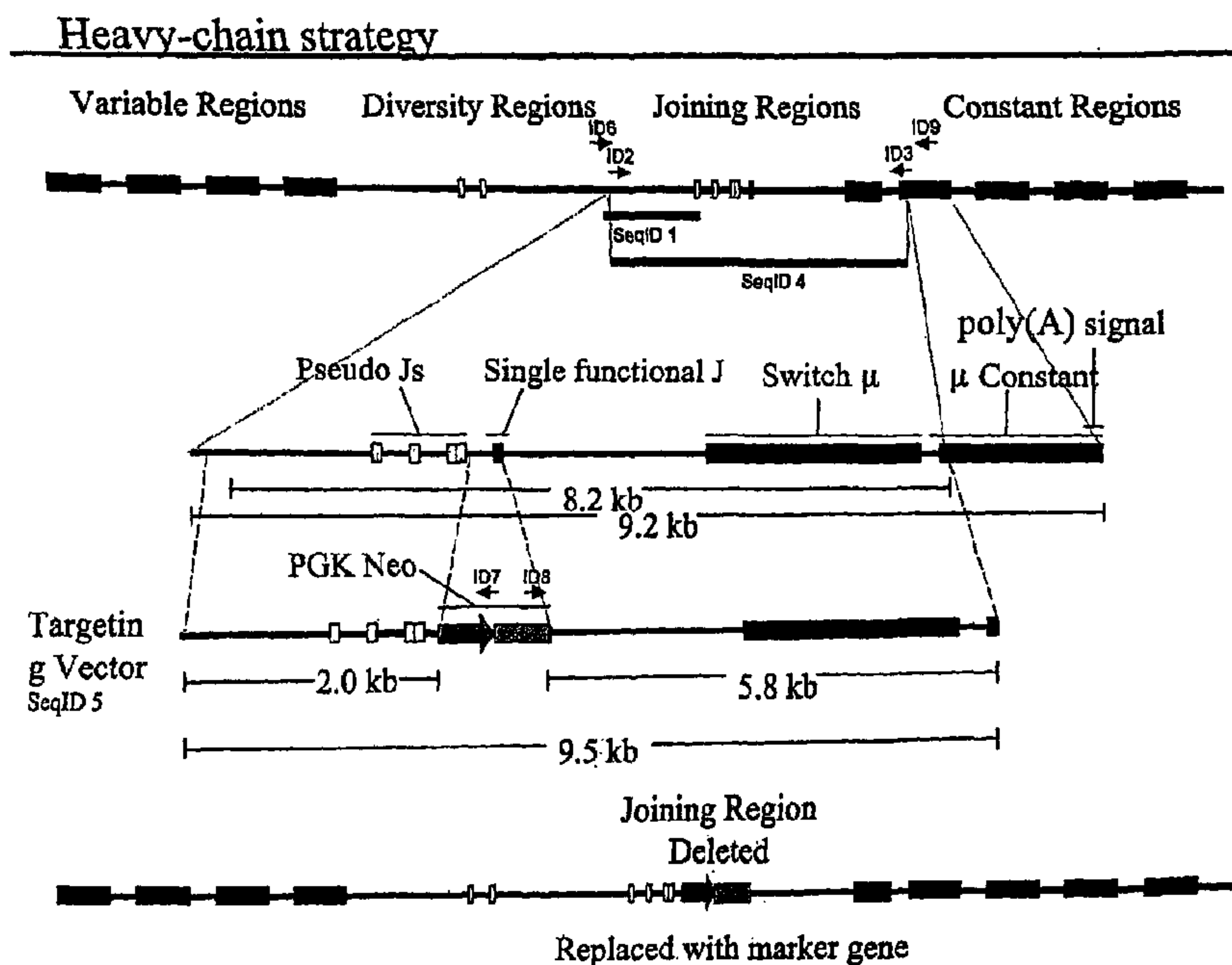




(86) Date de dépôt PCT/PCT Filing Date: 2005/10/24
(87) Date publication PCT/PCT Publication Date: 2006/05/04
(45) Date de délivrance/Issue Date: 2018/12/18
(85) Entrée phase nationale/National Entry: 2007/04/20
(86) N° demande PCT/PCT Application No.: US 2005/038582
(87) N° publication PCT/PCT Publication No.: 2006/047603
(30) Priorité/Priority: 2004/10/22 (US60/621,433)

(51) Cl.Int./Int.Cl. *C12N 15/13* (2006.01),
A01K 67/027 (2006.01), *C07K 16/00* (2006.01),
C12N 15/85 (2006.01), *C12N 5/10* (2006.01),
C12P 21/00 (2006.01)
(72) Inventeurs/Inventors:
WELLS, KEVIN, US;
AYARES, DAVID, US
(73) Propriétaire/Owner:
REVIVICOR, INC., US
(74) Agent: MBM INTELLECTUAL PROPERTY LAW LLP

(54) Titre : SEQUENCES GENOMIQUES PORCINES KAPPA ET LAMBDA A CHAINE LEGERE
(54) Title: PORCINE GENOMIC KAPPA AND LAMBDA LIGHT CHAIN SEQUENCES



(57) **Abrégé/Abstract:**

The present invention provides ungulate animals, tissue and organs as well as cells and cell lines derived from such animals, tissue and organs, which lack expression of functional endogenous immunoglobulin loci. The present invention also provides ungulate animals, tissue and organs as well as cells and cell lines derived from such animals, tissue and organs, which express xenogenous, such as human, immunoglobulin loci. The present invention further provides ungulate, such as porcine genomic DNA sequence of porcine heavy and light chain immunoglobulins. Such animals, tissues, organs and cells can be used in research and medical therapy. In addition, methods are provided to prepare such animals, organs, tissues, and cells.

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 May 2006 (04.05.2006)

PCT

(10) International Publication Number
WO 2006/047603 A2

(51) International Patent Classification:
A01K 67/027 (2006.01)

(21) International Application Number:
PCT/US2005/038582

(22) International Filing Date: 24 October 2005 (24.10.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/621,433 22 October 2004 (22.10.2004) US

(71) Applicant (for all designated States except US): **RE-VIVICOR, INC.** [US/US]; 1700 Kraft Drive, Suite 2400, Blacksburg, VA 24060 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **WELLS, Kevin** [US/US]; 635 Cameo Court, Christiansburg, VA 24073 (US). **AYARES, David** [US/US]; 300 Seminole Drive, Blacksburg, VA 24060 (US).

(74) Agent: **KNOWLES, Sherry, M.**; King & Spalding LLP, 191 Peachtree Steet, 45th Floor, Atlanta, GA 30303-1763 (US).

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

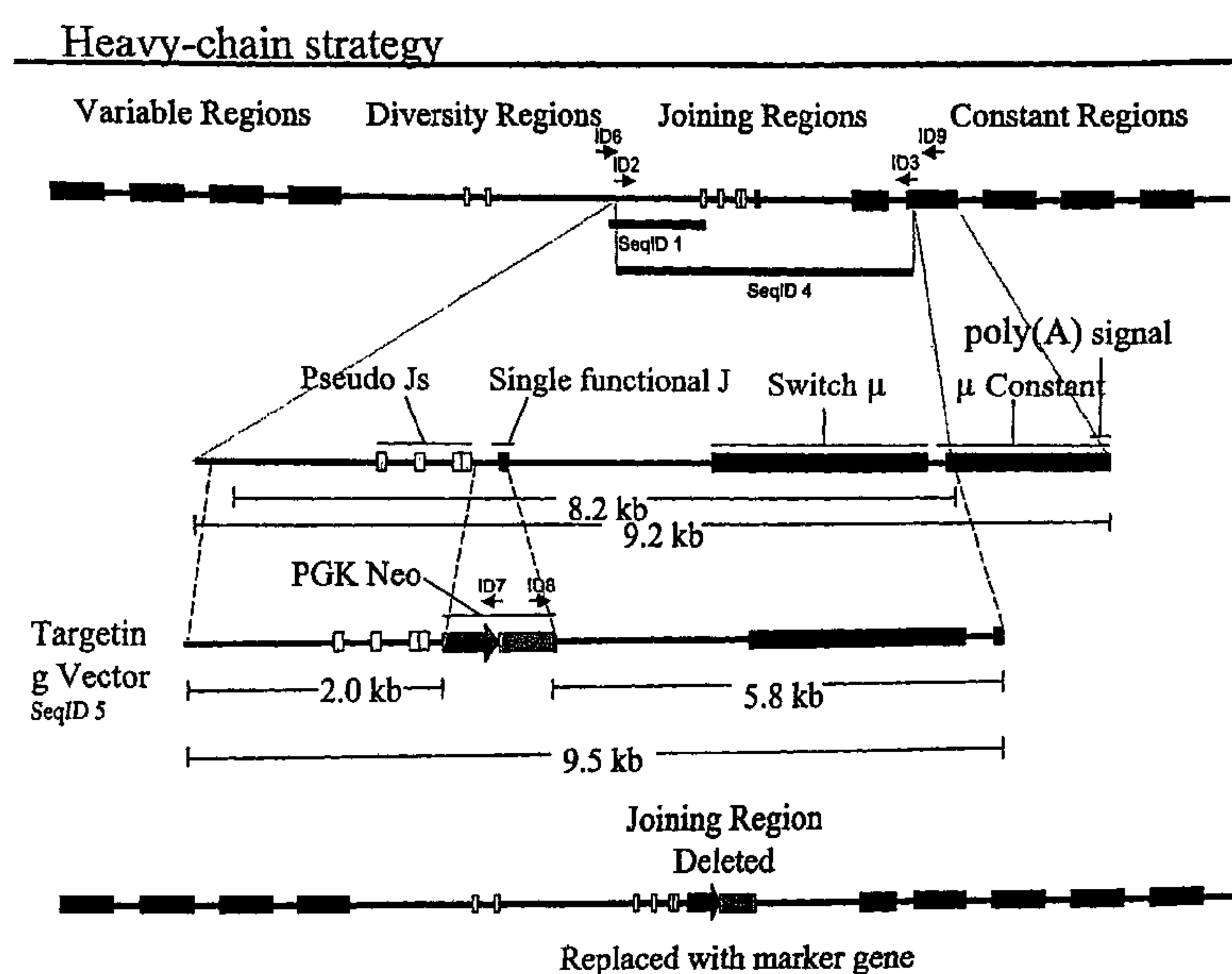
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT,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

Published:

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

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

(54) Title: UNGULATES WITH GENETICALLY MODIFIED IMMUNE SYSTEMS



(57) **Abstract:** The present invention provides ungulate animals, tissue and organs as well as cells and cell lines derived from such animals, tissue and organs, which lack expression of functional endogenous immunoglobulin loci. The present invention also provides ungulate animals, tissue and organs as well as cells and cell lines derived from such animals, tissue and organs, which express xenogenous, such as human, immunoglobulin loci. The present invention further provides ungulate, such as porcine genomic DNA sequence of porcine heavy and light chain immunoglobulins. Such animals, tissues, organs and cells can be used in research and medical therapy. In addition, methods are provided to prepare such animals, organs, tissues, and cells.

