGGGG 4050
 AGGTCAGAAT GGTTTCTTTA CTGTTTGTC AATTCTATTAT TTCAATACAG 4100
 AACAAATAGT TCTATAACTG AAATATATTT GCTATTGTAT ATTATGATTG 4150
 TCCCTCGAAC CATGAACACT CCTCCAGCTG AATTTCACAA TTCTCTGTCT 4200
 20 ATCTGCCAGG CCATTAAGTT ATTCATGGAA SATCTTTGAG GAACACTGCA 4250
 AGTTCAATC ATAAACACAT TTGAAATGA GTATTGTTTT GCATTGTATG 4300
 GAGCTATGTT TTGCTGTATC CTCAGAAAAA AAGTTTGTGA TAAAGCATTC 4350
 ACACCCATAA AAGATAGAT TTAATATTC CAGCTATAGG AAAGAAAGTG 4400
 CGTCTGCTCT TCACTCTAGT CTCAGTTGGC TCCTTCACAT GCATGCTTCT 4450
 25 TTATTTCTCC TATTTTGTGA AGAAAAATAT AGGTCACGTC TTGTTCTCAC 4500
 TTATGTCCTG CCTAGCATGG CTCAGATGCA CGTTGTAGAT ACAAGAAAGG 4550
 TCAAAATGAA CAGACTTCTG GTCTGTTACT ACAACCATAG TAATAAGCAC 4600
 ACTAACATAA AATTGCTAAT TATGTTTTCC ATCTCTAAGG TTCCACATTT 4650
 TTTCTGTTTT CTTAAAGATC CCATTATCTG GTTGTAACTG AAGCTCAATG 4700
 30 GAACATGAGC AATATTTCCC AGTCTTCTCT CCCATCCAAC AGTCTGTATG 4750
 GATTAGCAGA ACAGGCAGAA AACACATTGT TACCCAGAAT TAAAAACTAA 4800
 TATTTGCTCT CCATTCAATC CAAATGGAC CTATTGAAC TAAATCTAA 4850
 CCCAATCCCA TTAATGATT TCTATGGCGT CAAAGGTCAA ACTTCTGAAG 4900
 35 GGAACCTGTG GGTGGGTCAC AATTCAAGGT ATATATTTCC CAGGGCTCAG 4950
 CGGATCCATG GGCTCCATCG GCGCAGCAAG CATGGAATTT TGTTTTGATG 5000
 TATTCAAGGA GCTCAAAGTC CACCATGCCA ATGAGAACAT CTTCTACTGC 5050
 CCCATTGCCA TCATGTCCAG TCTAGCCATG GTATACCTGG GTGCAAAAGA 5100
 CAGCACCAGG ACACAGATAA ATAGGTTGT TCGCTTTGAT AAACCTTCCAG 5150
 40 GATTGCGAGA CAGTATTGAA GCTCAGTGTG GCACATCTGT AAACGTTTCC 5200
 TCTTCACTTA GAGACATCTT CAACCAAATC ACCAAACCAA ATGATGTTTA 5250
 TTGTTTCTAG CTTGCCAGTA GACTTTATGC TGAAGAGAGA TACCCAAATC 5300
 TGCCAGATA CTTGCACTGT GTGAAGGAAC TGTATAGAGG AGGCTTGGAA 5350
 CCTATCAACT TTCAAACAGC TGCAGATCAA GCCAGAGAGC TCATCAATTC 5400
 45 CTGGGTAGAA AGTCAGACAA ATGGAATTAT CAGAAATGTC CTTCAAGCAA 5450
 GCTCCGTGGA TTCTCAAATC GCAATGGTTC TGGTTAATGC CATTTCTTCT 5500
 AAAGGACTGT GGGAGAAAAC ATTTAAGGAT GAAGACACAC AAGCAATGCC 5550
 TTTCAAGATG ACTGAGCAAG AAAGCAAACC TGTGAGATG ATGTACCAGA 5600
 TTGTTTATTT TAGAGTGGCA TCAATGGCTT CTGAGAAAAT GAAGATCCTG 5650
 50 GAGCTTCCAT TTGCCAGTGG GACAATGAGC ATGTTGGTGC TGTGCTCTGA 5700
 TGAAGTCTCA GGCCTTGAGC AGCTTGAGAG TATAATCAAC TTTGAAAAAC 5750
 TGACTGAATG GACCAGTTCT AATGTTATGG AAGAGAGGAA GATCAAAGTG 5800
 TACTTACCTC GCATGAAGAT GGAGGAAAAA TACAACCTCA CATCTGTCTT 5850
 AATGGCTATG GGCATTACTG ACGTGTITAG CTCTTCAGCC AATCTGTCTG 5900
 GCATCTCCTC AGCAGAGAGC CTGAAGATAT CTCAGCTGT CCATGCAGCA 5950
 55 CATGCAGAAA TCAATGAAGC AGCAGAGAG GTGGTAGGGT CAGCAGAGGC 6000
 TGGAGTGGAT GCTGCAAGCG TCTCTGAAGA ATTTAGGGCT GACCATCCAT 6050
 TCCTCTTCTG TATCAAGCAC ATCGCAACCA ACGCCGTCTT CTTCTTTGGC 6100
 AGATGTGTTT CCCCTCCGGG GCCAGCAGAT GACGCACCAG CAGATGACGC 6150
 ACCAGCAGAT GACGCACCAG CAGATGACGC ACCAGCAGAT GACGCACCAG 6200
 60 CAGATGACGC AACCAATGT ATCCTGAAAG GCTCTTGTGG CTGGATCGGC 6250
 CTGCTGGATG ACGATGACAA ATTTGTGAAC CAACACCTGT GCGGCTCACA 6300

CCTGGTGGAA GCTCTCTACC TAGTGTGCGG GGAACGAGGC TTCTTCTACA 6350
 CACCCAGAAC CCGCCGGGAG GCAGAGGACC TGCAGGTGGG GCAGGTGGAG 6400
 CTGGCGGGGG GCCCTGGTGC AGGCAGCCTG CAGCCCTTGG CCTGGAGGG 6450
 GTCCCTGCAG AAGCGTGGCA TTGTGGAACA ATGCTGTACC AGCATCTGCT 6500
 CCCTCTACCA GCTGGAGAAC TACTGCAACT AGGGCGCCTG GATCCAGATC 6550
 5 ACTTCTGGCT AATAAAAGAT CAGAGCTCTA GAGATCTGTG TGTGTGTTTT 6600
 TTGTGGATCT GCTGTGCTT CTAGTTGCCA GCCATCTGTT GTTTGCCCTT 6650
 CCCCCGTGCC TTCCITGACC CTGGAAGGTG CCACTCCAC TGCTCTTTCC 6700
 TAATAAAATG AGGAAATTGC ATCGCATGTG CTGAGTAGGT GTCAATTCTAT 6750
 10 TCTGGGGGGT GGGGTGGGCG AGCAGCAGCA GGGGAGGAT TGGGAAGACA 6800
 ATAGCAGGCA TGCTGGGGAT GCGGTGGGCT CTATCGGTAC CTCTCTCTCT 6850
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCTCT CTCGAGGGGG 6900
 GGCCCGGTAC CCAATTCGCC CTATAGTGAG TCGTATTACG CGCGCTCACT 6950
 GGCGTCTGTT TTACAACGTC GTGACTGGGA AAACCTTGGC GTTACCCAAAC 7000
 15 TTAATCGCTT TGCAGCACAT CCCCTTTTCG CCAGCTGCGC TAATAGCGAA 7050
 GAGGCCGCCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC TGAATGGCGA 7100
 ATGGAAATTG TAAGCGTTAA TATTTTGTAA AAATTCGCGT TAAATTTTGG 7150
 TTAATCAGC TCATTTTTTA ACCAATAGGC CGAAATCGGC AAAATCCCTT 7200
 ATAAATCAAA AGAATAGACC GAGATAGGGT TGAGTGTGTG TCCAGTTTGG 7250
 20 AACAGAGTC CACTATTAAA GAACGTGGAC TCCAACGTCA AAGGCCGAAA 7300
 AACCGTCTAT CAGGGCGATG GCCCACTACT CCGGGATCAT ATGACAAGAT 7350
 GTGTATCCAC CTTAACTTAA TGATTTTAC CAAATCATT AGGGGATTCA 7400
 TCAGTGCTCA GGGTCAACGA GAATTAAAT TCCGTACAGG AAGCTTATGA 7450
 TGATGATGTG CTTAAAAACT TACTCAATGG CTGGTTATGC ATATCGCAAT 7500
 25 ACATCGGAAA AACCTAAAAG AGCTTGCCGA TAAAAAGGC CAATTTATTG 7550
 CTATTTACCG CGGCTTTTTA TTGAGCTTGA AAGATAAATA AAATAGATAG 7600
 GTTTTATTGG AAGCTAAATC TTCTTTATCG TAAAAAATGC CCTCTGGGT 7650
 TATCAAGAGG GTCAATTATAT TTCCGGGAAT AACATCATTT GGTGACGAAA 7700
 TAACATAAGCA CTTGTCTCCT GITTACTCCC CTGAGCTTGA GGGGTTAACA 7750
 30 TGAAGGTATC CGATAGCAGG ATAATAATAC AGTAAAACGC TAAACCAATA 7800
 ATCCAAATCC AGCCATCCCA AATTGGTAGT GAATGATTAT AAATAACAGC 7850
 AAACAGTAAT GGGCCAAATA CACCGGTTCG ATTGGTAAGG CTCACCAATA 7900
 ATCCCTGTAA AGCACCTTGC TGATGACTCT TTGTTTGAT AGACATCACT 7950
 35 CCCGTGAATG CAGGTAAAGC GATCCACCA CCAGCCAATA AAATTAALAC 8000
 AGGGAAAACT AACCAACCTT CAGATATAAA CGCTAAAAAG GCATATGCAC 8050
 TACTATCTGC AATAAATCCG AGCAGTACTG CCGTTTTTTC GCCCATTTA 8100
 GTGGCTATTG TTCTGCCAC AAAGGCTTGG AATACTGAGT GTAAAAAGAC 8150
 AAGACCCGCT AATGAAAAGC CAACCATCAT GCTATTCCAT CCAAAACGAT 8200
 40 TTTGGTAAA TAGCACCCAC ACCGTGCGG GAATTTGGCC TATCAATTGC 8250
 GCTGAAAAAT AAATAATCAA CAAATGGCA TCGTTTTAAA TAAAGTGATG 8300
 TATACCGAAT TCAGCTTTTG TTCCCTTTAG TGAGGGTTAA TTGCGCGCTT 8350
 GGCCTAATCA TGGTCATAGC TGTTCCTGT GTGAAATTGT TATCCGCTCA 8400
 CAATTCCACA CAACATACGA GCCGGAAGCA TAAAGTGTA AGCCTGGGGT 8450
 45 GCCTAATGAG TGAGCTAAT CACATTAATT GCGTTGCGCT CACTGCCCGC 8500
 TTTCCAGTCG GAAACCTGT CGTGCCAGCT GCATTAATGA ATCGGCCAAC 8550
 GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTCCGC TTCTCGCTC 8600
 ACTGACTCGC TCGGCTCGGT CGTTCGGCTG CCGCGAGCGG TATCAGCTCA 8650
 CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA 8700
 AGAACATGTG AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC 8750
 50 GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCTGACG AGCATCACAA 8800
 AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCGACAGGA CTATAAAGAT 8850
 ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TCGCTCTCC TGTTCGAC 8900
 CTGCCGCTTA CCGGATACCT GTCCGCTTTT CTCCCTTCGG GAAGCGTGGC 8950
 GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTGCGTG TAGGTCGTTT 9000
 55 GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTACGCC CGACCGCTGC 9050
 GCCTTATCCG GTAACATCG TCTTGAGTCC AACCCGGTAA GACACGACTT 9100
 ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG 9150
 TAGGCGGTGC TACAGATTC TTGAAGTGGT GGCCTAATA CGGCTACACT 9200
 AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG 9250
 60 AAAAGAGTT GGTAGCTCTT GATCCGCAA ACAAAACACC GCTGGTAGCG 9300
 GTGGTTTTT TGTTCGAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT 9350

CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAAACGA 9400
 AAACCTCACGT TAAGGGATT TGGTCATGAG ATTATCAAAA AGGATCTTCA 9450
 CCTAGATCCT TTAAATTA AAATGAAGTT TTAATCAAT CTAAAGTATA 9500
 TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC 9550
 5 TATCTCAGCG AICGTCTAT TTCTTCATC CATAGTTGCC TGACTCCCG 9600
 TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG CCCAGTGCT 9650
 GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT TATCAGCAAT 9700
 AAACCAAGCA GCGGAAGGG CCGAGCGCAG AAGTGGTCTT GCAACTTTAT 9750
 CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAGCTAG AGTAAGTAGT 9800
 10 TCGCCAGTTA ATAGTTTGGC CAACGTTGTT GCCATTGCTA CAGGCATCGT 9850
 GGTGTACGCG TCGTCGTTTG GTATGGCTTC ATTCACTCC CGTCCCAAC 9900
 GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGCAAAA AGCGGTTAGC 9950
 TCCCTCGGTC CTCGATCGT TGTGAGAAAT AAGTTGGCCG CAGTGTATC 10000
 ACTCATGGT ATGGCAGCAC TGCATAATC TCTACTGTC ATGCCATCG 10050
 15 TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCAGAGAA 10100
 TAGTGTATGC GCGGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA 10150
 TACCCGCGCA CATAGCAGAA CTTTAAAGT GCTCATCATT GGAAAACGTT 10200
 CTTGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTG 10250
 ATGTAACCA CTCGTGCACC CAACTGATCT TCAGCATCTT TACTTTTAC 10300
 20 CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAGG 10350
 GAATAGGGC GACACGAAA TGTGAATAC TCATACCTT CTTTTTCAA 10400
 TATTATTGAA GCATTTATCA GGGTTAATGT CTCATGAGCG GATACATATT 10450
 TGAATGTATT TAGAAAAATA AACAAATAGG GGTCCGCGC ACATTTCCCC 10500
 25 GAAAGTGCC AC 10512

SEQ ID NO:32 (pTnMod(Oval/ENT tag/Proins/PA) - QUAIL)