WO 2006/047603 A2

DEMANDES OU BREVETS VOLUMINEUX

**LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS
COMPREND PLUS D'UN TOME.**

CECI EST LE TOME __1__ DE __2__

NOTE: Pour les tomes additionels, veuillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

**THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE
THAN ONE VOLUME.**

THIS IS VOLUME __1__ OF __2__

NOTE: For additional volumes please contact the Canadian Patent Office.

PORCINE GENOMIC KAPPA AND LAMBDA LIGHT CHAIN SEQUENCES

FIELD OF THE INVENTION

The present invention provides ungulate animals, tissue and organs as well as cells and cell lines derived from such animals, tissue and organs, which lack expression of functional endogenous immunoglobulin loci. The present invention also provides ungulate animals, tissue and organs as well as cells and cell lines derived from such animals, tissue and organs, which express xenogenous, such as human, immunoglobulin loci. The present invention further provides ungulate, such as porcine genomic DNA sequence of porcine heavy and light chain immunoglobulins. Such animals, tissues, organs and cells can be used in research and medical therapy. In addition, methods are provided to prepare such animals, organs, tissues, and cells.

BACKGROUND OF THE INVENTION

An antigen is an agent or substance that can be recognized by the body as 'foreign'. Often it is only one relatively small chemical group of a larger foreign substance which acts as the antigen, for example a component of the cell wall of a bacterium. Most antigens are proteins, though carbohydrates can act as weak antigens. Bacteria, viruses and other microorganisms commonly contain many antigens, as do pollens, dust mites, molds, foods, and other substances. The body reacts to antigens by making antibodies. Antibodies (also called immunoglobulins (Igs)) are proteins that are manufactured by cells of the immune system that bind to an antigen or foreign protein. Antibodies circulate in the serum of blood to detect foreign antigens and constitute the gamma globulin part of the blood proteins. These antibodies interact chemically with the antigen in a highly specific manner, like two pieces of a jigsaw puzzle, forming an antigen/antibody complex, or immune complex. This binding neutralises or brings about the destruction of the antigen.

When a vertebrate first encounters an antigen, it exhibits a primary humoral immune response. If the animal encounters the same antigen after a few days the immune response is more

rapid and has a greater magnitude. The initial encounter causes specific immune cell (B-cell) clones to proliferate and differentiate. The progeny lymphocytes include not only effector cells (antibody producing cells) but also clones of memory cells, which retain the capacity to produce both effector and memory cells upon subsequent stimulation by the original antigen. The effector cells live for only a few days. The memory cells live for a lifetime and can be reactivated by a second stimulation with the same antigen. Thus, when an antigen is encountered a second time, its memory cells quickly produce effector cells which rapidly produce massive quantities of antibodies.

By exploiting the unique ability of antibodies to interact with antigens in a highly specific manner, antibodies have been developed as molecules that can be manufactured and used for both diagnostic and therapeutic applications. Because of their unique ability to bind to antigenic epitopes, polyclonal and monoclonal antibodies can be used to identify molecules carrying that epitope or can be directed, by themselves or in conjunction with another moiety, to a specific site for diagnosis or therapy. Polyclonal and monoclonal antibodies can be generated against practically any pathogen or biological target. The term polyclonal antibody refers to immune sera that usually contain pathogen-specific antibodies of various isotypes and specificities. In contrast, monoclonal antibodies consist of a single immunoglobulin type, representing one isotype with one specificity.

In 1890, Shibasaburo Kitazato and Emil Behring conducted the fundamental experiment that demonstrated immunity can be transmitted from one animal to another by transferring the serum from an immune animal to a non-immune animal. This landmark experiment laid the foundation for the introduction of passive immunization into clinical practice. However, wide scale serum therapy was largely abandoned in the 1940s because of the toxicity associated with the administration of heterologous sera and the introduction of effective antimicrobial chemotherapy. Currently, such polyclonal antibody therapy is indicated to treat infectious diseases in relatively few situations, such as replacement therapy in immunoglobulin-deficient patients, post-exposure prophylaxis against several viruses (e.g., rabies, measles, hepatitis A and B, varicella), and toxin neutralization (diphtheria, tetanus, and botulism). Despite the limited use of serum therapy, in the United States, more than 16 metric tons of human antibody are required each year for intravenous antibody therapy. Comparable levels of use exist in the economies of most highly industrialized countries, and the demand can be expected to grow rapidly in

developing countries. Currently, human antibody for passive immunization is obtained from the pooled serum of donors. Thus, there is an inherent limitation in the amount of human antibody available for therapeutic and prophylactic therapies.

The use of antibodies for passive immunization against biological warfare agents represents a very promising defense strategy. The final line of defense against such agents is the immune system of the exposed individual. Current defense strategies against biological weapons include such measures as enhanced epidemiologic surveillance, vaccination, and use of antimicrobial agents. Since the potential threat of biological warfare and bioterrorism is inversely proportional to the number of immune persons in the targeted population, biological agents are potential weapons only against populations with a substantial proportion of susceptible persons.

Vaccination can reduce the susceptibility of a population against specific threats, provided that a safe vaccine exists that can induce a protective response. Unfortunately, inducing a protective response by vaccination may take longer than the time between exposure and onset of disease. Moreover, many vaccines require multiple doses to achieve a protective immune response, which would limit their usefulness in an emergency to provide rapid prophylaxis after an attack. In addition, not all vaccine recipients mount a protective response, even after receiving the recommended immunization schedule.

Drugs can provide protection when administered after exposure to certain agents, but none are available against many potential agents of biological warfare. Currently, no small-molecule drugs are available that prevent disease following exposure to preformed toxins. The only currently available intervention that could provide a state of immediate immunity is passive immunization with protective antibody (Arturo Casadevall "Passive Antibody Administration (Immediate Immunity) as a Specific Defense Against Biological Weapons" from Emerging Infectious Diseases, Posted 09/12/2002).

In addition to providing protective immunity, modern antibody-based therapies constitute a potentially useful option against newly emergent pathogenic bacteria, fungi, virus and parasites (A. Casadevall and M. D. Scharff, Clinical Infectious Diseases 1995; 150). Therapies of patients with malignancies and cancer (C. Botti et al, Leukemia 1997; Suppl 2:S55-59; B. Bodey, S.E. Siegel, and H.E. Kaiser, Anticancer Res 1996; 16(2):661), therapy of steroid resistant rejection of transplanted organs as well as autoimmune diseases can also be achieved through the use of

monoclonal or polyclonal antibody preparations (N. Bonnefoy-Berard and J.P. Revillard, *J Heart Lung Transplant* 1996; 15(5):435-442; C. Colby, et al *Ann Pharmacother* 1996; 30(10):1164-1174; M.J. Dugan, et al, *Ann Hematol* 1997; 75(1-2):41 2; W. Cendrowski, *Boll Ist Sieroter Milan* 1997; 58(4):339-343; L.K. Kastrukoff, et al *Can J Neurol Sci* 1978; 5(2):175178; J.E. Walker et al *J Neurol Sci* 1976; 29(2-4):303309).

Recent advances in the technology of antibody production provide the means to generate human antibody reagents, while avoiding the toxicities associated with human serum therapy. The advantages of antibody-based therapies include versatility, low toxicity, pathogen specificity, enhancement of immune function, and favorable pharmacokinetics.

The clinical use of monoclonal antibody therapeutics has just recently emerged. Monoclonal antibodies have now been approved as therapies in transplantation, cancer, infectious disease, cardiovascular disease and inflammation. In many more monoclonal antibodies are in late stage clinical trials to treat a broad range of disease indications. As a result, monoclonal antibodies represent one of the largest classes of drugs currently in development.

Despite the recent popularity of monoclonal antibodies as therapeutics, there are some obstacles for their use. For example, many therapeutic applications for monoclonal antibodies require repeated administrations, especially for chronic diseases such as autoimmunity or cancer. Because mice are convenient for immunization and recognize most human antigens as foreign, monoclonal antibodies against human targets with therapeutic potential have typically been of murine origin. However, murine monoclonal antibodies have inherent disadvantages as human therapeutics. For example, they require more frequent dosing to maintain a therapeutic level of monoclonal antibodies because of a shorter circulating half-life in humans than human antibodies. More critically, repeated administration of murine immunoglobulin creates the likelihood that the human immune system will recognize the mouse protein as foreign, generating a human anti-mouse antibody response, which can cause a severe allergic reaction. This possibility of reduced efficacy and safety has lead to the development of a number of technologies for reducing the immunogenicity of murine monoclonal antibodies.

Polyclonal antibodies are highly potent against multiple antigenic targets. They have the unique ability to target and kill a plurality of "evolving targets" linked with complex diseases. Also, of all drug classes, polyclonals have the highest probability of retaining activity in the event of antigen mutation. In addition, while monoclonals have limited therapeutic activity

against infectious agents, polyclonals can both neutralize toxins and direct immune responses to eliminate pathogens, as well as biological warfare agents.

The development of polyclonal and monoclonal antibody production platforms to meet future demand for production capacity represents a promising area that is currently the subject of much research. One especially promising strategy is the introduction of human immunoglobulin genes into mice or large domestic animals. An extension of this technology would include inactivation of their endogenous immunoglobulin genes. Large animals, such as sheep, pigs and cattle, are all currently used in the production of plasma derived products, such as hyperimmune serum and clotting factors, for human use. This would support the use of human polyclonal antibodies from such species on the grounds of safety and ethics. Each of these species naturally produces considerable quantities of antibody in both serum and milk.

Arrangement of Genes Encoding Immunoglobulins

Antibody molecules are assembled from combinations of variable gene elements, and the possibilities resulting from combining the many variable gene elements in the germline enable the host to synthesize antibodies to an extraordinarily large number of antigens. Each antibody molecule consists of two classes of polypeptide chains, light (L) chains (that can be either kappa (κ) L-chain or lambda (λ) L-chain) and heavy (H) chains. The heavy and light chains join together to define a binding region for the epitope. A single antibody molecule has two identical copies of the L chain and two of the H chain. Each of the chains is comprised of a variable region (V) and a constant region (C). The variable region constitutes the antigen-binding site of the molecule. To achieve diverse antigen recognition, the DNA that encodes the variable region undergoes gene rearrangement. The constant region amino acid sequence is specific for a particular isotype of the antibody, as well as the host which produces the antibody, and thus does not undergo rearrangement.

The mechanism of DNA rearrangement is similar for the variable region of both the heavy- and light-chain loci, although only one joining event is needed to generate a light-chain gene whereas two are needed to generate a complete heavy-chain gene. The most common mode of rearrangement involves the looping-out and deletion of the DNA between two gene segments. This occurs when the coding sequences of the two gene segments are in the same orientation in the DNA. A second mode of recombination can occur between two gene segments that have

opposite transcriptional orientations. This mode of recombination is less common, although such rearrangements can account for up to half of all V_{κ} to J_{κ} joins; the transcriptional orientation of half of the human V_{κ} gene segments is opposite to that of the J_{κ} gene segments.

The DNA sequence encoding a complete V region is generated by the somatic recombination of separate gene segments. The V region, or V domain, of an immunoglobulin heavy or light chain is encoded by more than one gene segment. For the light chain, the V domain is encoded by two separate DNA segments. The first segment encodes the first 95-101 amino acids of the light chain and is termed a V gene segment because it encodes most of the V domain. The second segment encodes the remainder of the V domain (up to 13 amino acids) and is termed a joining or J gene segment. The joining of a V and a J gene segment creates a continuous exon that encodes the whole of the light-chain V region. To make a complete immunoglobulin light-chain messenger RNA, the V-region exon is joined to the C-region sequence by RNA splicing after transcription.

A heavy-chain V region is encoded in three gene segments. In addition to the V and J gene segments (denoted V_H and J_H to distinguish them from the light-chain V_L and J_L), there is a third gene segment called the diversity or D_H gene segment, which lies between the V_H and J_H gene segments. The process of recombination that generates a complete heavy-chain V region occurs in two separate stages. In the first, a D_H gene segment is joined to a J_H gene segment; then a V_H gene segment rearranges to DJ_H to make a complete V_H -region exon. As with the light-chain genes, RNA splicing joins the assembled V-region sequence to the neighboring C-region gene.

Diversification of the antibody repertoire occurs in two stages: primarily by rearrangement ("V(D)J recombination") of Ig V, D and J gene segments in precursor B cells resident in the bone marrow, and then by somatic mutation and class switch recombination of these rearranged Ig genes when mature B cells are activated. Immunoglobulin somatic mutation and class switching are central to the maturation of the immune response and the generation of a "memory" response.

The genomic loci of antibodies are very large and they are located on different chromosomes. The immunoglobulin gene segments are organized into three clusters or genetic loci: the κ , λ , and heavy-chain loci. Each is organized slightly differently. For example, in humans, immunoglobulin genes are organized as follows. The λ light-chain locus is located on

chromosome 22 and a cluster of V_{λ} gene segments is followed by four sets of J_{λ} gene segments each linked to a single C_{λ} gene. The κ light-chain locus is on chromosome 2 and the cluster of V_{κ} gene segments is followed by a cluster of J_{κ} gene segments, and then by a single C_{κ} gene. The organization of the heavy-chain locus, on chromosome 14, resembles that of the κ locus, with separate clusters of V_H , D_H , and J_H gene segments and of C_H genes. The heavy-chain locus differs in one important way: instead of a single C-region, it contains a series of C regions arrayed one after the other, each of which corresponds to a different isotype. There are five immunoglobulin heavy chain isotypes: IgM, IgG, IgA, IgE and IgD. Generally, a cell expresses only one at a time, beginning with IgM. The expression of other isotypes, such as IgG, can occur through isotype switching.

The joining of various V, D and J genes is an entirely random event that results in approximately 50,000 different possible combinations for VDJ(H) and approximately 1,000 for VJ(L). Subsequent random pairing of H and L chains brings the total number of antibody specificities to about 10^7 possibilities. Diversity is further increased by the imprecise joining of different genetic segments. Rearrangements occur on both DNA strands, but only one strand is transcribed (due to allelic exclusion). Only one rearrangement occurs in the life of a B cell because of irreversible deletions in DNA. Consequently, each mature B cell maintains one immunologic specificity and is maintained in the progeny or clone. This constitutes the molecular basis of the clonal selection; i.e., each antigenic determinant triggers the response of the pre-existing clone of B lymphocytes bearing the specific receptor molecule. The primary repertoire of B cells, which is established by V(D)J recombination, is primarily controlled by two closely linked genes, recombination activating gene (RAG)-1 and RAG-2.

Over the last decade, considerable diversity among vertebrates in both Ig gene diversity and antibody repertoire development has been revealed. Rodents and humans have five heavy chain classes, IgM, IgD, IgG, IgE and IgA, and each have four subclasses of IgG and one or two subclasses of IgA, while rabbits have a single IgG heavy chain gene but 13 genes for different IgA subclasses (Burnett, R.C et al. *EMBO J* 8:4047; Honjo, In Honjo, T, Alt. F.W. T.H. eds, *Immunoglobulin Genes* p. 123 Academic Press, New York). Swine have at least six IgG subclasses (Kacskovics, I et al. 1994 *J Immunol* 153:3565), but no IgD (Butler et al. 1996 *Inter. Immunol* 8:1897-1904). A gene encoding IgD has only been described in rodents and primates. Diversity in the mechanism of repertoire development is exemplified by contrasting the pattern

seen in rodents and primates with that reported for chickens, rabbits, swine and the domesticated Bovidae. Whereas the former group have a large number of V_H genes belonging to seven to 10 families (Rathbun, G. In Hongo, T. Alt. F.W. and Rabbitts, T.H., eds, *Immunoglobulin Genes*, p. 63, Academic press New York), the V_H genes of each member of the latter group belong to a single V_H gene family (Sun, J. et al. 1994 *J. Immunol.* 155:56118; Dufour, V et al.1996, *J Immunol.* 156:2163). With the exception of the rabbit, this family is composed of less than 25 genes. Whereas rodents and primates can utilize four to six J_H segments, only a single J_H is available for repertoire development in the chicken (Reynaud et al. 1989 *Adv. Immunol.* 57:353). Similarly, Butler et al. (1996 *Inter. Immunol* 8:1897-1904) hypothesized that swine may resemble the chicken in having only a single J_H gene. These species generally have fewer V, D and J genes; in the pig and cow a single VH gene family exists, consisting of less than 20 gene segments (Butler et al, *Advances in Swine in Biomedical Research*, eds: Tumbleson and Schook, 1996; Sinclair et al, *J. Immunol.* 159: 3883, 1997). Together with lower numbers of J and D gene segments, this results in significantly less diversity being generated by gene rearrangement. However, there does appear to be greater numbers of light chain genes in these species. Similar to humans and mice, these species express a single κ light chain but multiple λ light chain genes. However, these do not seem to affect the restricted diversity that is achieved by rearrangement.

Since combinatorial joining of more than 100 V_H , 20-30 D_H and four to six J_H gene segments is a major mechanism of generating the antibody repertoire in humans, species with fewer V_H , D_H or J_H segments must either generate a smaller repertoire or use alternative mechanisms for repertoire development. Ruminants, pigs, rabbits and chickens, utilize several mechanisms to generate antibody diversity. In these species there appears to be an important secondary repertoire development, which occurs in highly specialized lymphoid tissue such as ileal Peyer's patches (Binns and Licence, *Adv. Exp. Med. Biol.* 186: 661, 1985). Secondary repertoire development occurs in these species by a process of somatic mutation which is a random and not fully understood process. The mechanism for this repertoire diversification appears to be templated mutation, or gene conversion (Sun et al, *J. Immunol.* 153: 5618, 1994) and somatic hypermutation.

Gene conversion is important for antibody diversification in some higher vertebrates, such as chickens, rabbits and cows. In mice, however, conversion events appear to be infrequent

among endogenous antibody genes. Gene conversion is a distinct diversifying mechanism characterized by transfers of homologous sequences from a donor antibody V gene segment to an acceptor V gene segment. If donor and acceptor segments have numerous sequence differences then gene conversion can introduce a set of sequence changes into a V region by a single event. Depending on the species, gene conversion events can occur before and/or after antigen exposure during B cell differentiation (Tsai et al. International Immunology, Vol. 14, No. 1, 55-64, January 2002).

Somatic hypermutation achieves diversification of antibody genes in all higher vertebrate species. It is typified by the introduction of single point mutations into antibody V(D)J segments. Generally, hypermutation appears to be activated in B cells by antigenic stimulation.

Production of Animals with Humanized Immune Systems

In order to reduce the immunogenicity of antibodies generated in mice for human therapeutics, various attempts have been made to replace murine protein sequences with human protein sequences in a process now known as humanization. Transgenic mice have been constructed which have had their own immunoglobulin genes functionally replaced with human immunoglobulin genes so that they produce human antibodies upon immunization. Elimination of mouse antibody production was achieved by inactivation of mouse Ig genes in embryonic stem (ES) cells by using gene-targeting technology to delete crucial *cis*-acting sequences involved in the process of mouse Ig gene rearrangement and expression. B cell development in these mutant mice could be restored by the introduction of megabase-sized YACs containing a human germline-configuration H- and κ L-chain minilocus transgene. The expression of fully human antibody in these transgenic mice was predominant, at a level of several 100 $\mu\text{g/l}$ of blood. This level of expression is several hundred-fold higher than that detected in wild-type mice expressing the human Ig gene, indicating the importance of inactivating the endogenous mouse Ig genes in order to enhance human antibody production by mice.

The first humanization attempts utilized molecular biology techniques to construct recombinant antibodies. For example, the complementarity determining regions (CDR) from a mouse antibody specific for a hapten were grafted onto a human antibody framework, effecting a CDR replacement. The new antibody retained the binding specificity conveyed by the CDR sequences (P. T. Jones et al. Nature 321: 522-525 (1986)). The next level of humanization

involved combining an entire mouse VH region with a human constant region such as γ_1 (S. L. Morrison et al., Proc. Natl. Acad. Sci., 81, pp. 6851-6855 (1984)). However, these chimeric antibodies, which still contain greater than 30% xenogeneic sequences, are sometimes only marginally less immunogenic than totally xenogeneic antibodies (M. Bruggemann et al., J. Exp. Med., 170, pp. 2153-2157 (1989)).

Subsequently, attempts were carried out to introduce human immunoglobulin genes into the mouse, thus creating transgenic mice capable of responding to antigens with antibodies having human sequences (Bruggemann et al. Proc. Nat'l. Acad. Sci. USA 86:6709-6713 (1989)). Due to the large size of human immunoglobulin genomic loci, these attempts were thought to be limited by the amount of DNA, which could be stably maintained by available cloning vehicles. As a result, many investigators concentrated on producing mini-loci containing limited numbers of V region genes and having altered spatial distances between genes as compared to the natural or germline configuration (See, for example, U.S. Pat. No. 5,569,825). These studies indicated that producing human sequence antibodies in mice was possible, but serious obstacles remained regarding obtaining sufficient diversity of binding specificities and effector functions (isotypes) from these transgenic animals to meet the growing demand for antibody therapeutics.

In order to provide additional diversity, work has been conducted to add large germline fragments of the human Ig locus into transgenic mammals. For example, a majority of the human V, D, and J region genes arranged with the same spacing found in the unarranged germline of the human genome and the human C_μ and C_δ constant regions was introduced into mice using yeast artificial chromosome (YAC) cloning vectors (See, for example, WO 94/02602). A 22 kb DNA fragment comprising sequences encoding a human γ_2 constant region and the upstream sequences required for class-switch recombination was latter appended to the foregoing transgene. In addition, a portion of a human kappa locus comprising V_κ , J_κ and C_κ region genes, also arranged with substantially the same spacing found in the unarranged germline of the human genome, was introduced into mice using YACS. Gene targeting was used to inactivate the murine IgH & kappa light chain immunoglobulin gene loci and such knockout strains were bred with the above transgenic strains to generate a line of mice having the human V, D, J, C_μ , C_δ . and C_{γ_2} constant regions as well as the human V_κ , J_κ and C_κ region genes all on an inactivated murine immunoglobulin background (See, for example, PCT patent application WO 94/02602 to Kucherlapati et al.; see also Mendez et al., Nature Genetics 15:146-156 (1997)).