CTGACGCGCC CTGTAGCGGC GCATTAAGCG CGCGGGGTGT GGTGGTTACG 50
 CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGGCCG CTCCTTTGCG 100
 30 TTTCTTCCCT TCCTTTCTCG CCACGTTGCG CGGCATCAGA TTGGCTATTG 150
 GCCATTGCAT ACGTTGTATC CATATCATAA TATGTACATT TATATTGGCT 200
 CATGTCCAAC ATTACCGCCA TGTGACATT GATTATTGAC TAGTATTAA 250
 TAGTAATCAA TTACGGGGTC ATTAGTTTAT AGCCCATATA TGGAGTTCCG 300
 35 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC 350
 CCGGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA 400
 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT AACGCCCCA 450
 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCC CCTATTGACG 500
 40 TCAATGACCG TAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA 550
 TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC 600
 CATGGTGATG CGGTTTGGC AGTACATCAA TGGCCGTGGA TAGCGGTTTG 650
 ACTCACGGGG ATTTCCAAGT CTCACCCCA TTGACGTCAA TGGAGTTTG 700
 TTTTGGCACC AAAATCAAG GGACTTTCCA AAATGTCGTA ACAACTCCGC 750
 45 CCCATTGACG CAAATGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA 800
 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC 850
 TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCG 900
 GGAACGCTGC ATTGGAACGC GGAATCCCGG TGCCAAGAGT GACGTAAGTA 950
 CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA 1000
 50 CTGTTTTTGG CTTGGGGCCT ATACACCCCT GCTTCTTAT GCTATAGGTG 1050
 ATGGTATAGC TTAGCCTATA GGTGTGGGT ATTGACCATT ATTGACCACT 1100
 CCCCTATTGG TGACGATACT TTCCATTACT AATCCATAAC ATGGCTCTTT 1150
 GCCACAATA TCTCTATTGG CTATATGCCA ATACTCTGTC CTTGAGAGAC 1200
 TGACACGGAC TCTGTATTTT TACAGGATGG GGTCCCATTT ATTATTACA 1250
 AATTACATA TACAACAACG CCGTCCCCCG TGCCCGCAGT TTTTATTAAA 1300
 55 CATAGCGTGG GATCTCAAG CGAATCTCGG GTACGTGTTT CCGACATGGG 1350
 CTCCTCTCCG GTAGCGGCG AGCTTCACA TCCGAGCCCT GGTCCCATGC 1400
 TCCAGCGCG TCAATGGTGC TCGGCGCTC CTTGCTCCTA ACAGTGGAGG 1450
 CCAGACTTAG GCACAGCACA ATGCCACCA CCACCACTGT GCCGCACAAG 1500
 60 GCCGTGGCG TAGGGTATGT GTCTGAAAT GAGCGTGGAG ATTGGGCTCG 1550
 CACGCTGAC GCAGATGGA GACTTAAGGC AGCGGCAGAA GAAGATGCAG 1600
 GCAGCTGAGT TGTGTATTTC TGATAAGAGT CAGAGGTAAC TCCCGTTGCG 1650

GTGCTGTTAA CGGTGGAGGG CAGTGTAGTC TGAGCAGTAC TCGITGCTGC 1700
 CGCGCGCGCC ACCAGACATA ATAGCTGACA GACTAACAGA CTGTCCTTT 1750
 CCATGGGTCT TTTCTGCAGT CACCGTCGGA CCATGTGTGA ACTTGATATT 1800
 TTACATGATT CTCTTTACCA ATCTGCGCCC GAATTACACT TAAAAAGACT 1850
 CAACAGCTTA ACGTTGGCTT GCCACGCATT ACTTGACTGT AAAACTCTCA 1900
 CTCTTACCGA ACTTGGCCGT AACCTGCCAA CCAAAGCGAG AACAAAACAT 1950
 AACATCAAAAC GAATCGACCG ATGTGTAGGT AATCGTCACC TCCACAAAGA 2000
 GCGACTCGCT GTATACCGTT GGCATGCTAG CTTTATCTGT TCGGGAATAC 2050
 GATGCCCATT GTACTTGTG ACTGGTCTGA TATTGCTGAG CAAAAACGAC 2100
 TTATGGTATT GCGAGCTTCA GTCGCACTAC ACGGTCTTTC TGTTACTCTT 2150
 TATGAGAAAG CGTTCCTGCT TTCAGAGCAA TGTTCAAAGA AAGCTCATGA 2200
 CCAATTICTA GCCGACCTTG CGAGCATTTCT ACCGAGTAAC ACCACACCGC 2250
 TCATTGTGAG TGATGCTGGC TTTAAAGTGC CATGGTATAA ATCCGTTGAG 2300
 AAGCTGGGTT GGTACTGGTT AAGTCGAGTA AGAGGAAAAG TACATATATG 2350
 AGACCTAGGA GCGGAAAAC GGAACCTTAT CAGCAACTTA CATGATATGT 2400
 CATCTAGTCA CTCAAAGACT TTAGGCTATA AGAGGCTGAC TAAAAAGCAAT 2450
 CCAATCTCAT GCCAAATTTCT ATGTATATAA TCTCGCTCTA AAGCCCGAAA 2500
 AAATCAGCGC TCGACACGGA CTCATTGTCA CCACCCCTCA CCTAAATCT 2550
 ACTCAGCGTC GGCAAAAGAG CCAATGGTTC TAGCAACTAA CTTACCTGTT 2600
 GAAATTCGAA CACCCAAACA ACTTGTAAAT ATCTATTCGA AGCGAATGCA 2650
 GATTGAAGAA ACCTTCCGAG ACTTGAAAAG TCCTGCCTAC GGACTAGGCC 2700
 TACGCCATAG CCGAACGAGC AGCTCAGAGC GTTTTGATAT CATGCTGCTA 2750
 ATCGCCCTGA TGCTTCAACT AACATGTTGG CTTGCGGGCG TTCATGCTCA 2800
 GAAACAAGGT TGGGACAAGC ACTTCCAGGC TAACACAGTC AGAAATCGAA 2850
 ACGTACTCTC AACAGTTCCG TTAGGCATGG AAGTTTTCGG GCATCTGGC 2900
 TACACAATAA CAAGGGAAGA CTTACTCGTG GCTGCAACCC TACTAGTCA 2950
 AAATTTATTC ACACATGGTT ACGCTTTGGG GAAATATGA TAATGATCCA 3000
 GATCACTTCT GGCTAATAAA AGATCAGAGC TCTAGAGATC TGTGTGTTGG 3050
 TTTTGTGTTG ATCTGCTGTG CCTTCTAGTT GCCAGCCATC TGTGTGTTGC 3100
 CCCTCCCGCG TGCTTCTCTT GACCTCGGAA GGTGCCATC CCACCTGCTC 3150
 TTCTTAATAA AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCATT 3200
 CTATTTCTGGG GGGTGGGGTG GGGCAGCACA GCAAGGGGGA GGATGGGAA 3250
 GACAATAGCA GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCTCTCT 3300
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCGGTAC CTCTCTCTCT 3350
 CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CGGTACCAGG TGCTGAAGAA 3400
 TTGACCCGGT GACCAAGGTT GCCTTTTATC ATCACTTTAA AAATAAAAAA 3450
 CAATTAATCA GTGCTGTGTA TAAGCAGCAA TTAATTATGA TTGATGCCTA 3500
 CATCACAAAC AAAACTGATT TAACAAATGG TTGGTCTGCC TTAGAAAGTA 3550
 TATTTGAACA TTATCTTGAT TATATTATGG ATAATAATAA AAACCTTATC 3600
 CCTATCCAAG AAGTGATGCC TATCATTTGG TGGAAATGAA TTGAAAAAAA 3650
 TTAGCCTTGA ATACATTACT GGTAAAGTAA ACGCCATTCT CAGCAAAATTG 3700
 ATCCAAGAGA ACCAACTTAA AGCTTTCCTG ACGGAATGTT AATCTCTGTT 3750
 GACCTGAGC ACTGATGAAT CCCTAATGTA TTTTGGTAAA AATCAITTAAG 3800
 TTAAGGTGGA TACACATCTT GTCATATGAT CCGGTAAATG TGAGTTAGCT 3850
 CACTCATTAG GCACCCGAGC CTTTACACTT TATGCTTCCG GCTCGTATGT 3900
 TGTGTGGAAT TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC 3950
 CATGATTACG CCAAGCGCGC AATTAAACCT CACTAAAGGG AACAAAAGCT 4000
 GGAGCTCCAC CGCGGTGGCG GCGCTCTAG AACTAGTGGG TCCCCCGGGG 4050
 AGGTGAGAAT GGTTCCTTTA CTGTTTGTC AATTCTATTAT TTCAATACAG 4100
 AACAAAAGCT TCTATAACTG AAATATATTT GCTATTGTAT ATTATGATTG 4150
 TCCCTCGAAC CATGAACACT CCTCCAGCTG AATTTCACAA TTCTCTGTCT 4200
 ATCTGCCAGG CTGGAAGATC ATGGAAGATC TCTGAGGAAC ATTGCAAGTT 4250
 CATACCAATA ACTCATTTGG AATTGAGTAT TATTTTGCTT TGAATGGAGC 4300
 TATGTTTGGC AGTTCCCTCA GAAGAAAAGC TTGTTATAAA GCGCTACAC 4350
 CCATCMAAAG ATATATTTAA ATATTTCAAC TACAGAAAAG TTTTGTCTGC 4400
 TCTTCACTCT GATCTCAGTT GGTTCCTTCA CGTACATGCT TCTTTATTGG 4450
 CCTATTTTGT CAAGAAAATA ATAGGTCAAG TCCTGTCTCT ACTTATCTCC 4500
 TGCCTAGCAT GGCTTAGATG CACGTGTAC ATTCAAGAAG GATCAATGA 4550
 AACAGACTTC TGGTCTGTTA CAACAACCAT AGTAATAAAC AGACTAACTA 4600
 ATAATGCTA ATATATTTT CCATCTCTAA GGTCCACCA TTTTCTGTGT 4650
 TTAAGATCCC ATTATCTGGT TGTAACGTAA GCTCAATGGA ACATGAACAG 4700

TATTTCTCAG TCTTTTCTCC AGCAATCCTG ACGGATTAGA AGAACTGGCA 4750
 GAAACACTTT TGTTACCCAG AATTAAAAAC TAATATTTCG TCTCCCTTCA 4800
 ATCCAAATG GACCTATTGA AACTAAAAATC TGACCCCAATC CCATTAAATT 4850
 ATTTCTATGG CGTCAAAGGT CAAACTTTTG AAGGGAACCT GTGGGTGGGT 4900
 5 CCCAATTTCAG GCTATATATT CCCAGGGCT CAGCCAGTGG ATCCATGGGC 4950
 TCCATCGGTG CAGCAAGCAT GGAATTTTGT TTTGATGTAT TCAAGGAGCT 5000
 CAAAGTCCAC CATGCCAATG ACAACATGCT CTACTCCCCC TTTGCCATCT 5050
 TGTCAACTCT GGCCTATGGT TCCCTAGGTG CAAAAGACAG CACCAGGACC 5100
 CAGATAAATA AGGTTGTTCA CTTTGATAAA CTTCAGGAT TCGGAGACAG 5150
 10 TATTGAAGCT CAGTGTGGCA CATCTGTAAA TGTTCACTCT TCACCTAGAG 5200
 ACATACTCAA CCAATCACC AAACAAAATG ATGCTTATTC GTTCAGCCTT 5250
 GCCAGTAGAC TTTATGCTCA AGAGACATAC ACAGTCGTGC CGGAATACTT 5300
 GCAATGTGTG AAGGAACTGT ATAGAGGAGG CTTAGAATCC GTCAACTTTC 5350
 AAACAGCTGC AGATCAAGCC AGAGGCCTCA TCAATGCTG GGTAGAAAGT 5400
 15 CAGACAAAGC GAATTATCAG AAACATCCTT CAGCCAAGCT CCGTGGATTG 5450
 TCAAACTGCA ATGCTCCTGG TTAATGCCAT TGCCTTCAAG GGACTGTGGG 5500
 AGAAAGCATT TAAGGCTGAA GACACGCAAA CAATACCTTT CAGAGTGACT 5550
 GAGCAAGAAA GCACACCTGT GCAGATGATG TACCAGATTG GTTCATTAA 5600
 AGTGGCATCA ATGGCTTCTG AGAAAAAGAA GATCCTGGAG CTTCCATTG 5650
 20 CCAGTGAAGC AATGAGCATG TTGGTGTCTG TGCCTGATGA TGTCTCAGGC 5700
 CTTGAGCAGC TTGAGAGTAT AATCAGCTTT GAAAACTGA CTGAATGGAC 5750
 CAGTTCTAGT ATTATGGAAG AGAGGAAGGT CAAAGTGATC TTACCTCGCA 5800
 TGAAGATGGA GGAGAAATAC AACCTCACAT CTCTCTTAAT GGCTATGGGA 5850
 ATTACTGACC TGTTCAAGCTC TTCAGCCAAT CTGTCTGGCA TCTCCTCAGT 5900
 25 AGGGAGCCTG AAGATATCTC AAGCTGTCCA TGCAGCACAT GCAGAAATCA 5950
 ATGAAGCGGG CAGAGATGTG GTAGGCTCAG CAGAGGCTGG AGTGGATGCT 6000
 ACTGAAGAAAT TTAGGGCTGA CCATCCATTC CTCTTCTGTG TCAAGCACAT 6050
 CGAAACCAAC GCCATTCTCC TCTTTGGCAG ATGTTTCTCT CCGGCGCCAG 6100
 CAGATGACGC ACCAGCAGAT GACGCACCAG CAGATGACGC ACCAGCAGAT 6150
 30 GACGCACAGC CAGATGACGC ACCAGCAGAT GACGCACAA CATGTATCTT 6200
 GAAAGGCTCT TGTGGCTGGA TCGGCTGTCT GGATGACGAT GACAAATTTG 6250
 TGAACCAACA CCTGTGGGCG TCACACCTGG TGGAAAGCTCT CTACCTAGTG 6300
 TGGGGGGAAC GAGGCTTCTT CTACACACCC AAGACCCGCC GGGAGGCAGA 6350
 GGACCTGCAG GTGGGGCAGG TGGAGCTGGG CGGGGGCCCT GGTGCAGGCA 6400
 35 GCCTGCAGCC CTTGGCCCTG GAGGGGTCCC TGCAGAAGCG TGGCATTTGT 6450
 GAACAATGCT GTACAGCAT CTGCTCCCTC TACCAGCTGG AGAACTACTG 6500
 CAACTAGGGC GCCTGGATCC AGATCACTTC TGGCTAATAA AAGATCAGAG 6550
 CTCTAGAGAT CTGTGTCTTG GTTTTTTGTG GATCTGCTGT GCCCTCTAGT 6600
 40 TGCCAGCCAT CTGTTGTTTG CCCCTCCCCC GTGCCTTCCT TGACCTGGA 6650
 AGGTGCCACT CCCACTGTCC TTTCCTAATA AAATGAGGAA ATTGCATCGC 6700
 ATTGTCTGAG TAGGTGTCAT TCTATTCTGG GGGGTGGGT GGGGCGACAC 6750
 ACCAAGGGGG AGGATTGGGA AGACAATAGC AGGCATGCTG GGGATGCGGT 6800
 GGGCTCTATG GGTACCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC 6850
 45 TCTCTCGTA CCTCTCTGA GGGGGGGCCC GGTACCCAAT TCGCCCTATA 6900
 GTGAGTCGTA TTACGCGCGC TCACTGGCCG TCGTTTTACA ACGTCGTGAC 6950
 TGGGAAAAAC CTGSCGTAC CCAACTTAAT CGCCTTGAG CACATCCCCC 7000
 TTTGGCCAGC TGGCGTAATA GCGAAGAGGC CCGCACCGAT CGCCCTTCCC 7050
 AACAGTTGCG CAGCCTGAAT GCGGAATGGA AATTGTAAGC GTTAATATT 7100
 50 TGTTAAAAAT CGCGTTAAAT TTTTGTAAAA TCAGCTCATI TTTTAACCAA 7150
 TAGGCCGAAA TCGGCAAAAT CCTTATAAAA TCAAAAGAAAT AGACCGAGAT 7200
 AGGGTTGAGT GTTGTTCAG TTTGGAACAA GAGTCCACTA TTAAGAAGAC 7250
 TGGACTCCAA CGTCAAAGGG CGAAAAACCG TCTATCAGGG CGATGGCCCA 7300
 CTACTCCGGG ATCATATGAC AAGATGTGTA TCCACCTTAA CTTAATGATT 7350
 55 TTTACCAAAA TCATTAGGGG ATTATCAGT GCTCAGGGTC AACGAGAATT 7400
 AACATTCGGT CAGGAAAGCT TATGATGATG ATGTGCTTAA AACTTACTC 7450
 AATGGCTGGT TATGCATATC GCAATACATG CGAAAAACCT AAAAGAGCTT 7500
 GCCGATAAAA AAGGCCAATT TATTGCTATT TACCGCGGT TTTTATTGAG 7550
 CTTGAAAGAT AAATAAATA GATAGGTTT ATTTGAAGCT AAATCTTCTT 7600
 60 TATCGTAAAA AATGCCCTCT TGGTTATCA AGAGGGTCAT TATATTTCGC 7650
 GGAATAACAT CATTTGGTGA CGAAATAACT AAGCACTCT CTCTGTTTA 7700
 CTCCTCTGAG CTTGAGGGGT TAACATGAAG GTCATCGATA GCAGGATAAT 7750