Yeast artificial chromosomes as cloning vectors in combination with gene targeting of endogenous loci and breeding of transgenic mouse strains provided one solution to the problem of antibody diversity. Several advantages were obtained by this approach. One advantage was that YACs can be used to transfer hundreds of kilobases of DNA into a host cell. Therefore, use of YAC cloning vehicles allows inclusion of substantial portions of the entire human Ig heavy and light chain regions into a transgenic mouse thus approaching the level of potential diversity available in the human. Another advantage of this approach is that the large number of V genes has been shown to restore full B cell development in mice deficient in murine immunoglobulin production. This ensures that these reconstituted mice are provided with the requisite cells for mounting a robust human antibody response to any given immunogen. (See, for example, WO 94/02602.; L. Green and A. Jakobovits, J. Exp. Med. 188:483-495 (1998)). A further advantage is that sequences can be deleted or inserted onto the YAC by utilizing high frequency homologous recombination in yeast. This provides for facile engineering of the YAC transgenes.

In addition, Green et al. Nature Genetics 7:13-21 (1994) describe the generation of YACs containing 245 kb and 190 kb-sized germline configuration fragments of the human heavy chain locus and kappa light chain locus, respectively, which contained core variable and constant region sequences. The work of Green et al. was recently extended to the introduction of greater than approximately 80% of the human antibody repertoire through introduction of megabase sized, germline configuration YAC fragments of the human heavy chain loci and kappa light chain loci, respectively, to produce XenoMouseTM mice. See, for example, Mendez et al. Nature Genetics 15:146-156 (1997), Green and Jakobovits J. Exp. Med. 188:483-495 (1998), European Patent No. EP 0 463 151 B1, PCT Publication Nos. WO 94/02602, WO 96/34096 and WO 98/24893.

Several strategies exist for the generation of mammals that produce human antibodies. In particular, there is the "minilocus" approach that is typified by work of GenPharm International, Inc. and the Medical Research Council, YAC introduction of large and substantially germline fragments of the Ig loci that is typified by work of Abgenix, Inc. (formerly Cell Genesys). The introduction of entire or substantially entire loci through the use microcell fusion as typified by work of Kirin Beer Kabushiki Kaisha.

In the minilocus approach, an exogenous Ig locus is mimicked through the inclusion of pieces (individual genes) from the Ig locus. Thus, one or more V_H genes, one or more D_H genes,

one or more J_H genes, a mu constant region, and a second constant region (such as a gamma constant region) are formed into a construct for insertion into an animal. See, for example, U.S. Patent Nos. 5,545,807, 5,545,806, 5,625,825, 5,625,126, 5,633,425, 5,661,016, 5,770,429, 5,789,650, 5,814,318, 5,591,669, 5,612,205, 5,721,367, 5,789,215, 5,643,763; European Patent No. 0 546 073; PCT Publication Nos. WO 92/03918, WO 92/22645, WO 92/22647, WO 92/22670, WO 93/12227, WO 94/00569, WO 94/25585, WO 96/14436, WO 97/13852, and WO 98/24884; Taylor et al. Nucleic Acids Research 20:6287-6295 (1992), Chen et al. International Immunology 5:647-656 (1993), Tuailon et al. J. Immunol. 154:6453-6465 (1995), Choi et al. Nature Genetics 4:117-123 (1993), Lonberg et al. Nature 368:856-859 (1994), Taylor et al. International Immunology 6:579-591 (1994), Tuailon et al. J. Immunol. 154:6453-6465 (1995), and Fishwild et al. Nature Biotech. 14:845-851 (1996).

In the microcell fusion approach, portions or whole human chromosomes can be introduced into mice (see, for example, European Patent Application No. EP 0 843 961 A1). Mice generated using this approach and containing the human Ig heavy chain locus will generally possess more than one, and potentially all, of the human constant region genes. Such mice will produce, therefore, antibodies that bind to particular antigens having a number of different constant regions.

While mice remain the most developed animal for the expression of human immunoglobulins in humans, recent technological advances have allowed for progress to begin in applying these techniques to other animals, such as cows. The general approach in mice has been to genetically modify embryonic stem cells of mice to knock-out murine immunoglobulins and then insert YACs containing human immunoglobulins into the ES cells. However, ES cells are not available for cows or other large animals such as sheep and pigs. Thus, several fundamental developments had to occur before even the possibility existed to generate large animals with immunoglobulin genes knocked-out and that express human antibody. The alternative to ES cell manipulation to create genetically modified animals is cloning using somatic cells that have been genetically modified. Cloning using genetically modified somatic cells for nuclear transfer has only recently been accomplished.

Since the announcement of Dolly's (a cloned sheep) birth from an adult somatic cell in 1997 (Wilmut, I., et al (1997) Nature 385: 810-813), ungulates, including cattle (Cibelli, J et al 1998 Science 280: 1266-1258; Kubota, C. et al.2000 Proc. Nat'l. Acad. Sci 97: 990-995), goats

(Baguisi, A. et al., (1999) Nat. Biotechnology 17: 456-461), and pigs (Polejaeva, I.A., et al. 2000 Nature 407: 86-90; Betthauser, J. et al. 2000 Nat. Biotechnology 18: 1055-1059) have been cloned.

The next technological advance was the development of the technique to genetically modify the cells prior to nuclear transfer to produce genetically modified animals. PCT publication No. WO 00/51424 to PPL Therapeutics describes the targeted genetic modification of somatic cells for nuclear transfer.

Subsequent to these fundamental developments, single and double allele knockouts of genes and the birth of live animals with these modifications have been reported. Between 2002 and 2004, three independent groups, Immerge Biotherapeutics, Inc. in collaboration with the University of Missouri (Lai et al. (Science (2002) 295: 1089-1092) & Kolber-Simonds et al. (PNAS. (2004) 101(19):7335-40)), Alexion Pharmaceuticals (Ramsoondar et al. (Biol Reprod (2003)69: 437-445) and Revivicor, Inc. (Dai et al. (Nature Biotechnology (2002) 20: 251-255) & Phelps et al. (Science (2003) Jan 17;299(5605):411-4)) produced pigs that lacked one allele or both alleles of the alpha-1,3-GT gene via nuclear transfer from somatic cells with targeted genetic deletions. In 2003, Sedai et al. (Transplantation (2003) 76:900-902) reported the targeted disruption of one allele of the alpha-1,3-GT gene in cattle, followed by the successful nuclear transfer of the nucleus of the genetically modified cell and production of transgenic fetuses.

Thus, the feasibility of knocking-out immunoglobulin genes in large animals and inserting human immunoglobulin loci into their cells is just now beginning to be explored. However, due to the complexity and species differences of immunoglobulin genes, the genomic sequences and arrangement of Ig kappa, lambda and heavy chains remain poorly understood in most species. For example, in pigs, partial genomic sequence and organization has only been described for heavy chain constant alpha, heavy chain constant mu and heavy chain constant delta (Brown and Butler Mol Immunol. 1994 Jun;31(8):633-42, Butler et al Vet Immunol Immunopathol. 1994 Oct;43(1-3):5-12, and Zhao et al J Immunol. 2003 Aug 1;171(3):1312-8).

In cows, the immunoglobulin heavy chain locus has been mapped (Zhao et al. 2003 J. Biol. Chem. 278:35024-32) and the cDNA sequence for the bovine kappa gene is known (See, for example, U.S. Patent Publication No. 2003/0037347). Further, approximately 4.6kb of the bovine mu heavy chain locus has been sequenced and transgenic calves with decreased expression of heavy chain immunoglobulins have been created by disrupting one or both alleles

of the bovine mu heavy chain. In addition, a mammalian artificial chromosome (MAC) vector containing the entire unarranged sequences of the human Ig H-chain and κ L-chain has been introduced into cows (TC cows) with the technology of microcell-mediated chromosome transfer and nuclear transfer of bovine fetal fibroblast cells (see, for example, Kuroiwa et al. 2002 Nature Biotechnology 20:889, Kuroiwa et al. 2004 Nat Genet. Jun 6 Epub, U.S. Patent Publication Nos. 2003/0037347, 2003/0056237, 2004/0068760 and PCT Publication No. WO 02/07648).

While significant progress has been made in the production of bovine that express human immunoglobulin, little has been accomplished in other large animals, such as sheep, goats and pigs. Although cDNA sequence information for immunoglobulin genes of sheeps, goats and pigs is readily available in Genbank, the unique nature of immunoglobulin loci, which undergo massive rearrangements, creates the need to characterize beyond sequences known to be present in mRNAs (or cDNAs). Since immunoglobulin loci are modular and the coding regions are redundant, deletion of a known coding region does not ensure altered function of the locus. For example, if one were to delete the coding region of a heavy-chain variable region, the function of the locus would not be significantly altered because hundreds of other function variable genes remain in the locus. Therefore, one must first characterize the locus to identify a potential "Achilles heel".

Despite some advancements in expressing human antibodies in cattle, greater challenges remain for inactivation of the endogenous bovine Ig genes, increasing expression levels of the human antibodies and creating human antibody expression in other large animals, such as porcine, for which the sequence and arrangement of immunoglobulin genes are largely unknown.

It is therefore an object of the present invention to provide the arrangement of ungulate immunoglobulin germline gene sequence.

It is another object of the present invention to provide novel ungulate immunoglobulin genomic sequences.

It is a further object of the present invention to provide cells, tissues and animals lacking at least one allele of a heavy and/or light chain immunoglobulin gene.

It is another object of the present invention to provide ungulates that express human immunoglobulins.

It is a still further object of the present invention to provide methods to generate cells, tissues and animals lacking at least one allele of novel ungulate immunoglobulin gene sequences and/ or express human immunoglobulins.

SUMMARY OF THE INVENTION

The present invention provides for the first time ungulate immunoglobulin germline gene sequence arrangement as well as novel genomic sequences thereof. In addition, novel ungulate cells, tissues and animals that lack at least one allele of a heavy or light chain immunoglobulin gene are provided. Based on this discovery, ungulates can be produced that completely lack at least one allele of a heavy and/or light chain immunoglobulin gene. In addition, these ungulates can be further modified to express xenogenous, such as human, immunoglobulin loci or fragments thereof.