ATL11802.133492.1

AATACAGTAA AACGCTAAAC CAATAATCCA AATCCAGCCA TCCCAAAATTG 7800
 GTAGTGAATG ATTATAAATA ACAGCAAAACA GTAATGGGCC AATAACACCG 7850
 GTTGCAATTGG TAAGGCTCAC CAATAATCCC TGTAAGCAC CTTGCTGATG 7900
 ACTCTTTGTT TGGATAGACA TCACTCCCTG TAATGCAGGT AAAGCGATCC 7950
 5 CACCACCAGC CAATAAATT AAACACGGGA AAATAACCA ACCTTCAGAT 8000
 ATAAPCGCTA AAAAGGCAAA TGCACACTA TCTGCAATAA ATCCGAGCAG 8050
 TACTGCCGTT TTTTCGCCCT ATTTAGTGGC TATTCTTCCT GCCACAAAGG 8100
 CTTGGAATAC TGAGTGTAAG AGACCAAGAC CCGCTAATGA AAAGCCAACC 8150
 ATCATGCTAT TCCATCCAAA ACGATTTTCG GTAAATAGCA CCCACACCGT 8200
 10 TGCGGGAATT TGGCCTATCA ATTGCGCTGA AAAATAAATA ATCAACAAAA 8250
 TGGCATCGTT TTAATAAAG TGATGTATAC CGAATTCAGC TTTTGTTCCT 8300
 TTTAGTGAGG GTTAATTGCG CGCTTGGCGT AATCATGGTC ATAGCTGTTT 8350
 CCTGTGTGAA ATTGTTATCC GCTCACAATT CCACACAACA TACGAGCCGG 8400
 AAGCATAAAG TGTAAAGCCT GGGGTGCCA ATGAGTGAGC TAACACACAT 8450
 15 TAATTGCGTT GCGCTCACTG CCCGCTTTCC AGTCGGGAAA CCTGTCTGTC 8500
 CAGCTGCATT AATGAATCGG CCAACGCGCG GGGAGAGGCG GTTTCGTAT 8550
 TGGGCGCTCT TCCGCTTCTT CGCTCACTGA CTCGCTGCGC TCGCTCGTTC 8600
 GGCTGCGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT ACGGTATCC 8650
 ACAGAATCAG GGGATAACGC AGGAAGAAC ATGTGAGCAA AAGGCCAGCA 8700
 20 AAAGGCCAGG AACCGTAAAA AGGCGCGCTT GCTGGCGTTT TTCCATAGGC 8750
 TCCGCCCCCT TGACGAGCAT CACAAAAATC GACGCTCAAG TCAGAGGTGG 8800
 CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC CTGGAAGCTC 8850
 CCTGTGCGC TCTCTGTTC CGACCTGCC GCTTACCGGA TACCTGTCCG 8900
 CCTTTCTCCC TTGCGGAAGC GTGCGGCTTT CTCATAGCTC ACGCTGTAGG 8950
 25 TATCTCAGTT CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA 9000
 ACCCCCCGTT CAGCCCGACC GCTGCGCTTT ATCCGTAAC TATCGCTTGG 9050
 AGTCCAACCC GGTAAAGCAC GACTTATCGC CACTGGCAGC AGCCACTGGT 9100
 AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG AGTTCTTGAA 9150
 GTGGTGGCCT AACTACGGCT AACTAGAAG GACAGTATTT GGTATCTGCG 9200
 30 CTCTGCTGAA GCCAGTTACC TTGCGAAAAA GAGTTGTAGT CTCTGATCC 9250
 GGCAACAAA CCACCGCTGG TAGCGGTGGT TTTTGTGTTT GCAAGCAGCA 9300
 GATTACCGCG AGAAAAAAG GATCTCAAGA AGATCCITTT ATCTTTTCTA 9350
 CGGGGTCTGA CGCTCAGTGG AACGAAAACT CACGTTAAGG GATTTTGTTC 9400
 ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA ATTTAAAAATG 9450
 35 AAGTTTAAA TCAATCTAAA GTATATATGA GTAAACITGG TCTGACAGTT 9500
 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG 9550
 TCATCCATAG TTGCTGACT CCCGCTCGTG TAGATAACTA CGATACGGGA 9600
 GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA GACCCACGCT 9650
 CACCGGCTCC AGATTTATCA GCAATAAACCC AGCCAGCCGG AAGGGCCGAG 9700
 40 CGCAGAAAGT GTCTTCAAC TTTATCCGCC TCCATCCAGT CTATTAATTG 9750
 TTGCGGGGAA GCTAGAGTAA GTAGTTGCC AGTTAATAGT TTGCGCAACG 9800
 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGTGTATG 9850
 GCTTCATTCA GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC 9900
 CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCTCCG ATCGTTGTCA 9950
 45 GAAGTAAGTT GGCCGCGAGT TTAATCACTA TGGTTATGGC AGCACTGCAT 10000
 AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG TGAAGGTGA 10050
 GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT 10100
 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAATTTA 10150
 AAAGTGCTCA TCATTGGAAA ACGTTCCTCG GGGCGAAAA TCTCAAGGAT 10200
 50 CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT GCACCAACT 10250
 GATCTTCAGC ATCTTTTACT TTCACGAGCG TTTCTGGGTG AGCAAAAAACA 10300
 GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC GGAATGTTG 10350
 AATACTCATA CTCTCTCTT TTCAATATTA TTGAAGCATT TATCAGGGT 10400
 ATGTCTCAT GAGCGGATAC ATATTGAAT GTATTAGAA AAATAAACAA 10450
 55 ATAGGGGTTT CGCGCACATT TCCCGAAAA GTGCCAC 10487

SEQ ID NO:33 (conalbumin polyA)
 60 tctgcoattg ctgcttctct tgccttctct cgtcactctg aatgtggctt ctctgactact
 gccacagcaa gaaataaaat ctcaaatctt aaatgggttt cctgaggttt tccaagagtc
 gtttaagcaca ttccttcccc agcacccttt gctgcaggcc agtgccaggc accaacttgg

ctactgctgc ccatgagaga aatccagttc aatattttcc aaagcaaaat ggattacata
 tgccctagat cctgattaac aggcgtttgt attatctagt gctttcgtt caeccagatt
 atccattgc ctccc

5

SEQ ID NO:34 (exemplary antibody light chain sequence)

1 gagctcgtga tgaccagac tccatctccc ctgtctgcct ctctgggaga cagagtcacc
 61 atcagttgca gggcaaatca ggacattagc aattatttaa actggtatca gcagaaacca
 121 gatggaaactg ttaaaactct gatctactac acatcaagat tacactcagg ggtcccatca
 10 181 aggttcagtg gcagtgggtc tggaaacagat tattctctca ccatagcaa cctggagcaa
 241 gaagattttg ccacttactt ttgccaacag ggtaatacgc ttccgtggag gtccgtggga
 301 ggcaccaacc tggaaatcaa acgggctgat gctgcaccaa ctgtatccat ctcccacca
 361 tccagtgagc agttaacatc tggaggtgcc tcagtcgtgt gcttcttgaa caacttctac
 421 cccaaaagaca tcaatgtcaa gtggaagatt gatggcagtg aacgacaaaa tggcgtctctg
 15 481 aacagttgga ctgatcagga cagcaaacag agcacctaca gcattgagcag caccctcacg
 541 ttgaccaagg acgagtatga acgacataac agctatacct gtgaggccac tcacaagaca
 601 tcaacttcac ccattgtcaa gagcttcac aggaatgagt gttaa

20

SEQ ID NO:35 (exemplary antibody heavy chain sequence)

1 ctgagtcag gacctggcct ggtggcgccc tcacagaacc tgtccatcac ttgcaactgc
 61 tctgggtttt cattaaccag ctatggtgta cactgggttc gccagcctcc aggaagggt
 121 ctggaatggc tgggagtaat atggactggt agaagcacia ctataattc ggctctcatg
 181 tccagactga gcattcagcaa agacaactcc aagagccaag ttttcttaaa aatgaacagt
 25 241 ctgcaactcg atgacacagc catctactac tgtggcagag ggggtctgat tacgtccttt
 301 gctatggaact actggggtca aggaacctca gtcaccgtct cctcagccaa aacgacccc
 361 ccattctgtct atccactggc ccttgatctc gctgcccaca cttaactccat ggtgacctg
 421 ggatgcctgg tcaagggtca tttccctgag ccagtgcag tgacctggaa ctctggatcc
 481 ctgtccagcg gtgtgcacac ctccctagct gtcctgcagt ctgacctcta cactctgagc
 30 541 agctcagtg ctgtccctc cagcacctgg ccagcgaga ccgtcactg caacgttgcc
 601 cccccgccca gcagcaccaa ggtggacaag aaaattgtgc ccagggttg tactagt

SEQ ID NO:36 (pTnMCS)

35 1 ctgacgcgcc ctgtagcgge gcattaagcg cggcgggtgt ggtgggttac cgcagcgtga
 61 ccgtacact tgcagcgccc ctacgcgcc ctcccttccc tttcttccc tcttttctg
 121 ccaogttccc cggcatcaga ttggtattg gccattgcat acgttgatc catacataa
 181 tatgtacatt tatattggct catgtccaa attaccgcca tgttgacatt gattattgac
 241 tagttattaa tagtaataa ttacggggtc attagtcat agcccatata tggagttccg
 40 301 cttacataa ctacggtaa atggccgcc tggctgacg cccaacgacc cccgccatt
 361 gacgtcaata atgacgtatg tcccatagt aacgccaata gggactttcc attgacgtca
 421 atgggtggag tatttacggt aaactgccc ctggcagta catcaagtgt atcatatgcc
 481 aagtagcccc cctattgacg tcaatggcg taatggccc gctggcatt atgcccagta
 541 catgacctta tgggactttc ctactggca gtacatctac gtattagta tgcctattac
 45 601 catggtgatg cgtttttggc agtacctcaa tggcggtgga tagcggttg acrcacgggg
 661 atttccaagt ctccacccc ttgacgtcaa tggagtttg ttttggcacc aaaatcaag
 721 ggactttcca aaatgtcgt acaactccc cccattgacg caaatgggcy gtaggcgtgt
 781 acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagagc

132

ATLUB02 133492.1

841 ccatccacgc tgttttgacc tccatagaag acacccggac ccatccacgc tccgcggccg
 901 ggaacgggac attggaacgc ggattccccc tgccaagagt gacgtaagta cgcctatag
 961 actctatagg caccaccctt tggtcttat gcatgctata ctgtttttgg cttggggcct
 1021 atacaccccc gcttccttat gctataggt atggtatagc ttacgctata ggtgtgggtt
 1081 attgaccatt attgaccact cccctattgg tgacgatact ttcattact aatccataac
 1141 atggctcttt gccacaacta tctctattgg ctatatgcca atactctgtc cttcagagac
 1201 tgacacggac tctgtatttt tacaggatgg ggtcccaatt attatttaca aattcacata
 1261 racacacacg cgcctccccg tgcccgagct ttttattaaa catagcgtgg gatctccacg
 1321 cgaatctcgg gtacgtgttc cggacatggg ctctctccg gtacgcgggg agcttccaca
 1381 tccgagccct ggtcccatgc ctccacggc tcatggtcgc tcggcagctc cttgtcccta
 1441 acagtggagg coagacttag gcacagcaca atgcccacca ccaccagtgt gccgcacaag
 1501 ggcgtggcgg tagggatgtg gctgaaaat gagcgtggag attgggctcg caccgctgac
 1561 gcagatggaa gacttaaggc agcggcagaa gaagatgcag gcagctgagt tgtttatct
 1621 tgataagagt cagaggtaac tccgcttgcc gtcgtgttaa cgggtggagg caggtagtc
 1681 tgagcagtag tctgtgtgct cgcgcgcgct accagacata atagctgaca gactaacaga
 1741 ctgttccctt ccatgggtct tttctgcagt caccgtcgga ccatgtcgca actcgatatt
 1801 ttacacgact ctcttaccac attctgcccc gaattacact taaaacgact caacagctta
 1861 acgttgcctt gccacgcatt acttgaactg aaaactctca ctcttaccga acttggcgtt
 1921 aacctgcaaa ccaagcgagc aacaaaacat aacatacaac gaatcgacog attgttaggt
 1981 aatcgtcacc tccacaaaga gcgactcgct gtataccgtt ggcagctag ctttatctgt
 2041 tccggcaata cgtgcccat tctactgtt gactggtctg atattctgta gcaaaaaacg
 2101 ctatggtat tgcgagotto agtcgacta caggctcgtt ctgttactct tcatgagaaa
 2161 ggcgtccccc tttcagagca atgttcaaa aaagctcatg accaatttct agccgacctt
 2221 ggcagatctc taccagtaga caccacacgc ctcatlgtca gtgactgtg ctttaaaagt
 2281 ccatgggata aatccgttga gaagctgggt tggactgtgt taagtogagt aagaggaaaa
 2341 gtacaatatg cagacctagg agcggaacaa tggaaaacct tcagcaacct acatgatag
 2401 tcatctagtc actcaagac tttaggctat aagaggctga ctaaaagcaa tcaactctca
 2461 tgcaaaatc tattgtataa atctcgctct aaagcgccga aaactcagcg ctgcacacgg
 2521 acctattgtc accaccgctc acctaaaatc tactcagcgt cggcaaaagg gcaatgggtt
 2581 ctagcaacta acttactcgt tgaatttcga acccccaaac aacttgttaa tatctatcg
 2641 aagcgaatgc agattgaaga aaccttccga gacttgaaaa gtcctgccta cggactaggc
 2701 ctacgccata gccgaacgag cagctcagag cgttttgata tcatgtgctt aatcgccctg
 2761 atgtctcaac laacalgttg gcttgcgggc gttcatgctc agaacaagg tgggacaaag
 2821 cacttccagg ctaacacagt cagaatctga aacgtactct caacagttcg cttaggcatg
 2881 gaagttttgc ggcattctgg ctacacaata acaaggggaag acttactcgt ggcgtcaacc
 2941 ctactagctc aaaaatttat cacacatggt tacgctttgg ggaattatg aggggatacg
 3001 tctagagcga tccgggactc cgggaaaagc gttggtgacc aaaggtgcct tttatcatca
 3061 ctttaaaaac aaaaaacat tactcagtcg ctgttataag cagcaattaa tlatgattga
 3121 tgctctacac acaacaaaaa ctgatttaac aaatgggttg tctgccttag aaagtatat
 3181 tgaaacttat ctgattataa ttattgataa taataaaaaa ctatcccta tcaagaaagt
 3241 gatgcctatc attggttggg atgaacttga aaaaatttag ccttgaatac attactggt
 3301 aggttaacgc cattgtcagc aaattgatcc aagagaacca acttaaaagt ttcctgacgg
 3361 aatgttaatt ctctgtgacc ctgagcactg atgaatcccc taatgatttt ggtaaaaatc
 3421 attaagttaa ggtgataaca catctgttca tatgatcccg gtaatgtgag ttagctcact
 3481 cattaggcac cccaggcttt acactttatg ctccggctc gtatgttgtg tggaaattgtg
 3541 agcgataaac aatttcacac aggaacacgc tatgaccatg attacgccaac gcgcgcaatt
 3601 aacctcactc aaagggaaac aaagctggag ctccacccgc gtggcgcccg ctctagaact
 3661 agtgatccc cgggctgca ggaattcgat atcaagctta tgcataccgc tgacctcgag
 3721 cggggggccc gtaaccaaatt cgcctlatag tgagtcttat tacgcgcgct cactggccgt
 3781 cgttttcaaa cgtcgtgact gggaaaaacc tggcgttacc caacttaac gccttgcagc
 3841 catccccct ttcgccaagt ggcgtaatag cgaagaggcc cgcaccgatc gcccttccca
 3901 cagtttgcgc agcctgaatg gcgaatggaa attgtaagcg ttaataatctt gttaaaaatc
 3961 gcgttaaat tttgttaaat cagctcattt ttaaccaat aggcggaat cggcaaaatc
 4021 ccttataaat caaaagaata gaccgagata ggggttagtg ttgtccagt ttggaacaag
 4081 agtccactat taaagaacgt ggaactccac gtcaaaaggc gaaaaaccgt ctatcagggc
 4141 gatggccccc tactccggga ctatatgaca agatgtgtat ccaccttaac ttaatgattt
 4201 ttaccaaact cattagggga ttcatcagtg ctccaggtca acgagaatta acattccgtc
 4261 aggaagcct atgatgatga tgtgtttaa aacttactca atggctgggt atgcatactg
 4321 caatacatgc gaaaaacta aaagagcttg cgtataaaaa aggcacattt attgtatctt
 4381 accgggctt tttattgagc ttgaaagata aataaaatag ataggtttta ttgaaagcta
 4441 aatcttcttt atcttaaaaa atgcccctctt ggggtatcaa gagggtcatt atatttcgcg
 4501 gaataacatc atttgggtgac gaataacta agcaactgtc tccgttttac tcccctgagc
 4561 ttgaggggtt aacatgaagg tcatcgatag caggataata atacagtaaa acgtcaaac
 4621 aataatccaa atccagccat cccaaattgg tagtgaatga ttataataa cagcaaacag
 4681 taatgggcca ataacacggc tgcatttggt aaggctcacc aataatccct gtaaaagcag
 4741 ttgctgatga cctcttggtt ggaatagcat cactccctgt aatgcaggtt aagcgatccc
 4801 accaccagcc aataaaatta aaacagggaa aactaaccaa ccttcagata taaacgctaa
 4861 aaaggcaaat gcactactat ctgcaataaa tccgagcagt actgcgctt ttcgcccat

4921 ttatgtgcta tttctctgc cacaaaggct tggaaactq aqtgtaaaag accaagaccc
 4981 gtaatgaaa gccaccatc atgctattca tcatcacgat ttctgtaata gccaccaccc
 5041 gtgctgatt ggctatcaat gcgctgaaat aataatcaac aaatggcacc gttaaataag
 5101 :gatgtatac cgatcagctt ttgttccctt tagtgagggt taattgcgcg cttggcgtaa
 5161 :catggteat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata
 5221 cgagccgaaa gcataaagt taaagccttg ggtgcctaata gagtgcgcta actcacatta
 5281 attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgcgtgcca gctgcattaa
 5341 tgaatcgccg aacgcgcggg gagagggcgt ttgcgtattg ggctgtcttc cgcttctctg
 5401 ctactcactc cgctgcgctc ggtcgttcgg ctgcggcgag cggatcage tcaactcaag
 5461 cgggtaatac ggttatccac agaatcaagg gataaccgag gaaagaacat gtgagcaaaa
 5521 ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt coataggctc
 5581 cgcctccctg acgagcatca caaaaatcga cgtcaagtc agaggtggcg aaaccgcaca
 5641 ggactataaa gataccaggc gtttccctct ggaagctccc tgcgtgcctc tctgtctcgg
 5701 accctgcgcg ttaccggata cctgtccgcc ttctccctt cgggaagcgt ggcgctttct
 5761 catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttgcgtccaa gctgggctgt
 5821 gtgcacgaac cccctgttca gcccgaccgc tgcgccttat cggtaacta tgcctctgta
 5881 tccaaccggt taagacacga cttatcgcca ctgcgcagc cactcgttaa caggattagc
 5941 agagcgaggc atgtaggcgg tgcacagag ttcttgaagt ggtggcctaa ctacgctac
 6001 actgaagaga cagattttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga
 6061 gttggtgagc cttgatccgg caaaccaaac accgctggta cgggtggttt tttgtttg
 6121 agcagcaga ttacgcgcag aaaaaagga tctcaagaag alcctttgat ctttctacg
 6181 gggctctgac ctcatggaa cgaatactca cgttaaggga ttttggtcat gagattatca
 6241 aagagctctc tcacctagat ccttttaaat taaaaatgaa gttttaate aatcaaatg
 6301 atatatagc aaacttggc tgcagttac caatgcttaa tcatgagggc acctatctca
 6361 gcgatctgtc tatttcttcc atccatagtc gctgaactcc cgtcgtgta gataactacg
 6421 :tacgggagg gcttaccatc tggcccccag gctgcaatga taccgcgaga cccacgctca
 6481 ccgctccagc atttatcagc aataaacccg ccagccggaa gggccgagcg cagaagtggg
 6541 cctgcaactt tatccgcctc catccagctc attaatgttt gccgggaagc tagagtaagt
 6601 agttcgccag ttaatagttt ggcgaacgtt gttgocattg ctacaggcat cgtgggtgca
 6661 cgtcgtcgtc ttggtatggc ttcatcagc tccggttccc aacgatcaag gcgagttaca
 6721 tgatcccca tgttgtgcaa aaaagcgggt agctccttcg gtctccgat cgttgtcaga
 6781 agtaagtggc cgcagtggt atcactcatg gttatggcag cactgcataa ttctcttact
 6841 gteatgccat ccgtaagatg cttttctgtg actggtgagt actcaacca gtcattctga
 6901 gaatagtga tggggcgacc gaggttgctc tgcccgcggt caatacggga taataccgcg
 6961 ccacatagca gaactttaaa agtgcctatc attggaatac gttcttcggg gcgaaaaatc
 7021 tcaaggatct taccgctgtt gagatccagt tcatgtaac ccaactcgtc acccaactga
 7081 tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagc aaggcaaaat
 7141 gccgcaaaaa agggataaag ggcgacacgg aatgttgaa tactcatact ctctctttt
 7201 caatattatt gaagcattta tcagggttat tgtctatga cgggatacat atttgatgt
 7261 atttagaaaa ataaacaaat aggggttccg cgcacatttc ccgaaaaagt gccac

SEQ ID NO:37 (chicken ovalbumin ehancer)
 cggggtgca gaaaaatgcc aggtgggacta tgaactcaca tccaaaggag
 cttgaccta tacctgattt tcttcaaat ggggaaccaa cacaatccca caaacagct
 45 cagagagaaa ccactactga tggctacagc accaagggtat gcaatggcaa tccattcgac
 attcatctgt gacctgagca aaatgattta tctctccatg aatgggtgct tcttccctc
 atgaaaggc aatttccaca ctacacatat gcaacaaaga caaacagaga acaattaatg
 tgcctcttcc taatgtcaaa attgtagtgg caaagaggag aacaaaatct caagttctga
 50 gtagggttta gtgattggat aagagggctt gacctgtgag ctacactgga cttcatatcc
 ttttgataa aaagtgcctt tataactttc aggtctccga gttcttatc atgagactgt
 tgggtttagg acagaccac aatgaaatgc ctggcatagg aaagggcagc agagccttag
 ctgacctttt cttgggacaa gcattgtcaa acaatgtgtg acaaaactat ttgtactgct
 tgcacagct gtgctgggca gggcaatcca ttgccacctc tccaggtaa ccttccaact
 55 gcaagaagat tgttgcttac tctctctaga

SEQ ID NO:38 (5' untranslated region)
 GTGGATCAACATACAGTAGAAAGCTGTATTGCCTTTAGCACTCAAGCTCAAAAGACAACACTCAGAGTTC
 ACC

60 SEQ ID NO:39 (putative cap site)
 ACATACAGCTAG AAAGCTGTAT TGCCTTTAGC ACTCAAGCTC AAAAGACAAC TCAGAGTTCA

65 SEQ ID NO:40 (fragment of ovalbumin promoter - chicken)
 GAGGTGAGAAT GGTTCCTTTA CTGTTTGTCATTCTATTAT TTCAATACAG
 AACATAGCT TCTATAACTG AAATATATTT GCTATTGTAT ATTATGATTG

TCCCTCGAAC CATGAACACT CCTCCAGCTG AATTTCACAA TTCCTCTGTC
 ATCTGCCAGG CCATTAAAGT ATTCAATGAA GATCTTTGAG GAACACTGCA
 AGTTCAIATC ATAAACACAT TTGAATTTGA GTATTGTTTT GCATTGTATG
 GAGCTATGTT TTGCTGTATC CTCAGAAAAA AAGTTTGTTA TAAAGCATTG
 5 ACACCCATAA AAGATAGAT TTAATATTTC CAGCTATAGG AAAGAAAGTG
 CGTCTGCTCT TCACTCTAGT CTCAGTTGGC TCCTTCACAT GCAIGCTTCT
 TTATTTCTCC TATTTTGTCA AGAAAAATAT AGGTCACGTC TTGTTCTCAC
 TTATGTCCTG CCTAGCATGG CTCAGATGCA CGTTGTAGAT ACAAGAAGGA
 TCAATGAAA CAGACTTCTG GTCGTTTACT ACAACCATAG TAATAAGCAC
 10 ACTAACIAAT AATTGCTAAT TATGTTTTCC ATCTCTAAGG TTTCCACATT
 TTTCTGTTTT CTTAAAGATC CCATTATCTG GTTCTAACTG AAGCTCAATG
 GAACATGAGC AATATTTCCC AGTCTTCTCT CCCATCCAAC AGTCTGTATG
 GATTAGCAGA ACAGGCAGAA AACACATTGT TACCCAGAAT TAAAACTAA
 TATTTGCTCT CCATTCAATC CAAAATGGAC CTATTTGAAAC TAAATCTAA
 15 CCCAATCCCA TTAATGATT TCTATGGCGT CAAAGGTCAA ACTTCTGAAG
 GGAACCTGTG GGTGGGTAC AATTCAGGCT ATATATTTCC CAGGGCTCAG
 C