In one aspect of the present invention, a transgenic ungulate that lacks any expression of functional endogenous immunoglobulins is provided. In one embodiment, the ungulate can lack any expression of endogenous heavy and/ or light chain immunoglobulins. The light chain immunoglobulin can be a kappa and/ or lambda immunoglobulin. In additional embodiments, transgenic ungulates are provided that lack expression of at least one allele of an endogenous immunoglobulin wherein the immunoglobulin is selected from the group consisting of heavy chain, kappa light chain and lambda light chain or any combination thereof. In one embodiment, the expression of functional endogenous immunoglobulins can be accomplished by genetic targeting of the endogenous immunoglobulin loci to prevent expression of the endogenous immunoglobulin. In one embodiment, the genetic targeting can be accomplished via homologous recombination. In another embodiment, the transgenic ungulate can be produced via nuclear transfer.

In other embodiments, the transgenic ungulate that lacks any expression of functional endogenous immunoglobulins can be further genetically modified to express an xenogenous immunoglobulin loci. In an alternative embodiment, porcine animals are provided that contain an xenogenous immunoglobulin locus. In one embodiment, the xenogenous immunoglobulin loci can be a heavy and/ or light chain immunoglobulin or fragment thereof. In another embodiment, the xenogenous immunoglobulin loci can be a kappa chain locus or fragment thereof and/ or a

lambda chain locus or fragment thereof. In still further embodiments, an artificial chromosome (AC) can contain the xenogenous immunoglobulin. In one embodiment, the AC can be a yeast AC or a mammalian AC. In a further embodiment, the xenogenous locus can be a human immunoglobulin locus or fragment thereof. In one embodiment, the human immunoglobulin locus can be human chromosome 14, human chromosome 2, and human chromosome 22 or fragments thereof. In another embodiment, the human immunoglobulin locus can include any fragment of a human immunoglobulin that can undergo rearrangement. In a further embodiment, the human immunoglobulin loci can include any fragment of a human immunoglobulin heavy chain and a human immunoglobulin light chain that can undergo rearrangement. In still further embodiment, the human immunoglobulin loci can include any human immunoglobulin locus or fragment thereof that can produce an antibody upon exposure to an antigen. In a particular embodiment, the exogenous human immunoglobulin can be expressed in B cells to produce xenogenous immunoglobulin in response to exposure to one or more antigens.

In another aspect of the present invention, transgenic ungulates are provided that expresses a xenogenous immunoglobulin loci or fragment thereof, wherein the immunoglobulin can be expressed from an immunoglobulin locus that is integrated within an endogenous ungulate chromosome. In one embodiment, ungulate cells derived from the transgenic animals are provided. In one embodiment, the xenogenous immunoglobulin locus can be inherited by offspring. In another embodiment, the xenogenous immunoglobulin locus can be inherited through the male germ line by offspring. In still further embodiments, an artificial chromosome (AC) can contain the xenogenous immunoglobulin. In one embodiment, the AC can be a yeast AC or a mammalian AC. In a further embodiment, the xenogenous locus can be a human immunoglobulin locus or fragment thereof. In one embodiment, the human immunoglobulin locus can be human chromosome 14, human chromosome 2, and human chromosome 22 or fragments thereof. In another embodiment, the human immunoglobulin locus can include any fragment of a human immunoglobulin that can undergo rearrangement. In a further embodiment, the human immunoglobulin loci can include any fragment of a human immunoglobulin heavy chain and a human immunoglobulin light chain that can undergo rearrangement. In still further embodiment, the human immunoglobulin loci can include any human immunoglobulin locus or fragment thereof that can produce an antibody upon exposure to an antigen. In a particular

embodiment, the exogenous human immunoglobulin can be expressed in B cells to produce xenogenous immunoglobulin in response to exposure to one or more antigens.

In another aspect of the present invention, novel genomic sequences encoding the heavy chain locus of ungulate immunoglobulin are provided. In one embodiment, an isolated nucleotide sequence encoding porcine heavy chain is provided that includes at least one variable region, two diversity regions, at least four joining regions and at least one constant region, such as the mu constant region, for example, as represented in Seq ID No. 29. In another embodiment, an isolated nucleotide sequence is provided that includes at least four joining regions and at least one constant region, such as the mu constant region, of the porcine heavy chain genomic sequence, for example, as represented in Seq ID No. 4. In a further embodiment, nucleotide sequence is provided that includes 5' flanking sequence to the first joining region of the porcine heavy chain genomic sequence, for example, as represented in Seq ID No 1. Still further, nucleotide sequence is provided that includes 3' flanking sequence to the first joining region of the porcine heavy chain genomic sequence, for example, as represented in the 3' region of Seq ID No 4. In further embodiments, isolated nucleotide sequences as depicted in Seq ID Nos 1, 4 or 29 are provided. Nucleic acid sequences at least 80, 85, 90, 95, 98 or 99% homologous to Seq ID Nos 1, 4 or 29 are also provided. In addition, nucleotide sequences that contain at least 10, 15, 17, 20, 25 or 30 contiguous nucleotides of Seq ID Nos 1, 4 or 29 are provided. In one embodiment, the nucleotide sequence contains at least 17, 20, 25 or 30 contiguous nucleotides of Seq ID No 4 or residues 1- 9,070 of Seq ID No 29.

In another embodiment, the nucleotide sequence contains residues 9,070-11039 of Seq ID No 29. Further provided are nucleotide sequences that hybridizes, optionally under stringent conditions, to Seq ID Nos 1, 4 or 29, as well as, nucleotides homologous thereto.

In another embodiment, novel genomic sequences encoding the kappa light chain locus of ungulate immunoglobulin are provided. The present invention provides the first reported genomic sequence of ungulate kappa light chain regions. In one embodiment, nucleic acid sequence is provided that encodes the porcine kappa light chain locus. In another embodiment, the nucleic acid sequence can contain at least one joining region, one constant region and/or one enhancer region of kappa light chain. In a further embodiment, the nucleotide sequence can include at least five joining regions, one constant region and one enhancer region, for example, as represented in Seq ID No. 30. In a further embodiment, an isolated nucleotide sequence is

provided that contains at least one, at least two, at least three, at least four or five joining regions and 3' flanking sequence to the joining region of porcine genomic kappa light chain, for example, as represented in Seq ID No 12. In another embodiment, an isolated nucleotide sequence of porcine genomic kappa light chain is provided that contains 5' flanking sequence to the first joining region, for example, as represented in Seq ID No 25. In a further embodiment, an isolated nucleotide sequence is provided that contains 3' flanking sequence to the constant region and, optionally, the 5' portion of the enhancer region, of porcine genomic kappa light chain, for example, as represented in Seq ID Nos. 15, 16 and/or 19.

In further embodiments, isolated nucleotide sequences as depicted in Seq ID Nos 30, 12, 25, 15, 16 or 19 are provided. Nucleic acid sequences at least 80, 85, 90, 95, 98 or 99% homologous to Seq ID Nos 30, 12, 25, 15, 16 or 19 are also provided. In addition, nucleotide sequences that contain at least 10, 15, 17, 20, 25 or 30 contiguous nucleotides of Seq ID Nos 30, 12, 25, 15, 16 or 19 are provided. Further provided are nucleotide sequences that hybridizes, optionally under stringent conditions, to Seq ID Nos 30, 12, 25, 15, 16 or 19, as well as, nucleotides homologous thereto.

In another embodiment, novel genomic sequences encoding the lambda light chain locus of ungulate immunoglobulin are provided. The present invention provides the first reported genomic sequence of ungulate lambda light chain regions. In one embodiment, the porcine lambda light chain nucleotides include a concatamer of J to C units. In a specific embodiment, an isolated porcine lambda nucleotide sequence is provided, such as that depicted in Seq ID No. 28. . In one embodiment, nucleotide sequence is provided that includes 5' flanking sequence to the first lambda J/C region of the porcine lambda light chain genomic sequence, for example, as represented by Seq ID No 32. Still further, nucleotide sequence is provided that includes 3' flanking sequence to the J/C cluster region of the porcine lambda light chain genomic sequence, for example, approximately 200 base pairs downstream of lambda J/C, such as that represented by Seq ID No 33. Alternatively, nucleotide sequence is provided that includes 3' flanking sequence to the J/C cluster region of the porcine lambda light chain genomic sequence, for example, as represented by Seq ID No 34, 35, 36, 37, 38, and/or 39. In a further embodiment, nucleic acid sequences are provided that encode bovine lambda light chain locus, which can include at least one joining region-constant region pair and/or at least one variable region, for example, as represented by Seq ID No. 31. In further embodiments, isolated nucleotide

sequences as depicted in Seq ID Nos 28, 31, 32, 33, 34, 35, 36, 37, 38, or 39 are provided. Nucleic acid sequences at least 80, 85, 90, 95, 98 or 99% homologous to Seq ID Nos 28, 31, 32, 33, 34, 35, 36, 37, 38, or 39 are also provided. In addition, nucleotide sequences that contain at least 10, 15, 17, 20, 25 or 30 contiguous nucleotides of Seq ID Nos 28, 31, 32, 33, 34, 35, 36, 37, 38, or 39 are provided. Further provided are nucleotide sequences that hybridizes, optionally under stringent conditions, to Seq ID Nos 28, 31, 32, 33, 34, 35, 36, 37, 38, or 39, as well as, nucleotides homologous thereto.

In another embodiment, nucleic acid targeting vector constructs are also provided. The targeting vectors can be designed to accomplish homologous recombination in cells. These targeting vectors can be transformed into mammalian cells to target the ungulate heavy chain, kappa light chain or lambda light chain genes via homologous recombination. In one embodiment, the targeting vectors can contain a 3' recombination arm and a 5' recombination arm (i.e. flanking sequence) that is homologous to the genomic sequence of ungulate heavy chain, kappa light chain or lambda light chain genomic sequence, for example, sequence represented by Seq ID Nos. 1, 4, 29, 30, 12, 25, 15, 16, 19, 28 or 31, as described above. The homologous DNA sequence can include at least 15 bp, 20 bp, 25 bp, 50 bp, 100 bp, 500 bp, 1kbp, 2 kbp, 4 kbp, 5 kbp, 10 kbp, 15 kbp, 20 kbp, or 50 kbp of sequence homologous to the genomic sequence. The 3' and 5' recombination arms can be designed such that they flank the 3' and 5' ends of at least one functional variable, joining, diversity, and/or constant region of the genomic sequence. The targeting of a functional region can render it inactive, which results in the inability of the cell to produce functional immunoglobulin molecules. In another embodiment, the homologous DNA sequence can include one or more intron and/or exon sequences. In addition to the nucleic acid sequences, the expression vector can contain selectable marker sequences, such as, for example, enhanced Green Fluorescent Protein (eGFP) gene sequences, initiation and/or enhancer sequences, poly A-tail sequences, and/or nucleic acid sequences that provide for the expression of the construct in prokaryotic and/or eukaryotic host cells. The selectable marker can be located between the 5' and 3' recombination arm sequence.