SEQ ID NO:41 pInMCS (CMV-CHOVg-ent-ProInsulin-synPA)
 20 1 ctgacgcgcg ctgtagcggc gcattaagcg cggcggtgt ggtggttacg cgcagcgtga
 61 cgcctacact tgcacgcgcg ctgacgcgcg ctccttttgc tctcttccct tctcttctcg
 121 ccagcttcgc cggcatcaga ttggtatttg gccattgcac acgttgatcc catatcataa
 181 tatgtacatt tataattggc catgtccaac attaccgcca tgttgacatt gattattgac
 241 tagttattaa tagtaataaa ttacgggggtc attagtcat agcccatata tggagtcccg
 25 301 cgttacataa cttacggtaa atggccgcgc tggctgacgc cccaacgacc cccgccatt
 361 gacgtcaata atgacgtatg ttcccatagt aacgcataa gggactttcc attgacgtca
 421 attggttgga tatttacggt aaactgcaca cttggcagta catcaagtgt atcatatgcc
 481 aagtagcgcg cctattgacg tcaatgacgg taaatggccc gcttggtcatt atgcccagta
 541 catgaacctt tgggaacttc ctacttgcca gtacatctac gtattagta tgcattata
 30 601 catggtgatg cggttttggc agtacatcaa tgggcgttga tagcgttttg actcagggg
 661 atttccaaagt ctcaccccca ttgacgtcaa tggaggtlly ltllygcacc aaatcaacg
 721 ggaacttcca aaatgtcgta acaactcgc ccatgtgac caaatggggg gtaggcgtgt
 781 a:ggtaggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagagc
 841 c:atccacgc tgttttgacc tccatagaag acacgggac cgtaccgac tccgcggcgc
 35 901 ggaacggtgc attggaaacg ggaattcccg tgcgaagat gacgtaaata ccgctatag
 961 a:ctctatagg cacacccctt tggctcttat gcatgctata ctgtttttgg cttggggcct
 1021 atacaccccg ctttccctat gctatagggt atggtatagc ttacgctata ggttggtgt
 1081 attgaccatt attgacact cccctattgg tgacgatact ttccattact aatccataac
 1141 ttggtctctt gccacaacta tctctattgg ctatagcca atactctgtc ctcagagac
 40 1201 (gacacggac tctgtatttt tacaggatgg ggtccattt attattaca aattcacata
 1261 tacaacaacg cgttcccccg tgcgcgcagt ttttattaaa catagcgttg gatctccacg
 1321 cgaatctcgg gtacgtgttc cggacatggg cctcttctcg gtacggcgg agcttccaca
 1381 tccgagccct ggtcccatgc ctcacgcggc tcatggctgc tccgagctc cttgctcta
 1441 acagtggagg ccagacttag gcacagcaca atgcccacca ccaccagtgt gccgcacaag
 45 1501 gccgtggcgg tagggtatgt gtctgaaat gagcgtggag attgggctcg cacggctgac
 1561 gcagatggaa gacttaaggg agcgcagaa gaagatcgag gcagctgagt tgttgtatc
 1621 tgataagagt cagaggtaac tccgttgcg gtgctgttaa cgttgagggg cagtgtatgc
 1681 tgagcagtae tggttgtctg cgcgcgcgcg accagacata atagctgaca gactaacaga
 1741 ctgttctctt ccatgggtct tttctgcagt caccgtcgga ccatgtgoga actcgatatt
 50 1801 ttacagact cttcttacc aattctgcgc gaattacact taaaacgact caacagctta
 1861 acgttggtgt gccacgcatt acttgactgt aaaaacttca cttctacga acttgccgt
 1921 aacctgcaca ccaagcgag acaaaaacat aacatcaaac gaatcgacg attgttaggt
 1981 aatcgtcacc tccacaaga gcagctcgct gtataccgtt ggcagctag ctttatctgt
 2041 tccggcaata cgaatgccat tgtacttgtt gactggtctg atattctgga gcaaaaacga
 55 2101 cttatggtat tgcgagcttc agtcgcacta caggtctgct ctgttactct ttatgagaaa
 2161 gcgttcccg tttcagagca atgttcaaa aaagctcatg accaatttct agccgacctt
 2221 cgcagcttc taccagtaa caccacacg ctcattgcta gtgactctgg ctttaagtg
 2281 ccatggtata aatcogttga gaagctgggt tggtaactgt taagtcagt aagagcaaaa
 2341 gtacaaatg cagacatagg agcggaaac tggaaaccta tcaagcaact acatgatga
 2401 reatctagtc actcaagac ttttagctat aagaggctga ctaaaagcaa tccaatctca
 60 2461 (gccaatttc tattgtataa atctcgtct aaaggccgaa aaaaatcagc ctcgacacgg
 2521 cctcattgtc accacccgct accataaate taactcagct cggcaaaagg gccattgggt
 2581 ctacgaacta acttaacctgt tgaatttoga acaccaaac aacttgttaa tatctattcg
 2641 aagcgaatgc agattgaaga aaccttcga gaactgaaaa gtctcgtcta cggactaggg
 65 2701 ctacgcata gccgaacgag cagctcagag cgttttgata tcatgctgct aatcgccctg
 2761 ctgttcaac taacatgttg gcttgcgggc gttcatgctc agaaacaagg ttgggacaa

135

ATLLIB02 133492.1

2821 cacttccagg ctaaacacagt cagaaatcga aacgtactct caacagttcg cttaggcacg
 2881 gaagttttgc ggcattctgg ctacacaata acaagggaag acttactcgt ggtgcacac
 2941 ctactagctc aaaatttatt cacacatggt tacgctttgg ggaattatg aggggatcgc
 3001 tctagagcga tccgggatct cgggaaagc gttggtgacc aaagggtgct ttatcatca
 3061 ctttaaaat aaaaaacaat tactcagtg cgtttataag cagcaattaa ttatgattga
 3121 tgcctacatc acaaaaaaaa ctgatttaac aaatggttgg tctgccttag aaagtatatt
 3181 tgaacattat ctgtattata ttattgataa taataaaaac cttatcccta tccaagaagt
 3241 gatgcctatc attggttggg atgaacttga aaaaaattag cttgaatc atactggtga
 3301 aggttaacgc cattgtcagc aaattgatcc aagagaacca acttaagct ttctgacgg
 3361 aatgttaatt ctctgtgacc ctgagcactg atgaatcccc taatgatttt ggtaaaaatc
 3421 attaaagtta ggtggataga catcttctga tatgatcccg gtaatgtgag tttagtcaat
 3481 cattaggcac cccaggcttt acactttatg ctccggctc gtatgttgtg tgaatttgtg
 3541 agcggataac aatttcacac aggaacacgc tatgaccatg attacgccaa gcgcgaatt
 3601 aacccctcact aaaggaaca aaagctggag ctccaccgag gtggcggcgc ctctagaact
 3661 agtggatccc cccggcatca gatggctatg tggccattgc atactgtga tccatcatc
 3721 aatagtaca tttatatggg ctcatgtcca acattaccgc catgttgaca ttgattattg
 3781 actagtattt aatagtaatc aattacgggg tcatagtctc atagcccata tatggagtcc
 3841 cgcgttcatc aacttacggt aaatggcccg cctggctgac cgcaccaaga ccccccacca
 3901 ttgacgtcaa taatgaagta tgttcccata gtaacgcaa tagggacttt ccatgaagct
 3961 caatgggtgg agtatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg
 4021 ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgctgggca ttatgccacg
 4081 tacatgacct tatgggactt tctacttgg cagtacatct acgtattagt cgtctattt
 4141 accatggtga tgcggttttg gcagctacac aatggcggtg gatagcgtt tgaactcagg
 4201 ggaatttccaa gtctccaccc cattgaagtc aatggagttt tgttttggca ccaaaatcaa
 4261 cgggactttc caaaatgtcg taacaactcc gcccatatga cgcacaaagg cggtaggcgt
 4321 gtacgggtgg aggtcttatat aagcagagct cgtttatgta accgtcagat cgcctggaga
 4381 cgcctaccac gctgttttga cctccataga agacacccgg accgatccag tctcccgagg
 4441 cgggaacggg gcatgtgaac gcggattccc cgtgccaaag gtgcgttaag taccgcctat
 4501 agactctata ggcacacccc ttgggtctct atgcatgcta tactgttttt ggcttggggc
 4561 ctatcacccc ccgcttctct atgctatagg tgaatgtata gcttagccta taggtgtggg
 4621 ttattgacaa ttattgacaa ctcccctatt ggtgacgata ctttccatta ctaatccata
 4681 acatggtctc ttgccaacac tatctctatt ggcttatatg caatctctgt tcttccagag
 4741 actgacacgg actctgtatt ttacaggat ggggtcccat ttattattta caaattcaaa
 4801 tatacaacaa cgcctctccc cgtgcccgca gtttttatta aacatagcgt gggatctcca
 4861 cgcgaatctc ggttactgtg tccggacatg ggtctctctc cgttagcggc ggagcttcca
 4921 catccagacc ctgggtcccat gccctcagcg gctcatggtc gctcggcagc tctctgctcc
 4981 taacagtgga ggccagacgt agccacagca caatgccac caccaccagt gtgcgcgaca
 5041 aggcctggc ggtagggtat ggtctgtaaa atgagcgtgg agattgggct cgcacggctg
 5101 acgcagatgg aagacttaag gtagcggcag aagaagatgc aggcagctga gttgttgtat
 5161 tctgataaga gtccagagta actccggttg cgtgtgtgtt aacggtggag ggcagtgtag
 5221 tctgagcagt actcgttctt gccgcggcgg ccacagacaa taatagctga cagactaaca
 5281 gaetgttctt ttccatgggt cttttctgca gtcacgctcg ggtccatgg gctccatcgg
 5341 cgcagcaagc atggaatttt gttttgtatg attcaaggag ctcaaatgac accatgccaa
 5401 tgagaaacac ttctactgac coattgcatc atgtcagct ctgagccatgg tatacctggg
 5461 tgcaaaagac agcaccagga cacagataaa taaggttgtt cgttttgata aacttccagg
 5521 attcggagac agtattgaag ctcaagtgtg cacactctga aacgttcaat ctccacttag
 5581 agacatcttc aaccaaatac ccaaaccaaa tgatgtttat togttcagcc ttgccagtag
 5641 actttatgct gaagagagat acccaatcct gccagaatac ttgcagctgt tgaaggaaat
 5701 gtatagagga ggtctgggac ctatcaactt tcaaacagct gcagatccag ccagagagct
 5761 catcaattcc tgggtagaaa gtccagacaa tggaaatcac agaaatgtcc ttccgcaag
 5821 ctccgtggat tctcaaacgt caatgggtct ggtaaatgcc attgtcttca aaggactgtg
 5881 ggagaaaaaa ttttaaggatg aagacacaca agcaatgcct ttccagatga ctgagcaaga
 5941 aagcaaacct gtgcagatga tgtaccagat tgggtttatt agagtggcat caatggcttc
 6001 tgagaaaatg aagatcctgg agcttccatt tgcagtgagg acaatgagca tggttgtgtc
 6061 gttgcctgat gaagtctcag gccctgagca gcttgagagt ataatacct ttgaaaaact
 6121 gaetgaatgg accagtctta atgttatgga agagaggaag atcaaatgtt acttacctcg
 6181 catgaagatg gaggaaaaat acaacctcac atctgtctta atggctatgg gcatctactg
 6241 cgtgttttag tcttcagcca atctgtctgg catctctca gcagagagcc tgaagatalc
 6301 tcaagctgct catgcagcac atgcagaaat caatgaagca ggcagagagg ttgtagggct
 6361 agcagaggct ggagtggatg ctgcaagcgt ctctgaagaa tttagggctg accatccatt
 6421 cctcttctgt atcaagcaca tgcacaacaa cgccttctc ttctttggca gatgtgttcc
 6481 ccgcccagag cagatgacgc accagcagat gacgcaccag cagatgacgc accagcagat
 6541 gacgcaccag cagatgacgc accagcagat gacgcaccaa catgtatcct gaaggctct
 6601 tgtggtgga tgcgctgct ggaatgagat gacaaaattg tgaaccaaca cctgtgcggc
 6661 tccacactgg tgaagctct ctacttagtg tgcggggaac gaggtctctt ctacacaccc
 6721 aagaccgccc gggaggcaga ggaactgacg gtggggcagg tggagctggg cgggggccc
 6781 ggtgacgca gccctgagcc ctggccctg gaggggtccc tgcagaagcg tggcattgtg
 6841 gaacaatgct gtaccagcat ctgctccctc taccagctgg agaactactg caactagggc

6901 gcttaaaagg cgaattatcg cggccgctct agaccaggcg cctggatcca gatcactttc
 6951 ggctaataaa agatcagagc tctagagatc tgtgtgttgg tttttgtgg atctgctgtg
 7021 ccttctagtt gccagccatc tgtgtgttgc ccttcccccg tgccttccctt gacctcggaa
 7081 ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgcttgagt
 7141 aggtgtcatt ctattctggg ggtgggggtg gggcagcaca gcaaggggga ggaattggaa
 7201 gacaatagca ggcattgctg ggcattgctg ggcattgctg ggcattgctg ggcattgctg
 7261 ctctctctct ctctctctct ctctcgttac ctctctcgag ggggggcccg gtaccaaat
 7321 cgcctatag tgagtcgtat tacgctgctt cactggcctt cgttttaca cgtcgtgact
 7381 gggaaaaacc tggcgttacc caacttaate gccttgacg acatccccct ttgcgcagct
 7441 ggcataatag cgaagaggcc cgcaccgac gcccttccca acagtgtgcg agcctgaatg
 7501 gcgaatggaa attgtaagcg ttaattttt gttaaaatc gcgttaaat tttgttaaat
 7561 cagctcattt ttaaccaat aggcgaaat cggcaaaatc ccttataaat caaagaata
 7621 gaccagata gggttgagt ttgttccagt ttggaacaag agtccactat taaagaacgt
 7681 ggcctccaac gtcaaaaggg gaaaaacgt ctatcagggc gatggccac taactcggga
 7741 tcatatgaca agatgtgtat ccacttaac ttaatgattt ttacaaaaat cattagggga
 7801 tcatcagtag ctacgggtca acgagaatta acattccgtc aggaagacct atgatgatga
 7861 tgtgtctaaa aacttactca atggctggtt atgcataatc caatacagc gaaaaccta
 7921 aaagagcttg cagataaaaa aggcacattt attgctattt accggtgctt tttatgagc
 7981 tgaagagata aataaaatag ataggtttta tttgaagcta aatctcttt atctgaaaaa
 8041 atgcctcctt gggttatcaa gagggtcatt atatttcgag gaataacatc atttggtgac
 8101 gaaataacta agcactgttc tctctgttac tccctgagc ttgaggggtt aacatgaagg
 8161 tcatcgatag caggataata atacagtaaa acgctaaacc aataatccaa ttccagccat
 8221 ccaaatagg tagtgaatga ttaataataa cagcaaaacg taatgggcca ataacacggg
 8281 ttgctattgt aaggctcacc aataatcctt gtaagcacc ttgctgaat cctttgttt
 8341 ggcataagat cactcctgtt aatgcaggta aagcgatccc accaccagcc aataaaatta
 8401 aaacagggaa aactaaccaa ccttcagata taaacgctaa aaaggcaaat gcactactat
 8461 ctgcaataaa tccgagcagt actgcggtt tttgcccac ttagtggcta ttcttctgac
 8521 cacaagagct tgaataactg agtgaataag accaagaccc gtaatgaaa gccaacatc
 8581 atgctattca tcatcagat ttctgtaata gccacacacc gtgctggatt ggctatcaat
 8641 gcgctgaaat aataatcaac aaatggcacc gtaataaag tgatgtatcc cgtacagctt
 8701 ttgttccctt tagtgagggt taatlgtcgc ctltggcgtaa tcatggtcat agctgtttcc
 8761 tgtgtgaaat tggtatccgc tcaaatatcc acacaacata cagacgggaa gcataaagt
 8821 taaagcctgg ggtgctaat gagtgaatga actcacatta attgctgtgc gctcactgcc
 8881 cgctttccag tccggaaacc tgtgtgcca gctgcattaa tgaatcgcc aacgcgggg
 8941 gagaggcgtt ttgctattg ggcgtctctt cgttctctg cctactgact cgtgcgctc
 9001 ggtcgttcgg ctgcggcgag cggtatcagc tcaatcaaa ggcgtaaac ggttatccac
 9061 agaatcaggg gataacgcag gaaagacat gtgagcaaaa ggccagcaaa aggcaggaa
 9121 cgttaaaaag cgcgcttgc tggcgttttt ccataggctc cgcctccctg acgagcatca
 9181 caaaaatcga cgtcagatc agaggtggcg aaaaaccgca ggcataataa gataccaggc
 9241 gtttccctct ggaagctccc tctgtcgcctc tctgttccg accctgcgcg ttaccggata
 9301 cctgtccgcc tttctcctt cgggaagcgt ggcgtttct catagctcac gctgtaggta
 9361 tctcagttcg gtgtaggctg ttgcctccaa gctgggctgt gtgcacgaac ccccggttca
 9421 gccgcagcgc tgcgccttat ccggttaacta tctgtttgag tccaaaccg taagacacga
 9481 cttatgccca ctggcagcag ccaactgtaa caggattagc agagcgaggt atgtaggcgg
 9541 tgctacagag tcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg
 9601 tatctgcgtc ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgacccgg
 9661 caaaacaaac accgctggtt cgggtggttt tttgtttgc aagcagcaga ttacgcggag
 9721 aaaaaaagga tctcaagaag atcctttgat cttttctacg ggtgtctgag ctcagtgga
 9781 cgaaaactca cgttaaggga ttttggctat gagattatca aaaggatct tcacctagat
 9841 ccttttaaat taaaaatgaa gttttaaatc aatcaaatg atatatgagt aaacttggtc
 9901 tgacagttac caatgcttaa tcaatgaggg accatctcga gogatctgtc tatttctgtc
 9961 atccatagtt gccctgctc cgtctgtgta gataactacg atacgggagg gottaccatc
 10021 tggcccccagt gctgcaatga taccgagaga ccaacgctca cggctccag atttatcagc
 10081 aataaaccag ccagccggaa gggccgagcg cagaagtgtt cctgcaactt tatccgcctc
 10141 catccagctt attaatgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt
 10201 ggcgaacggt gttgcattg ctacaggcat cgtgtgtgta cgtcgtcgt ttggtatggc
 10261 ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca tgtgtgcaa
 10321 aaaaagcgggt agctccttcg gtctccgat cgtgtctaga agtaagtgg ccgcagltgt
 10381 atcactcatg gttatggcag cactgcataa ttctcttact gtcagtcctt ccgtgaatg
 10441 cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtga tgcggcgacc
 10501 gagtgtctct tgcggcggt caatacggga taataccgag ccacatagca gaactttaa
 10561 agtgcctcgc attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt
 10621 gagatccagt tcatgtaaac ccaactcgtc acccaactga tcttcagcat cttttacttt
 10681 caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gcccaaaaa agggaaataa
 10741 ggcgacacgg aatgttgaa tactcatact cttcctttt caatattat gaagcattta
 10801 tcagggttat tgcctcatga cgggatacat attgaaagt atttagaaaa ataaacaaat
 10861 aggggttcgg cgcacatttc cccgaaaagt gccac

SEQ ID NO:42 (pTnMOD (CMV-CHOVg-ent-ProInsulin-synPA))

1 ctgacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtggttaac cgcagcgtga
 5 61 ccgctacact tgcagcggc ctgagcggc ctcctttcgc tttcttccct tcccttctcg
 121 ccacgttcgc cggcatcaga ttgctatttg gccattgcac acgttgatc catatcataa
 181 tatgtacatt tatattggct catgtccaac attaccgcca tggtagacatt gattattgac
 241 tagttattaa tagtaataaa ttacgggggc attagtccat agcccatata tggagttcgc
 301 cgttacataa cttacggtaa atggccggcc tggctgaccg cccaacgaac cccgccatt
 361 gacgtcaata atgacgtatg ttcccatagt aacgccaaata gggactttcc attgacgtca
 10 421 atgggtggag tatttacggt aaactgccca cttggcagta catcaagttg atcatatgcc
 481 aagtagcggc cctattgacg tcaatgacgg taaatggccc gccctggcatt atggccagta
 541 catgacctta tgggaatttc ctacttggca gtacatctac gtattagtca tgcctattac
 601 catgggtatg cgggttttgg agtacatcaa tgggcgttga tagcggtttg acccagggg
 661 atttccaaat ctcaccccca ttgacgtcaa tgggagtttg ttttggcaac aaatcaaacg
 721 ggaactttcca aaatgtcgta acaactcgcg cccattgacg caaatggggc gttagcggtg
 15 781 acgggtggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg
 841 ccatccacgc tgttttgacc tccatagaag acaccgggac cgtaccagcc tcccgccgcg
 901 ggaacgggtg attggaacgc ggaattcccg tgccaagagt gacgtaaata cgcctatag
 961 actctatagg cacacccctt tggctcttat gcatgcta ctgttttttg cttggggcct
 1021 atacaacccc gcttccttat gctatagggt atggtatagc tttagcctata ggtgtgggtt
 1081 attgaccatt attgaccact cccctattgg tgaagatact ttccattact aalcataaac
 1141 atggctctttt gccacaacta tctctatttg ctatattgca atactctgtc cttcagagac
 1201 tgacacggac tctgtatttt tacaggatgg ggtccattt attatttaca aattcaacta
 1261 tacaacaacg ccgtcccccg tgcccgcagt ttttattaaa catagcgttg gatctccacg
 1321 cgaatctcgg gtacgtgttc cggacatggg ctcttctccg ttagcggcgg agcttcacaa
 1381 tccagacccct ggtccctatg ctcacggcgc tcatgttgcg tcggcagctc cttgctccta
 1441 acagctggag ccagacttag gcacagcaca atgcccacca ccaccagtg gcgcacaaag
 1501 ccgctggcgg tacggtatgt gtctgaaaat gacgtggag attggcgtcg cagcgctgac
 1561 gcagatggaa gacttaaggc agcggcagaa gaagatgcag gcagctgagt tglgtatttc
 1621 tgataaagat cagaggtaac tcccgttgcg gtgctgttaa cgggtggaggc cagtgtatgc
 1681 tgagcagtag tcgttgcgtc cgcggcgccc accagacata atagctgaca gactaacaga
 1741 ctgttctctt ccatgggtct tttctgcagt caccgtcgga ccatgtgtga actlgtatt
 1801 ttacatgatt ctctttacaa attctgcccc gaattacact taaaacgact caacagctta
 1861 acgttggctt gccacgcatt acttgactgt aaaactctca ctcttaccga acttggccgt
 1921 aactctgcaa ccaagcgag acaaaacat aacatcaaac gaatcgaccg altgttaggt
 1981 aatcgtaacc tccacaaga gcgactcgct gtatccggtt ggcattgtag cttatctctg
 2041 tccggcaata cgtatcccat tgtacttgtt gactgtgtgt atattcgtga gcaaaaaagc
 2101 ctattggtat tgcgagcttc agtcgcacta caggtcgct ctgttactct ttatgagaaa
 2161 gcgttcccgc tttcagacga atgttcaag aaagctcatg accaatttct agccgcacct
 2221 gcgagcattc taccgagtaa caaccacccg ctcattgtca gtgatgtcgg ctttaaagtg
 2281 ccatggtata aatccgttga gaagctgggt tggtagctgt taagtcgagt aagaggaaaa
 2341 gtacaatatg cagacctagg agcggaaaaa tggaaaccta tcagcaactt acatgatatg
 2401 tcatctagtc actcaagac tttaggctat aagaggctga ctaaaagcaa tccaattctca
 2461 tgccaaatcc tatlgtataa atctcgctct aaaggccgaa aaatcagcg ctcgacacgg
 2521 actcattgtc accacccgtc acctaaaatc tactcagcgt cggcaaaagga gccatgggtt
 2581 ctgacaaacta acttactctg tgaatttcga acacccaaac aacttgttaa tatctattcg
 2641 aagcgaatgc agattgaaga aaccttcga gacttgaaaa gtcttgctta oggactaggc
 2701 ctacgccata gccgaacgag cagctcagag cgttttgata tcatgtgtgt aatcgccctg
 2761 atgcttcaac taacatgttg gcttgggggc gttcatgctc agaaacaagg ttgggacaa
 2821 cacttccagg ctaacacagt cagaatacga aacgtactct caacagttcg cttaggcatg
 2881 gaagttttgc ggcattcttg ctacacaata acaagggaag acttactctg gcttgcaaac
 2941 ctactagctc aaaatttatt cacacatggt taagcttttg ggaattatg ataattgatc
 3001 agatcacttc tggctaataa aagatcagag ctctagagat ctgtgtgtgt gtttttttgg
 3061 gatctgtgtg gccttctagt tgcacgcat ctgttgtttg cccctccccg gtgccttctc
 3121 tgacccctga aggtgccact cccactgtcc tttcttaata aaatgaggaa attgcatcgc
 3181 attgtctgag taggtgtcat tctattctgg ggggtggggg ggggcagcac agcaaggggg
 3241 aggattggga agacaatagc aggcattgct gggatggcgt gggctctatg ggtacctctc
 3301 tctctctctc tctctctctc tctctctctc tctctcggtc cctctctctc tctctctctc
 3361 tctctctctc tctctctctc tctgtaccag gtgctgaaga attgacccgg tgacaaaagg
 3421 tgccttttat catcacttta aaaataaaaa acaattactc agtgccctgtt ataagcagca
 3481 attaatatg attgatgcct acatcacaac aaaaactgat ttaacaaatg gttgtgtctg
 3541 cttagaagat atatttgaac attatcttga ttatattatt gataataata aaaaccttat
 3601 ccatatccaa gaagtgatgc ctatcattgg ttggaatgaa cttgaaaaaa attagccttg
 3661 aatacattac tggtaaggtg aacgcatlgt tcagcaaat gatccaagag aaccaactta
 3721 aagctttccct gacggaatgt taattctctg tgacccctgag cactgatgaa tccctaatg
 3781 attttggtaa aaatcattaa gtttaaggtg atacacatct tgtcatatga tcccggttaat
 3841 gtgagtttag tcaactatta ggcacccag gcttcaact ttatgcttcc ggctcgtatg

3901 ttgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac
 3951 gccagcgcg caattaaccc tcaactaaagg gaacaaaagg tggagctcca cgcggtggc
 4021 ggccgctcta gaactagtg atccccggg catcagattg getattggcc attgcatacg
 4081 ttgtatccat atcataatat gtacatttat attggctcat gtccaacatt accggcatgt
 4141 tgacattgat tattgactag ttactaatag taatcaattt cggggctatt agtctatagc
 4201 ccatatatgg agttccgctg tacataactt acggtaaatg gcccgcttgg ctgaccgccc
 4261 aacgaccccc gcccatggac gtcaataatg acgtatgttc ccatagtaac gccaataggg
 4321 actttccatt gaactcaatg ggtggagtat ttacggtaaa ctgccactt ggcagtacat
 4381 caagtctatc atatgccaa gtaacccctt attgaactga atgacggtta atggcccgcc
 4441 tggcattatg cccagtaaat gaacttatgg gaactttcta ctggcagta catctacgta
 4501 ttatgtcatg ctattaccat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag
 4561 cgggttgact cagcggtatt tccaaagtct caccctattg acgtcaatgg gagttrgttt
 4621 tggcaccaaa atcaacggga ctttccaaaa tgcgttaaca actccgcccc attgacgcaa
 4681 atggcggtla ggcgtgaag gtgggaggtc tatataagca gagctcgttt agtgaaccgt
 4741 cagatcgctt ggagagcaca tccacgctgt tttgacctcc atagaagaca cggggacgga
 4801 tccagctccc cgggcgggga acggtgcatt ggaacgcgga ttccccgtgc caagatggac
 4861 gtaagtacgg cctatagact ctataggcac acccctttgg ctcttatgca tgcataactg
 4921 tttttggctt ggggcctata ccccccgct tcttatgtat ataggtgatg gtaagtctta
 4981 gccctatagg gtgggttatt gaacttatgt gaactctccc ctattgggtg cgtacttttc
 5041 cattatcaat ccaatacatg gctctttgcc acaactatct ctattggcta tatgccaata
 5101 cctctgctct cagagactga caggaactct gtatttttac aggtatgggt cccatttatc
 5161 atttacaatc tccatataac aacaacggcg tccccgtgc cgcagttttt tattataaat
 5221 agcgtgggat cttccacgga atctcgggta cgtgttccgg acatgggctc ttctcgggta
 5281 cggcgcgagc ttccacatcc gagccctggg ccatatgctc cagcggtcca tggctcctg
 5341 gcagctcctt gctcctaaca gtggagggca gaactaggca cagcaaatg cccacaccgt
 5401 ccagctgtgc gccacaagcc gtggcggtag ggtatgtgtc tgaanaatgag cgtggagatt
 5461 ggcgtcgcac ggcgtgacga gatggaagac ttaaggcagc ggcagaagaa gatgacgga
 5521 gctgagttgt tgtattctga taagagtcag aggtactccc cgttgcgggt ctgttaacgg
 5581 tggagggcag tgtattctga gcagtaactg ttgctgcggc cggcgccacc agacataata
 5641 gctgacagac taacagactg ttcttttcca tgggtctttt ctgcagtcac cgtcgggac
 5701 atcggctccc atcggcgag caagcatgga attttgtttt gatgtattca aggaactcaa
 5761 agtccacatc gccaatgaga acatcttcta ctgccctatt gccatcatgt cagctctagc
 5821 catggtatcc ctgggtgcaa aagacagcac caggacacag ataaataaag ttgttcgctt
 5881 tgataaactt ccaggattcg gagacagtat tgaagctcag tgtggcacat ctgtataact
 5941 tcaactcttc cttagagaca tctccaacca aatcaccaaa ccaaatgatg tttattcgtt
 6001 cagcctttgc agtagacttt atgctgaaga gagataccca atcctgccag aatacttga
 6061 gtgtgtgaag gaactgtata gaggaggtct ggaactatc aactttcaaa cagctgcaga
 6121 tcaagccaga gagctcatca attcctgggt agaaagttag acaaatggaa ttatcagaaa
 6181 tgtctctcag ccaagctccg tggatttcca aactgcaatg gttctggtta atgcaatgt
 6241 ctccaagga clglgggaga aaacatttaa ggtatgaagc acacaagcaa tgcctttcag
 6301 agtgactgag caagaaagca aactgtgca gatgatgac cagattgggt tatttagagt
 6361 ggcataaatg gcttctgaga aaatgaagat cctggagctt ccatttgcca gtgggcaaat
 6421 gagcatgttg gtgctgttgc ctgatgaagt ctacggcttt gagcagcttg agagtataat
 6481 caactttgaa aaactgactg aatggaccag ttctaatgtt atggaagaga ggaagatcaa
 6541 agtgaactta cctcgcatga agatggagga aaaaataaac ctacatctg tcttaattgg
 6601 tatggcattt actgacgtgt ttagctcttc agccaatctg tctggcatct cctcagcaga
 6661 gagcctgaag atatctcaag ctgtccatgc agcacatgca gaaatcaatg aagcaggcag
 6721 agaggttgta ggcgtcagcag aggcgtggag ggtatgtgca agcgtctctg aagaatttag
 6781 ggcctgacct ccatctctct tctgtatcaa gcacatgca accaacgccc ttctctctct
 6841 tggcagatgt gtttcccgcg gccagcagat gacgcaccag cagatgacgc accagcagat
 6901 gacgcaccag cagatgacgc accagcagat gacgcaccag cagatgacgc aacaacatgt
 6961 atctgaaa gctctgttgg ctggatggcg ctgctggatg acgatgacaa atttgtgaac
 7021 caacacctgt gcggctcaca cctggtggaa gctctctacc tagtgtcggg ggaacgagc
 7081 ttcttttaca caccacagac ccgcccggag gcagaggacc tgcaggtggg gcaggtggag
 7141 ctggcggggg gccctgggtg aggcagcctg cagcccttgg ccttggaggg gtccctgagc
 7201 aagcgtggca ttgtggaaca atgctgtacc agcatctgct cctctacca gctggagaac
 7261 tactgcaact agggcgctta aaggcgcaat tatcgcgccc gctctagacc aggcgctgg
 7321 atccagatca ctctcggtta ataaaagatc agagctctag agatctgtgt gttgggtttt
 7381 tgttgatctg ctgtgcttcc tagttgccag ccatctgttg ttggcccttc ccccgctcct
 7441 tcttgacccc tggaaagtgc caactccact gtcctttcct aataaaatga ggaatttga
 7501 tgcattgttc tgaagagtg tcatctatt ctgggggggt ggggtgggga gcacagcag
 7561 ggggaggatt ggaagacaa tagcaggcat gctgggagtg cgttgggttc tatgggtacc
 7621 tctctctctc tctctctctc tctctctctc tctctctctc ggcactctct ctgagggggg
 7681 gggccggtag ccaattcgcc ctatagttag tctattatcg cgcgtcact ggcgtcgtt
 7741 ttacaacgct gtgactggga aaacctggc gttacccaac ttaactcgct tgcagacat
 7801 ccccccttgc ccagctggcg taatagcgaa gagggcgca cagatcgccc tccccacag
 7861 ttgcgcagcc tgaatggcga atggaaattg taagcgttaa tattttgta aaattcgctg
 7921 taaattcttg ttaaatcagc tcattttcta accaataggc cgaatcgcc aaattccctt