In one particular embodiment, the targeting vector can contain 5' and 3' recombination arms that contain homologous sequence to the 3' and 5' flanking sequence of the J6 region of the porcine immunoglobulin heavy chain locus. Since the J6 region is the only functional joining region of the porcine immunoglobulin heavy chain locus, this will prevent the expression of a

functional porcine heavy chain immunoglobulin. In a specific embodiment, the targeting vector can contain a 5' recombination arm that contains sequence homologous to genomic sequence 5' of the J6 region, including J1-4, and a 3' recombination arm that contains sequence homologous to genomic sequence 3' of the J6 region, including the mu constant region (a "J6 targeting construct"), see for example, Figure 1. Further, this J6 targeting construct can also contain a selectable marker gene that is located between the 5' and 3' recombination arms, see for example, Seq ID No 5 and Figure 1. In other embodiments, the targeting vector can contain a 5' recombination arm that contains sequence homologous to genomic sequence 5' of the diversity region, and a 3' recombination arm that contains sequence homologous to genomic sequence 3' of the diversity region of the porcine heavy chain locus. In a further embodiment, the targeting vector can contain a 5' recombination arm that contains sequence homologous to genomic sequence 5' of the mu constant region and a 3' recombination arm that contains sequence homologous to genomic sequence 3' of the mu constant region of the porcine heavy chain locus.

In another particular embodiment, the targeting vector can contain 5' and 3' recombination arms that contain homologous sequence to the 3' and 5' flanking sequence of the constant region of the porcine immunoglobulin heavy chain locus. Since the present invention discovered that there is only one constant region of the porcine immunoglobulin kappa light chain locus, this will prevent the expression of a functional porcine kappa light chain immunoglobulin. In a specific embodiment, the targeting vector can contain a 5' recombination arm that contains sequence homologous to genomic sequence 5' of the constant region, optionally including the joining region, and a 3' recombination arm that contains sequence homologous to genomic sequence 3' of the constant region, optionally including at least part of the enhancer region (a "Kappa constant targeting construct"), see for example, Figure 2. Further, this kappa constant targeting construct can also contain a selectable marker gene that is located between the 5' and 3' recombination arms, see for example, Seq ID No 20 and Figure 2. In other embodiments, the targeting vector can contain a 5' recombination arm that contains sequence homologous to genomic sequence 5' of the joining region, and a 3' recombination arm that contains sequence homologous to genomic sequence 3' of the joining region of the porcine kappa light chain locus.

In another embodiment, primers are provided to generate 3' and 5' sequences of a targeting vector. The oligonucleotide primers can be capable of hybridizing to porcine

immunoglobulin genomic sequence, such as Seq ID Nos. 1, 4, 29, 30, 12, 25, 15, 16, 19, 28 or 31, as described above. In a particular embodiment, the primers hybridize under stringent conditions to Seq ID Nos. 1, 4, 29, 30, 12, 25, 15, 16, 19, 28 or 31, as described above. Another embodiment provides oligonucleotide probes capable of hybridizing to porcine heavy chain, kappa light chain or lambda light chain nucleic acid sequences, such as Seq ID Nos. 1, 4, 29, 30, 12, 25, 15, 16, 19, 28 or 31, as described above. The polynucleotide primers or probes can have at least 14 bases, 20 bases, 30 bases, or 50 bases which hybridize to a polynucleotide of the present invention. The probe or primer can be at least 14 nucleotides in length, and in a particular embodiment, are at least 15, 20, 25, 28, or 30 nucleotides in length.

In one embodiment, primers are provided to amplify a fragment of porcine Ig heavy-chain that includes the functional joining region (the J6 region). In one non-limiting embodiment, the amplified fragment of heavy chain can be represented by Seq ID No 4 and the primers used to amplify this fragment can be complementary to a portion of the J-region, such as, but not limited to Seq ID No 2, to produce the 5' recombination arm and complementary to a portion of Ig heavy-chain mu constant region, such as, but not limited to Seq ID No 3, to produce the 3' recombination arm. In another embodiment, regions of the porcine Ig heavy chain (such as, but not limited to Seq ID No 4) can be subcloned and assembled into a targeting vector.

In other embodiments, primers are provided to amplify a fragment of porcine Ig kappa light-chain that includes the constant region. In another embodiment, primers are provided to amplify a fragment of porcine Ig kappa light-chain that includes the J region. In one non-limiting embodiment, the primers used to amplify this fragment can be complementary to a portion of the J-region, such as, but not limited to Seq ID No 21 or 10, to produce the 5' recombination arm and complementary to genomic sequence 3' of the constant region, such as, but not limited to Seq ID No 14, 24 or 18, to produce the 3' recombination arm. In another embodiment, regions of the porcine Ig heavy chain (such as, but not limited to Seq ID No 20) can be subcloned and assembled into a targeting vector.

In another aspect of the present invention, ungulate cells lacking at least one allele of a functional region of an ungulate heavy chain, kappa light chain and/or lambda light chain locus produced according to the process, sequences and/or constructs described herein are provided. These cells can be obtained as a result of homologous recombination. Particularly, by inactivating at least one allele of an ungulate heavy chain, kappa light chain or lambda light

chain gene, cells can be produced which have reduced capability for expression of ungulate antibodies. In other embodiments, mammalian cells lacking both alleles of an ungulate heavy chain, kappa light chain and/or lambda light chain gene can be produced according to the process, sequences and/or constructs described herein. In a further embodiment, porcine animals are provided in which at least one allele of an ungulate heavy chain, kappa light chain and/or lambda light chain gene is inactivated via a genetic targeting event produced according to the process, sequences and/or constructs described herein. In another aspect of the present invention, porcine animals are provided in which both alleles of an ungulate heavy chain, kappa light chain and/or lambda light chain gene are inactivated via a genetic targeting event. The gene can be targeted via homologous recombination.

In other embodiments, the gene can be disrupted, i.e. a portion of the genetic code can be altered, thereby affecting transcription and/or translation of that segment of the gene. For example, disruption of a gene can occur through substitution, deletion ("knock-out") or insertion ("knock-in") techniques. Additional genes for a desired protein or regulatory sequence that modulate transcription of an existing sequence can be inserted. To achieve multiple genetic modifications of ungulate immunoglobulin genes, in one embodiment, cells can be modified sequentially to contain multiple genetic modifications. In other embodiments, animals can be bred together to produce animals that contain multiple genetic modifications of immunoglobulin genes. As an illustrative example, animals that lack expression of at least one allele of an ungulate heavy chain gene can be further genetically modified or bred with animals lacking at least one allele of a kappa light chain gene.

In embodiments of the present invention, alleles of ungulate heavy chain, kappa light chain or lambda light chain gene are rendered inactive according to the process, sequences and/or constructs described herein, such that functional ungulate immunoglobulins can no longer be produced. In one embodiment, the targeted immunoglobulin gene can be transcribed into RNA, but not translated into protein. In another embodiment, the targeted immunoglobulin gene can be transcribed in an inactive truncated form. Such a truncated RNA may either not be translated or can be translated into a nonfunctional protein. In an alternative embodiment, the targeted immunoglobulin gene can be inactivated in such a way that no transcription of the gene occurs. In a further embodiment, the targeted immunoglobulin gene can be transcribed and then translated into a nonfunctional protein.

In a further aspect of the present invention, ungulate, such as porcine or bovine, cells lacking one allele, optionally both alleles of an ungulate heavy chain, kappa light chain and/or lambda light chain gene can be used as donor cells for nuclear transfer into recipient cells to produce cloned, transgenic animals. Alternatively, ungulate heavy chain, kappa light chain and/or lambda light chain gene knockouts can be created in embryonic stem cells, which are then used to produce offspring. Offspring lacking a single allele of a functional ungulate heavy chain, kappa light chain and/or lambda light chain gene produced according to the process, sequences and/or constructs described herein can be bred to further produce offspring lacking functionality in both alleles through mendelian type inheritance.

In one aspect of the present invention, a method is provided to disrupt the expression of an ungulate immunoglobulin gene by (i) analyzing the germline configuration of the ungulate heavy chain, kappa light chain or lambda light chain genomic locus; (ii) determining the location of nucleotide sequences that flank the 5' end and the 3' end of at least one functional region of the locus; and (iii) transfecting a targeting construct containing the flanking sequence into a cell wherein, upon successful homologous recombination, at least one functional region of the immunoglobulin locus is disrupted thereby reducing or preventing the expression of the immunoglobulin gene. In one embodiment, the germline configuration of the porcine heavy chain locus is provided. The porcine heavy chain locus contains at least four variable regions, two diversity regions, six joining regions and five constant regions, for example, as illustrated in Figure 1. In a specific embodiment, only one of the six joining regions, J6, is functional. In another embodiment, the germline configuration of the porcine kappa light chain locus is provided. The porcine kappa light chain locus contains at least six variable regions, six joining regions, one constant region and one enhancer region, for example, as illustrated in Figure 2. In a further embodiment, the germline configuration of the porcine lambda light chain locus is provided.

In further aspects of the present invention provides ungulates and ungulate cells that lack at least one allele of a functional region of an ungulate heavy chain, kappa light chain and/or lambda light chain locus produced according to the processes, sequences and/or constructs described herein, which are further modified to express at least part of a human antibody (i.e. immunoglobulin (Ig)) locus. In additional embodiments, porcine animals are provided that express xenogenous immunoglobulin. This human locus can undergo rearrangement and

express a diverse population of human antibody molecules in the ungulate. These cloned, transgenic ungulates provide a replenishable, theoretically infinite supply of human antibodies (such as polyclonal antibodies), which can be used for therapeutic, diagnostic, purification, and other clinically relevant purposes. In one particular embodiment, artificial chromosomes (ACs), such as yeast or mammalian artificial chromosomes (YACS or MACS) can be used to allow expression of human immunoglobulin genes into ungulate cells and animals. All or part of human immunoglobulin genes, such as the Ig heavy chain gene (human chromosome 414), Ig kappa chain gene (human chromosome #2) and/or the Ig lambda chain gene (chromosome #22) can be inserted into the artificial chromosomes, which can then be inserted into ungulate cells. In further embodiments, ungulates and ungulate cells are provided that contain either part or all of at least one human antibody gene locus, which undergoes rearrangement and expresses a diverse population of human antibody molecules.

In additional embodiments, methods of producing xenogenous antibodies are provided, wherein the method can include: (a) administering one or more antigens of interest to an ungulate whose cells comprise one or more artificial chromosomes and lack any expression of functional endogenous immunoglobulin, each artificial chromosome comprising one or more xenogenous immunoglobulin loci that undergo rearrangement, resulting in production of xenogenous antibodies against the one or more antigens; and/ or (b) recovering the xenogenous antibodies from the ungulate. In one embodiment, the immunoglobulin loci can undergo rearrangement in a B cell.

In one aspect of the present invention, an ungulate, such as a pig or a cow, can be prepared by a method in accordance with any aspect of the present invention. These cloned, transgenic ungulates (e.g., porcine and bovine animals) provide a replenishable, theoretically infinite supply of human polyclonal antibodies, which can be used as therapeutics, diagnostics and for purification purposes. For example, transgenic animals produced according to the process, sequences and/or constructs described herein that produce polyclonal human antibodies in the bloodstream can be used to produce an array of different antibodies which are specific to a desired antigen. The availability of large quantities of polyclonal antibodies can also be used for treatment and prophylaxis of infectious disease, vaccination against biological warfare agents, modulation of the immune system, removal of undesired human cells such as cancer cells, and modulation of specific human molecules.

In other embodiments, animals or cells lacking expression of functional immunoglobulin, produced according to the process, sequences and/or constructs described herein, can contain additional genetic modifications to eliminate the expression of xenoantigens. Such animals can be modified to eliminate the expression of at least one allele of the alpha-1,3-galactosyltransferase gene, the CMP-Neu5Ac hydroxylase gene (see, for example, USSN 10/863,116), the iGb3 synthase gene (see, for example, U.S. Patent Application 60/517,524), and/or the Forssman synthase gene (see, for example, U.S. Patent Application 60/568,922). In additional embodiments, the animals disclosed herein can also contain genetic modifications to express fucosyltransferase and/or sialyltransferase. To achieve these additional genetic modifications, in one embodiment, cells can be modified to contain multiple genetic modifications. In other embodiments, animals can be bred together to achieve multiple genetic modifications. In one specific embodiment, animals, such as pigs, lacking expression of functional immunoglobulin, produced according to the process, sequences and/or constructs described herein, can be bred with animals, such as pigs, lacking expression of alpha-1,3-galactosyl transferase (for example, as described in WO 04/028243).

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates the design of a targeting vector that disrupts the expression of the joining region of the porcine heavy chain immunoglobulin gene.

Figure 2 illustrates the design of a targeting vector that disrupts the expression of the constant region of the porcine kappa light chain immunoglobulin gene.

Figure 3 illustrates the genomic organization of the porcine lambda immunoglobulin locus, including a concatamer of J-C sequences as well as flanking regions that include the variable region 5' to the JC region. Bacterial artificial chromosomes (BAC1 and BAC2) represent fragments of the porcine immunoglobulin genome that can be obtained from BAC libraries.

Figure 4 represents the design of a targeting vector that disrupts the expression of the JC cluster region of the porcine lambda light chain immunoglobulin gene. "SM" stands for a selectable marker gene, which can be used in the targeting vector.

Figure 5 illustrates a targeting strategy to insert a site specific recombinase target or recognition site into the region 5' of the JC cluster region of the porcine lambda immunoglobulin locus. "SM" stands for a selectable marker gene, which can be used in the targeting vector. "SSRRS" stands for a specific recombinase target or recognition site.

Figure 6 illustrates a targeting strategy to insert a site specific recombinase target or recognition site into the region 3' of the JC cluster region of the porcine lambda immunoglobulin locus. "SM" stands for a selectable marker gene, which can be used in the targeting vector. "SSRRS" stands for a specific recombinase target or recognition site.

Figure 7 illustrates the site specific recombinase mediated transfer of a YAC into a host genome. "SSRRS" stands for a specific recombinase target or recognition site.

DETAILED DESCRIPTION

The present invention provides for the first time ungulate immunoglobulin germline gene sequence arrangement as well as novel genomic sequences thereof. In addition, novel ungulate cells, tissues and animals that lack at least one allele of a heavy or light chain immunoglobulin gene are provided. Based on this discovery, ungulates can be produced that completely lack at least one allele of a heavy and/or light chain immunoglobulin gene. In addition, these ungulates can be further modified to express xenoogenous, such as human, immunoglobulin loci or fragments thereof.

In one aspect of the present invention, a transgenic ungulate that lacks any expression of functional endogenous immunoglobulins is provided. In one embodiment, the ungulate can lack any expression of endogenous heavy and/ or light chain immunoglobulins. The light chain immunoglobulin can be a kappa and/ or lambda immunoglobulin. In additional embodiments, transgenic ungulates are provided that lack expression of at least one allele of an endogenous immunoglobulin wherein the immunoglobulin is selected from the group consisting of heavy chain, kappa light chain and lambda light chain or any combination thereof. In one embodiment, the expression of functional endogenous immunoglobulins can be accomplished by genetic targeting of the endogenous immunoglobulin loci to prevent expression of the endogenous immunoglobulin. In one embodiment, the genetic targeting can be accomplished via

homologous recombination. In another embodiment, the transgenic ungulate can be produced via nuclear transfer.

In other embodiments, the transgenic ungulate that lacks any expression of functional endogenous immunoglobulins can be further genetically modified to express an xenogenous immunoglobulin loci. In an alternative embodiment, porcine animals are provided that contain an xenogenous immunoglobulin locus. In one embodiment, the xenogenous immunoglobulin loci can be a heavy and/ or light chain immunoglobulin or fragment thereof. In another embodiment, the xenogenous immunoglobulin loci can be a kappa chain locus or fragment thereof and/ or a lambda chain locus or fragment thereof. In still further embodiments, an artificial chromosome (AC) can contain the xenogenous immunoglobulin. In one embodiment, the AC can be a yeast AC or a mammalian AC. In a further embodiment, the xenogenous locus can be a human immunoglobulin locus or fragment thereof. In one embodiment, the human immunoglobulin locus can be human chromosome 14, human chromosome 2, and human chromosome 22 or fragments thereof. In another embodiment, the human immunoglobulin locus can include any fragment of a human immunoglobulin that can undergo rearrangement. In a further embodiment, the human immunoglobulin loci can include any fragment of a human immunoglobulin heavy chain and a human immunoglobulin light chain that can undergo rearrangement. In still further embodiment, the human immunoglobulin loci can include any human immunoglobulin locus or fragment thereof that can produce an antibody upon exposure to an antigen. In a particular embodiment, the exogenous human immunoglobulin can be expressed in B cells to produce xenogenous immunoglobulin in response to exposure to one or more antigens.

In another aspect of the present invention, transgenic ungulates are provided that expresses a xenogenous immunoglobulin loci or fragment thereof, wherein the immunoglobulin can be expressed from an immunoglobulin locus that is integrated within an endogenous ungulate chromosome. In one embodiment, ungulate cells derived from the transgenic animals are provided. In one embodiment, the xenogenous immunoglobulin locus can be inherited by offspring. In another embodiment, the xenogenous immunoglobulin locus can be inherited through the male germ line by offspring. In still further embodiments, an artificial chromosome (AC) can contain the xenogenous immunoglobulin. In one embodiment, the AC can be a yeast AC or a mammalian AC. In a further embodiment, the xenogenous locus can be a human immunoglobulin locus or fragment thereof. In one embodiment, the human immunoglobulin

locus can be human chromosome 14, human chromosome 2, and human chromosome 22 or fragments thereof. In another embodiment, the human immunoglobulin locus can include any fragment of a human immunoglobulin that can undergo rearrangement. In a further embodiment, the human immunoglobulin loci can include any fragment of a human immunoglobulin heavy chain and a human immunoglobulin light chain that can undergo rearrangement. In still further embodiment, the human immunoglobulin loci can include any human immunoglobulin locus or fragment thereof that can produce an antibody upon exposure to an antigen. In a particular embodiment, the exogenous human immunoglobulin can be expressed in B cells to produce xenogenous immunoglobulin in response to exposure to one or more antigens.

Definitions

The terms "recombinant DNA technology," "DNA cloning," "molecular cloning," or "gene cloning" refer to the process of transferring a DNA sequence into a cell or organism. The transfer of a DNA fragment can be from one organism to a self-replicating genetic element (e.g., bacterial plasmid) that permits a copy of any specific part of a DNA (or RNA) sequence to be selected among many others and produced in an unlimited amount. Plasmids and other types of cloning vectors such as artificial chromosomes can be used to copy genes and other pieces of chromosomes to generate enough identical material for further study. In addition to bacterial plasmids, which can carry up to 20 kb of foreign DNA, other cloning vectors include viruses, cosmids, and artificial chromosomes (e.g., bacteria artificial chromosomes (BACs) or yeast artificial chromosomes (YACs)). When the fragment of chromosomal DNA is ultimately joined with its cloning vector in the lab, it is called a "recombinant DNA molecule." Shortly after the recombinant plasmid is introduced into suitable host cells, the newly inserted segment will be reproduced along with the host cell DNA.

"Cosmids" are artificially constructed cloning vectors that carry up to 45 kb of foreign DNA. They can be packaged in lambda phage particles for infection into *E. coli* cells.

As used herein, the term "mammal" (as in "genetically modified (or altered) mammal") is meant to include any non-human mammal, including but not limited to pigs, sheep, goats, cattle (bovine), deer, mules, horses, monkeys, dogs, cats, rats, mice, birds, chickens, reptiles, fish, and insects. In one embodiment of the invention, genetically altered pigs and methods of production thereof are provided.

The term "ungulate" refers to hoofed mammals. Artiodactyls are even-toed (cloven-hooved) ungulates, including antelopes, camels, cows, deer, goats, pigs, and sheep. Perissodactyls are odd toes ungulates, which include horses, zebras, rhinoceroses, and tapirs. The term ungulate as used herein refers to an adult, embryonic or fetal ungulate animal.

As used herein, the terms "porcine", "porcine animal", "pig" and "swine" are generic terms referring to the same type of animal without regard to gender, size, or breed.

A "homologous DNA sequence or homologous DNA" is a DNA sequence that is at least about 80%, 85%, 90%, 95%, 98% or 99% identical with a reference DNA sequence. A homologous sequence hybridizes under stringent conditions to the target sequence, stringent hybridization conditions include those that will allow hybridization occur if there is at least 85, at least 95% or 98% identity between the sequences.

An "isogenic or substantially isogenic DNA sequence" is a DNA sequence that is identical to or nearly identical to a reference DNA sequence. The term "substantially isogenic" refers to DNA that is at least about 97-99% identical with the reference DNA sequence, or at least about 99.5-99.9% identical with the reference DNA sequence, and in certain uses 100% identical with the reference DNA sequence.

"Homologous recombination" refers to the process of DNA recombination based on sequence homology.

"Gene targeting" refers to homologous recombination between two DNA sequences, one of which is located on a chromosome and the other of which is not.

"Non-homologous or random integration" refers to any process by which DNA is integrated into the genome that does not involve homologous recombination.

A "selectable marker gene" is a gene, the expression of which allows cells containing the gene to be identified. A selectable marker can be one that allows a cell to proliferate on a medium that prevents or slows the growth of cells without the gene. Examples include antibiotic resistance genes and genes which allow an organism to grow on a selected metabolite. Alternatively, the gene can facilitate visual screening of transformants by conferring on cells a phenotype that is easily identified. Such an identifiable phenotype can be, for example, the production of luminescence or the production of a colored compound, or the production of a detectable change in the medium surrounding the cell.

The term "contiguous" is used herein in its standard meaning, i.e., without interruption, or uninterrupted.

"Stringent conditions" refers to conditions that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50°C, or (2) employ during hybridization a denaturing agent such as, for example, formamide. One skilled in the art can determine and vary the stringency conditions appropriately to obtain a clear and detectable hybridization signal. For example, stringency can generally be reduced by increasing the salt content present during hybridization and washing, reducing the temperature, or a combination thereof. See, for example, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York, (1989).