139

ATLLB02 133492.1

7991 ataatcaaa agaataagacc gagatagggc tgagtgtgtg tccagtttgg aacaagaagc
 8041 cactattaaa gaacgtggac tccacagtc aagggcgaaa aacgcgtctat cagggcgatg
 8101 gccactact cgggatcat atgacaagat gtgtatccac cttacttaa tgatttttac
 8161 caaatcatt aggggatcca tcatgtctca gggccaagca gaattacat tccgtcagga
 8221 aagcttatga tgatgatgtg cttaaaaact tactcaatgg ctggttatgc atatcgcaat
 8281 acatcgaaaa aacctaaaag agcttgccga taaaaaaggc caatttatg ctatttacg
 8341 cggcttttta ttgagcttga aagataaata aatatagatg gttttattg aagctaaatc
 8401 ttctttatcg taaaaaatgc cctcttgggt tatcaagagg gtctattat ttgcgggaat
 8461 aacatcattt ggtgacgaaa taactaagca cttgtctcct gtttactccc ctgagcttga
 8521 ggggttaaca tgaaggtcat cgtagcagg ataataatcc agtaaaacgc taaaccaata
 8581 atccaaatcc agccatccca aattggtagt gaattgattt aaataacagc aaacagtaat
 8641 gggccaataa caccggttgc attggttaagg ctcaccaata atccctgtaa agcaccctgc
 8701 tgatgactct ttgtttggat agacatcact cctgttaagt caggtaaaag gatcccaaca
 8761 ccagccaata aattlaaaac agggaaaact aaccaacctt cagatataaa cgtaaaaaag
 8821 gcaaatgcac tactatctgc aataaatccg agcagtagct cgttttttc gccatttag
 8881 tggctattct tctgccaca aaggcttggg atactgagtg taaaagacca agaccgtgac
 8941 tgaagaagcca accatcatgc tttcatcat cagcatttct gtaataagac cacaccgtgc
 9001 tggattggct atcaatgcgc tgaataata atcaacaagt ggcacgttta aataagtgat
 9061 gtataccgat caggttttgt tcccttttagt gagggttaat tgcgcgttg gctgaatcat
 9121 ggtcatagct gtttctgtg tgaattgtt atccgtccac aattccacac aacatcagag
 9181 ccggaagcat aaagtgtaaa gcctgggggt cctaagtgt gagctaaact acattaatg
 9241 cgttgccgct actgccgct tccaagtcgg gaaacctgt gtgccagctg cattaatga
 9301 tccgccaacg cgcggggaga ggcggtttgc gtattggggc ctcttcgct tctcgtctca
 9361 ctgactcgtc gcgctcggct gttcggctgc ggcagcgtt atcagctcac tcaaaagcgg
 9421 taatacggct atccacagaa tcaggggata acgagagaaa gaacatgtga gcaaaagggc
 9481 agcaaaaggg caggaaacgt aaaaaggcgc cgtgtctggc gtttttccat aggcctcgcg
 9541 cccctgacga gcatcacaaa aatcgacgct caagtacagc gtggcgaaac ccgacaggac
 9601 tataaagata ccagcgcttt cccctgtgaa gctccctcgt gcgctctcct gttccgaccc
 9661 tcccgcttac cggatacctt tccgcttttc tcccttcggg aagcgtggcg ctttctcata
 9721 gctcagctgt taggtatctc agttcgtgt aggtcgttcg ctccaagctg ggcgtgtgtg
 9781 acgaaccccc cgttcagccc gacgcgtgcg ccttatccgg taactatcgt cttgagttca
 9841 acccggttaag acacgactta tgcacactgg cagcagccac tggtaacagc attagcagag
 9901 caggttatgt aggcggtgct acagagttct tgaagtggt gccataacta ggcataacta
 9961 gaaggacagt atttggtatc tgcctctgc tgaagccagt taccttcgga aaaagagttg
 10021 gtatgctctg atccggcaaa caaacaccgc ctggttagcg tggttttttt gtttgcgaagc
 10081 agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt
 10141 ctgacgctca gtggaacgaa aactcagctt aagggatctt ggtcatgaga ttatcaaaaa
 10201 ggatcttcac ctgatcctt ttaattaaa aatgaagttt taaatcaatc taaagtatat
 10261 atgagtaaac ttggtctgac agttaaccaat gcttaacag tgaggcactt atctcagcga
 10321 tctgtctatt tctgtcatcc atagtgtcct gactccccgt cgtgtagata actacgatac
 10381 gggaggggctt accatctggc cccagtgtgc caatgatacc gcgagaccca cgtcaccggt
 10441 ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgaga agtggtctctg
 10501 caactttatc cgcctccatc cagtctatta attgttgccc ggaagctaga gtaagtatt
 10561 ccagagttaa tagtttgccc aacgttggtt ccatgtctac aggcacgtgt ggtcaccgt
 10621 cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat
 10681 cccccattgt gtgcacaaaa gcggttagct ccttcgggtc tccgatcgtt gtcagaagta
 10741 agttggccgc agtgttatca ctcatgttta tggcagcact gcataattct cttactgtca
 10801 tgccatccgt aagatgcttt tctgtgactg gtgagttact aaccaagtca tcttgagaat
 10861 agtgtatgcg gcgacggagt tgccttgccc cggcgtcaat acgggataat accgcgccac
 10921 atagcagaac tttaaaagt ctcatcattg gaaaacgttc ttccggggca aaactctcaa
 10981 ggaatttacc gctgttgaga tccagttcga tgaacccac tctggcacc aaactgatctt
 11041 cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaatgcgc
 11101 caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat
 11161 attattgaag catttatcag ggttattgtc tcatgagcgg atacatat ttgaatgatt
 11221 agaaaaataa acaaataggg gttccgcgca catttcccg aaaaagtgca c

SEQ ID NO:43 (pTnMOD(Chicken OVep+OVg'+ENT+proins+syn polyA))

60
 1 ctgacgcgcc ctgtagcgcg gcattaagcg cggcggtgtg ggtggttacg cgcagcgtga
 61 ccgctacact tgcacgcgcc cttagccgcg ctctttctgc tttcttctct tctttctctg
 121 ccagcttgcg cggcatcaga ttggctattg gccattgcac acgttgatct catatcataa
 181 tatgtacatt tatattggct catgtccaac attaccgcca tgttgacatt gattattgac
 241 tagttattaa tagtaataca ttaacggggtc attagttcat agccatata tggagttccg
 301 cgttacataa ctacgcgtaa atggccgcgc tggctgaccg cccaacgacc cccgccatt
 361 gcgctcaata atgacgtatg ttccatagt aacgccaata gggactttcc attgacgtca

140

ATLUB0213492.1

421 atgggtggag tatttacggt aaactgccca ctggcagta catcaagtgt atcatatgcc
 481 aagtaagccc cctattgacg tcaatgaagg taaatggccc gectggcatt atgccagta
 541 catgaacctta tgggaatttc ctacttggca gtacacttac gtattagta tegtattac
 601 catggtgatg cgggttttgc agtacatcaa tgggcgttga taggggttg actcacggg
 661 atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaatcaacg
 721 ggactttcca aaatgtogta acaactccgc cccattgacg caaatgggag gtaggcgtgt
 781 acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg cctggagacg
 841 ccatccacgc tgttttgacc tccatagaag acacgggac cgtaccagcc tccggcgccg
 901 ggaacggtag attggaacgc ggattcccg tgccaagagt gacgtaagta cggctatag
 961 actctatagg cacacccctt tggctcttat gcatgtata ctgttttttg cttggggcct
 1021 atacaccccc gcttccttat gctatagggt atggtatagc ttacgtata ggtgtgggtt
 1081 attgaccatt attgaccact cccctatttg tgacgatact ttccattact aatccataac
 1141 atgggtcttt gccacaccta tctctatttg ctatatgcca ataetctgtc cttcagagac
 1201 tgacacggac tctgtatttt tacaggatgg ggtccattt attatttaca aattcacata
 1261 tacaacaacg cgtcccccgc tgcccgcagt ttttattaaa catagcgtgg gatctccag
 1321 cgaatctcgg gtacgtgttc cggacatggg ctcttctcgc gtacggcgag agcttccaca
 1381 tccagaccct ggtcccctgc ctccagcgcc tcatggtcgc tccggagctc ctgtctcta
 1441 acagtgaggg ccagacttag gcacagcaca atgccacaca ccacagtggt gccgacaag
 1501 gccgtggcgg tagggtagt gtctgaaat gagcgtgga attgggtcgc cagcggtgac
 1561 gcagatggaa gacttaagc agcggcagaa gaagatgcag gcagctgagt tgtgtattc
 1621 tgataagagt cagaggtaac tccctgtgoc gtgctgttaa cgttggaggg cagtgtagtc
 1681 tgagcagtag tctgtctgoc cggcgccgccc accagacata atagctgaga gactaacac
 1741 ctgttccctt ccatgggtct tttctgcagt caccgtcgga ccagtgcga actgatatt
 1801 ttacacgact ctctttacca atcttgcgcc gaattacact taaaacgact caacagctta
 1861 acgttgggtt gccacgactt acttgactgt aaaaacttca ctcttaccga acttggccgt
 1921 aacctgccaa ccaaaagcag acaaaaacat aacatcaaac gaatcgacgc attgttagt
 1981 aatcgtcacc tccacaaaga gcgactcgtt gtataccgtt ggcagtctag ctttactgt
 2041 tccggcaata cgtatcccat tgtacttgtt gactgtctgt atattctgta gcaaaaaag
 2101 ctatgggtat tgcgagcttc agtcgcacta caggtcgtt ctgttactct ttatgacaaa
 2161 cggttcccgc tttcagayca atgttcaaa aaagctcatg accaatttct agccgcacct
 2221 gcgagatcgc taccgagtaa caccacacgc ctcatrgta gtgagctgg ctttaagtg
 2281 ccatggtata aatccgttga gaagctgggt tggtagctgt taagtccagt aagagaaaa
 2341 gtacaatatg cagacctagg agcggaaac tggaaccta tcagcaactt acatgatatt
 2401 tcatctagtc actcaagac tttaggctat aagaggctga ctaaaagcaa tccacttca
 2461 tgccaaatc tattgtataa atctcgtct aaaggcgaa aaaaacagc ctcgacacgg
 2521 actcattgtc accacccgct acctaaaatc tactcagct cggcaaaaga gccatgggtt
 2581 ctacgaacta acttacctgt tgaatttga acacccaaac aacttgttaa tatctattc
 2641 aagcgaatgc agattgaaga aaccttcgga gacttgaaaa gtctcgtcta cggactaggc
 2701 ctacgccata gccgaacgag cagctcagag cgttttgata tcatgtgtgt aatcgccgt
 2761 atggttcaac taacatgttg gcttgcgggc gttcatgctc agaaaacaag ttgggacaag
 2821 caactccagg ctacacagct cagaatcga aacgtactct caacagttcg cttaggcatg
 2881 gaagtcttgc ggcattctgg ctacacaata acaggggaag acttactcgt ggtcgcaacc
 2941 ctactagctc aaaatttatt cacacatggt tacgcttttg ggaatttatg aggggactgc
 3001 tctagagaga tccgggatct cgggaaaagc gttgggtgac aaaggtgcct ttatcatca
 3061 ctttaaaaat aaaaacaaat tactcagctc ctgttataag cagcaattaa ttatgatga
 3121 tgctacatc acacaaaaaa ctgatttaac aaatgggttg tctgcttag aaagtattt
 3181 tgaacattat cttgattata ttattgataa taataaaaa cttatcccta tccaagaagt
 3241 gatgcctatc attggttga atgaacttga aaaaaattg ccttgaatac attactggta
 3301 aggttaacgc cattgtcagc aaattgatcc aagagaacca acttaagct ttcctgacgg
 3361 aatgttaatt ctctgtgacc ctgagcactg atgaatcccc taatgatttt ggtaaaaatc
 3421 attaaagttaa ggtggatata catcttgcga tatgatcccc gtaatgtgag tttagctact
 3481 cattaggcac cccaggtttt acactttatg cttccggctc gtatgttgtg tggaaattgt
 3541 agcggataac aatttcacac aggaacacgc tatgaccatg attacgcaaa gcgcgcaatt
 3601 aacctcactt aaaggaaca aaagctggag ctccaccgag gtggcggcgc ccttagaact
 3661 agtggatccc cggggttga gaaaaatgcc aggtggacta tgaactcaca tccaaaggag
 3721 ctgacactga tacctgattt tcttcaact ggggaacaaa cacaatccca caaaacagct
 3781 cagagagaaa ccatcactga tggctacagc accaaggtat gcaatggcaa tccattcgac
 3841 attcatctgt gacctgagca aaatgattta tctctccatg aatggttctt tcttccctc
 3901 atgaaaagcc aatttcacac ctacacatat gcaacaaaga caaacagaga acaattaatg
 3961 tgctccttcc taatgtcaaa attgtagtgg caaagaggag aacaaaatct caagtctga
 4021 ttaggtttta gtattggat aagaggtttt gacctgtgag ctacactgga cttcatatcc
 4081 ttttggataa aaagtgtttt tataactttc aggtctccga gtctttattc atgagactgt
 4141 tgggttaggg acagaccac aatgaatgc ctggcatagg aaagggcagc agagccttag
 4201 ctgacctttt cttgggacaa gcatgttcaa acaatgtgtg acaaaactat ttgactgtct
 4261 tgcacagct gtgtgtggca gggcaatcca ttgccacctc tcccaggtaa ctttccaaact
 4321 gcaagaagat tgtgtttac tctcttga aaagctctgc agactgacat gcatctcata
 4381 ggtagagata acattttactg ggaagcacat ctatcatcat aaaaagcagg caagattttc
 4441 agacttctct agtegtgaa atagaagcaa aagacgtgat taaaaacaaa atgaacaaaa

4501 aaaaatcagt tgatacctgt ggtgtagaca tccagcaaaa aaatattatt tgcactaaca
 4561 tcttgtctta agtctctaga cttggcaagg agaattgaga ttctacacgt atatatgttt
 4621 tccaaaaagg aaggagagaa acaaaagaaa atggcactga ctaaaactta gctagtggta
 4681 taggaaagta attctgctta acagagattg cagtgatctc tatgtatgtc ctgaagaatt
 4741 abgttgtact tttttccccc atttttaaat caaacagtgc ttacagagg ttcaaatggc
 4801 ttctttactg ttgtcaatt ctattatttc aatcagaaac aatagcttct ataatgaaa
 4861 tatatttgc attgtatatt atgattgtcc ctgaacctat gaacactcct ccagctgaat
 4921 ttccaaattc ctctgtcttc tggcaggcca ttaagtatt catggaagat ctttgaggaa
 4981 cactgcaagt tcatatcata aacacatttg aaattgagta ttgttttgca ttgtatggag
 5041 ctatgttttg ctgtatcttc agaaaaaaag ttgtttataa accattcaca cccataaaaa
 5101 gatagattta aatattccag ctataggaaa gaaagtgcgt ctgctcttca ctctagtctc
 5161 agttggctcc ttccatgca tgcctcttta ttctctctat ttgtcaaga aaataatagg
 5221 tccagctctg ttctcaacta tgcctgcctc agcatggctc agatgcaagt tgtgatgaca
 5281 agaaggatca aatgaaacag acttctggtc tgttactaca accatagtaa laagcacaac
 5341 aactaataat tgcataatt gttttccatc tctaaggctc ccacattttt ctgtttctct
 5401 aaagatccca ttatctgggtt gtaactggaag ctcaatggaa catgagcaat atttccaggt
 5461 ctctctctcc atccaacagt cctgatggat tagcagaaca ggcagaaaac acattgttac
 5521 ccagaattaa aaactaatat ttgctctcca ttcaatcmaa aatggacctc ttgaaacagaa
 5581 aatcaacccc aatccattta aatgatttct atggcgtcaa aggtcaaaact tctgaaggga
 5641 acctgtgggt gggtcacaat tcaggctata tattcccccag ggcctcagcca gtggatcaac
 5701 atacagctag aaagctgtat tgccttttagc actcaagctc aaagacaaac tcagagtcca
 5761 ccatgggctc catcgcccca gcaagcatgg aattttgttt tcatgtatcc aaggagctca
 5821 aagtcaccca tgcgaatgag aacatcttct actgcccact tggcatcatg tcagctctag
 5881 ccatgggtata cctgggtgca aaagacagca ccaggacaca gataaataag gttgtctgct
 5941 ttgataaact tccaggattc ggagacagta ttgaagctca gtgtggcaca tctgtcaacg
 6001 ttcactcttc acttagagac atctccaacc aaatcaccaa accaaatgat gtttattcgt
 6061 tccagccttc cagttagact ttatctggaag agagataccc aatcctgcca gaataactgc
 6121 agtgtgtgaa ggaactgtat agaggaggct tggaaacctat caactttcaa acagctcag
 6181 atcaagccag agagctcatc aattcctggg tagaaagtcn gacaaatgga attatcagaa
 6241 atgtccttca gccaaagctc gtggattctc aaactgcaat ggttctgggt aatgcaattg
 6301 tcttcaaaag actgtgggag aaaaacttta aggatgaaga cacacaagca atgcttttca
 6361 gagtactga gcaaaaagc aaactgtgc agatgatgta ccagatttgt ttatttagag
 6421 tggcatcaat ggtctctgag aaaaagaaga tcttggagct tccatttgcc agtgggacaa
 6481 tgagcatgtt ggtgtgtgtg cctgatgaag tctcaggctc tgagcagctt gagagtataa
 6541 tcaactttga aaactgact gaatggacca gttctaattg tatggaagag aggaagatca
 6601 aagtgtactt acctgcctag aagatggagg aaaaatacaa cctcacatct gctttaatgg
 6661 ctatgggcat tactgactgt tttagctctt cagccaactc gtctggcctc tctcagcag
 6721 agagcctgaa gatattctaa gctgtccatg cagcacatgc agaaatcaat gaagcagga
 6781 gagagtggtt aggttcagca gaggctggag tggatgctgc aagcgtctct gaagaattta
 6841 gggctgacca tccattctct ttctgtatca agcacatgc aaccaacgcc gttctcttct
 6901 ttggcagatg tgtttctccg cggccagcag atgacgcacc agcagatgac gccaccagcag
 6961 atgacgcacc agcagatgac gccaccagcag atgacgcacc agcagatgac gcaacaacat
 7021 gtatcctgaa aggtctctgt ggtggatgc gctctgtgga tgacgatgac aaatttgtga
 7081 accaacacct gtgctgtcca cactgtgtg agactctctc cctagtgcgc ggggaacgag
 7141 gcttctctca cacaccacag acccgccggg aggcagagga cctgcagggt gggcaggtg
 7201 agctggcgcg gggcctctgt gccagccagcc tgcagcctt ggcctggag ggttccctgc
 7261 agaagcgtgg cattgtggaa caatgctgta ccagcatctg ctccctctac cagctggaga
 7321 actactgcaa ctaggcgccc taaaggcgga attatcgogc cgcctctaga ccaggcgctc
 7381 ggatccagat caclclggc taataaaaga tcagagctct agagatctgt gtgttggttt
 7441 ttgtgggac tgtgtgtgct tctagtgtcc agccatctgt tgtttgcccc tccccgtgc
 7501 ctctcttgac cctggaaggt gccactccca ctgtcttctc ctaataaaat gaggaatttg
 7561 catcgcttg tctgagtagg tgcattctca ttctgggggg tgggtgggg cagcacagca
 7621 agggggagga ttgggaagac aatagcagc atgctgggga tgcgtgggg tctatgggtg
 7681 cctctctctc tctctctctc tctctctctc tctctctctc tggtaacctc tctcagggg
 7741 gggcccggtta ccaattctgc cctatagtga gtctattac ggcgcctcac tggcgtctgt
 7801 ttacacagct cgtgactggg aaacccctgg cgttacccaa cttaatgccc ttgcagcaca
 7861 tcccccttcc gccagctggc gtaatagcga agaggccgcg accgatcgcc ctccccaa
 7921 gttgcgcagc ctgaatggcg aatggaaatt gtaagcgta atattttgtt aaaaatcgcg
 7981 tataaatttt gttaatcag ctctattttt aaccaatagg ccgaatcgg caaatccct
 8041 tataaatcaa aagaatagac cgagataggg ttgagtgttg ttccagtttg gaacaagagt
 8101 ccaactattaa agaagctgga ctccaacgct aaaggcgcaa aaacgcctca tcaaggcgat
 8161 gggccactac tccgggata tatgacaaga tgtgtatoca ccttaacctta atgattttta
 8221 ccaaaatcat taggggattc atcagtgtct aggtcaacg agaattaaca ttcctgcagg
 8281 aaagcttatg atgatgatg gcttaaaaac ttactcaatg gctggttatg catatgcgaa
 8341 tacatgcgaa aaacctaaaa gagcttgccg ataaaaagg ccaatttatt gctatttacc
 8401 gcggcttttt attgagcttg aaagataaat aaaaatagata ggttttattt gaagctaaat
 8461 ctcttttctc gtaaaaaatg cctcttggg ttatcaagag ggtcattata tttcgcggaa
 8521 taacatcatt tgggtgacgaa ataactaagc acttgcctcc tgttactctc cctgagcttg

3581 agggggttaac atgaaggtca tcatagtcag gataataata cagtaaaacg ctaaaaccaat
 3641 aatccaaatc cagccatccc aaattggtag tgaatgatta taataaacag caaacagtaa
 8701 tgggccaata acacgggttg cattggtaag gctcaccaat aatccctgta aagcaccttg
 8761 ctgatgactc ttgttttgga tagacatcac tccctgtaat gcaggtaaaag cgtatccccc
 5 3821 accagccaat aaaattaaaa cagggaatac taaccaacct tcagatataa acgctaaaaa
 3881 ggcaaatgca ctactatctg caataaatcc gagcagtagt gccgtttttt cggccattta
 3941 gtggctatcc ttctgcccac aaaggcttgg aatactgagt gtaaaagacc aagacccgta
 9001 atgaaaagcc aaccatcatg ctattcatca tcacgatttc tgtaatagca ccacaccgtg
 9061 ctggattggc tatcaatgag ctgaataaat aatcaacaaa tggcatcgtt aaataagtga
 10 9121 tgtataccga tcagcttttg ttcccttag tgagggttaa ttgcccgtt ggctaatca
 9181 tggcatagc tgttccctgt gtgaattgt tatccgtca caattccaca caacatacga
 9241 gccggaagca taaagtgtaa agcctgggtt gcctaatgag tgagctaact cacattaact
 9301 gcgttgctgt cactgcccgc ttccagtcg ggaacctgt cgtgcccagt gcattaatga
 9361 atcgcccaac gcggggggag aggggtttg cgtattgggc gctcttcgc ttccctgctc
 15 9421 actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaaggc
 9481 gtaatacgtt tatccacaga atcaggggat aacgcaggaa agaactgtg agcaaaagcc
 9541 cagcaaaagg ccagggaacg taaaaaggcc gcgttgctgg cgtttttcca taggctcgc
 9601 cccctgacg agcatcacaa aaatcgacgc tcaagttaga ggtggcgaaa ccgacagga
 9661 ctataaagat accagcggtt tccctcgga agctccctg tgcgtctcc tgttcgaccc
 20 9721 ctgcccgtta ccgataacct gtccgcttt ctcccttgg gaagcgtggc gctttctcat
 9781 agctcacgct gtatgtatct cagtccggt taggtcgttc gctccaagct gggctgtgtg
 9841 cagcaacccc cgttcagcc cgaccgctgc gccctatcc gtaactatcg tcttgatccc
 9901 aaccggtaa gacacgactt atcgccactg gcagcagcca ctgctaacc gatttagcga
 9961 gcgaggtatg tagcggtgct tacagagttc ttgaagtgtt ggctaacta cggctacact
 25 10021 agaagcagag tatttggtat ctgcgctcgt ctgaagccag ttaccttcgg aaaaagagt
 10081 ggtagctctt gatccggcaa acaaacccac gctggtagcg tgggtttttt tgtttgcaag
 10141 cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg
 10201 tctgacgctc agtgaacga aaactcacgt taagggtatt tggtaatgag attatcaaaa
 10261 aggatcttca cctagatcct tttaattaa aaatgaagtt ttaaatcaat cttaagtata
 30 10321 tatgagtaaa cttggtctga cagttaacca tgcctaatca gtgagcgacc tatctcagcg
 10381 atctgtctat ttggttctac catagttgcc tgaactcccc tctgttagat aactacgata
 10441 cggggggggt taccatctgg cccnagtgct gcaatgatac cgcgagacc accgtcaccc
 10501 gctccagatt taccagcaat aaaccagcca gccggaaggc cgcagcgagc aagtgttctc
 10561 gcaactttat ccgctcccat ccagttctatt aattgttgcc gggagcag agtaagtagt
 35 10621 tgcacagtta atagtttgcc caacytltgt gccattgcta caggcatcgt ggtgtcacgc
 10681 tctgtgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga
 10741 tccccatgt tgtgcaaaaa agcggttagc tccctcggtc ctccgatcgt tgcagaagt
 10801 aagttggcgc cagtgttctc actcatggtt atggcagcac tgcataatc tcttactgtc
 10861 atgcatccgc taagatgctt ttctgtgact ggtgagtagt caaccaagtc attctgagaa
 40 10921 tagtgtatgc ggccagcgag ttgctcttgc ccggcgtaaa tacgggataa taccgcgcca
 10981 catagcagaa ctttaaaagt gctcatcatt ggaatacgtt cttcggggcg aaaaactetca
 11041 aggatcttac cgtgtttgag atccagttcg atgtaaccca ctgctgcacc caactgatct
 11101 tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacagggaag gcaaatgtcc
 11161 gcaaaaaagg gaataagggc gacacggaaa tgttgaaatc tcatactctt cctttttcaa
 45 11221 tattattgaa gcatttatca ggtttattgt ctcatgagc gatcacatatt tgaatgtatt
 11281 tagaaaaata aacaaatagc ggttccgcgc acatttcccc gaaaagtgc ac