I. Immunoglobulin Genes

In one aspect of the present invention, a transgenic ungulate that lacks any expression of functional endogenous immunoglobulins is provided. In one embodiment, the ungulate can lack any expression of endogenous heavy and/ or light chain immunoglobulins. The light chain immunoglobulin can be a kappa and/ or lambda immunoglobulin. In additional embodiments, transgenic ungulates are provided that lack expression of at least one allele of an endogenous immunoglobulin wherein the immunoglobulin is selected from the group consisting of heavy chain, kappa light chain and lambda light chain or any combination thereof. In one embodiment, the expression of functional endogenous immunoglobulins can be accomplished by genetic targeting of the endogenous immunoglobulin loci to prevent expression of the endogenous immunoglobulin. In one embodiment, the genetic targeting can be accomplished via homologous recombination. In another embodiment, the transgenic ungulate can be produced via nuclear transfer.

In another aspect of the present invention, a method is provided to disrupt the expression of an ungulate immunoglobulin gene by (i) analyzing the germline configuration of the ungulate heavy chain, kappa light chain or lambda light chain genomic locus; (ii) determining the location of nucleotide sequences that flank the 5' end and the 3' end of at least one functional region of the locus; and (iii) transfecting a targeting construct containing the flanking sequence into a cell wherein, upon successful homologous recombination, at least one functional region of the

immunoglobulin locus is disrupted thereby reducing or preventing the expression of the immunoglobulin gene.

In one embodiment, the germline configuration of the porcine heavy chain locus is provided. The porcine heavy chain locus contains at least four variable regions, two diversity regions, six joining regions and five constant regions, for example, as illustrated in Figure 1. In a specific embodiment, only one of the six joining regions, J6, is functional.

In another embodiment, the germline configuration of the porcine kappa light chain locus is provided. The porcine kappa light chain locus contains at least six variable regions, six joining regions, one constant region and one enhancer region, for example, as illustrated in Figure 2.

In a further embodiment, the germline configuration of the porcine lambda light chain locus is provided.

Isolated nucleotide sequences as depicted in Seq ID Nos 1-39 are provided. Nucleic acid sequences at least 80, 85, 90, 95, 98 or 99% homologous to any one of Seq ID Nos 1-39 are also provided. In addition, nucleotide sequences that contain at least 10, 15, 17, 20, 25 or 30 contiguous nucleotides of any one of Seq ID Nos 1-39 are provided. Further provided are nucleotide sequences that hybridizes, optionally under stringent conditions, to Seq ID Nos 1-39, as well as, nucleotides homologous thereto.

Homology or identity at the nucleotide or amino acid sequence level can be determined by BLAST (Basic Local Alignment Search Tool) analysis using the algorithm employed by the programs blastp, blastn, blastx, tblastn and tblastx (see, for example, Altschul, S.F. et al (1997) Nucleic Acids Res 25:3389-3402 and Karlin et al, (1990) Proc. Natl. Acad. Sci. USA 87, 2264-2268) which are tailored for sequence similarity searching. The approach used by the BLAST program is to first consider similar-segments, with and without gaps, between a-query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only those matches which satisfy a preselected threshold of significance. See, for example, Altschul et al, (1994) (Nature Genetics 6, 119-129). The search parameters for histogram, descriptions, alignments, expect (ie., the statistical significance threshold for reporting matches against database sequences), cutoff, matrix and filter (low complexity) are at the default settings. The default scoring matrix used by blastp, blastx, tblastn, and tblastx is the BLOSUM62 matrix (Henikoff et al, (1992) Proc. Natl. Acad. Sci. USA 89,

10915-10919), which is recommended for query sequences over 85 in length (nucleotide bases or amino acids).

Porcine Heavy Chain

In another aspect of the present invention, novel genomic sequences encoding the heavy chain locus of ungulate immunoglobulin are provided. In one embodiment, an isolated nucleotide sequence encoding porcine heavy chain is provided that includes at least one variable region, two diversity regions, at least four joining regions and at least one constant region, such as the mu constant region, for example, as represented in Seq ID No. 29. In another embodiment, an isolated nucleotide sequence is provided that includes at least four joining regions and at least one constant region, such as the mu constant region, of the porcine heavy chain genomic sequence, for example, as represented in Seq ID No. 4. In a further embodiment, nucleotide sequence is provided that includes 5' flanking sequence to the first joining region of the porcine heavy chain genomic sequence, for example, as represented in Seq ID No 1. Still further, nucleotide sequence is provided that includes 3' flanking sequence to the first joining region of the porcine heavy chain genomic sequence, for example, as represented in the 3' region of Seq ID No 4. In further embodiments, isolated nucleotide sequences as depicted in Seq ID Nos 1, 4 or 29 are provided. Nucleic acid sequences at least 80, 85, 90, 95, 98 or 99% homologous to Seq ID Nos 1, 4 or 29 are also provided. Further provided are nucleotide sequences that hybridizes, optionally under stringent conditions, to Seq ID Nos 1, 4 or 29, as well as, nucleotides homologous thereto.

In addition, nucleotide sequences that contain at least 10, 15, 17, 20, 25 or 30 contiguous nucleotides of Seq ID Nos 1, 4 or 29 are provided. In one embodiment, the nucleotide sequence contains at least 17, 20, 25 or 30 contiguous nucleotides of Seq ID No 4 or residues 1- 9,070 of Seq ID No 29. In other embodiments, nucleotide sequences that contain at least 50, 100, 1,000, 2,500, 4,000, 4,500, 5,000, 7,000, 8,000, 8,500, 9,000, 10,000 or 15,000 contiguous nucleotides of Seq ID No. 29 are provided. In another embodiment, the nucleotide sequence contains residues 9,070-11039 of Seq ID No 29.

In further embodiments, isolated nucleotide sequences as depicted in Seq ID Nos 1, 4 or 29 are provided. Nucleic acid sequences at least 80, 85, 90, 95, 98 or 99% homologous to Seq ID Nos 1, 4 or 29 are also provided. In addition, nucleotide sequences that contain at least 10,

15, 17, 20, 25 or 30 contiguous nucleotides of Seq ID Nos 1, 4 or 29 are provided. Further provided are nucleotide sequences that hybridizes, optionally under stringent conditions, to Seq ID Nos 1, 4 or 29, as well as, nucleotides homologous thereto.

In one embodiment, an isolated nucleotide sequence encoding porcine heavy chain is provided that includes at least one variable region, two diversity regions, at least four joining regions and at least one constant region, such as the mu constant region, for example, as represented in Seq ID No. 29. In Seq ID No. 29, the Diversity region of heavy chain is represented, for example, by residues 1089-1099 (D(pseudo)), the Joining region of heavy chain is represented, for example, by residues 1887-3352 (for example: J(psuedo): 1887-1931, J(psuedo): 2364-2411, J(psuedo): 2756-2804, J (functional J): 3296-3352), the recombination signals are represented, for example, by residues 3001-3261 (Nonamer), 3292-3298 (Heptamer), the Constant Region is represented by the following residues: 3353-9070 (J to C mu intron), 5522-8700 (Switch region), 9071-9388 (Mu Exon 1), 9389-9469 (Mu Intron A), 9470-9802 (Mu Exon 2), 9830- 10069 (Mu Intron B), 10070-10387 (Mu Exon 3), 10388-10517 (Mu Intron C), 10815-11052 (Mu Exon 4), 11034-11039 (Poly(A) signal).

Seq ID No. 29	tctagaagacgctggagagaggccagacttcctcggaacagctcaaagag ctctgtcaaagccagatcccatcacacgtgggcaccaataggccatgcc gcctccaagggccgaactgggttctccacggcgccacatgaagcctgcagc ctggcttatccttccgtggtgaagaggcaggccgggactggacgagg ggctagcaggggtgtgtaggcaccttgccccccacccggcaggaacc agagaccctggggctgagagtgcctccaaacaggatgccccacccttc aggccaccttcaatccagctacactccacctgccattcctctgggca cagggcccagcccctggatcttggccttggctcgacttgacccacgcgc acacacacacttctaactgtgtgctcgctcaccctccccagcgtggtc catgggcagcacggcagtgccgctccggcggtagtgcagaggtccc tccccccccaggagccccagggtgtgtgcagatctgggggctcctg tcccttacaccttcatgccccctccctcataccaccctccaggcgggag gcagcgagaccttggcccagggactcagccaacgggcacacgggaggcca gccctcagcagctggctcccaaaggagggtgggaggttaggtccacagct gccacagagagaaacctgacggacccacaggggccacgccagccggaa ccagctccctcgtgggtgagcaatggccaggggccccgcccggccaccacgg ctggccttgcgccagctgagaactcacgtccagtgaggagactcaaga cagcctgtgcacacagcctcggatctgctcccatttcaagcagaaaaagg aaaccgtgcaggcagccctcagcattcaaggattgtgcagcgccaac tattcgtcggcagtgccgattagaatgaccgtggagaaggcggaagg tctctgtgggctctgcggccaacaggccctggctccacctgcccgtgc cagcccaggggcttggggcagccaggaaccacagtgtcaccgggacc acagtgcagcaaaactccggccagagcagccccaggccagccgggct ctcgccctggaggactaccatcagatgcacaagggggcgagtgtgaag agacgtgtcggcggccatttgggaaggcgaaggacctccaggtgga caggaggtgggacgcactccaggcaagggactgggtcccaaggcctggg
---------------	--

gaaggggtactggcttgggggttagcctggccagggaacggggagcgggg
 cggggggctgagcagggaggacctgacctcgtgggagcgaggcaagtcag
 gcttcaggcagcagccgcacatcccagaccaggaggtgaggcaggaggg
 gcttcagcggggcggggcctgcctggctccggggctcctgggggacg
 ctggctcttgttccgtgtcccgcagcacaggggcagctcgtgggcta
 tgcctacctgatgtctggggccggggcgtcagggtcgtcctcctcag
 gggagagtccctgaggctacgtgggg*ggggactatggcagctccacc
 aggggcctggggaccaggggcctggaccaggctgcagcccggaggacggg
 cagggtcttggctctccagcatctggccctcggaatggcagaaccctg
 gcgggtgagcgagctgagagcgggtcagacagacaggggccggccgaaa
 ggagaagtggggcagagcccggcaggggccaggcccaagggtctgtgt
 gccagggcctgggtgggcacattggtgtggccatggctacttagattcgt
 ggggccagggcacctgtgtaccgtctcctcaggtgagcctggtgtctga
 tgtccagctaggcgctggtggggcgggtgggctgtctcaggctaggg
 caggggctgggatgtgtattgtcaaggaggggcaacagggtgcagactg
 tgccctggaaactgaccactggggcaggggcgtcctggtcacgtctcc
 tcaggtaagacggccctgtgcccctctctcggggactgaaaaggaatt
 ttccaagattccttggctgtgtggggccctctggggccccgggggtgg
 ctcccctcctgcccagatggggcctcgccctgtggagcacgggctgggca
 cacagctcagcttagggccacagaggcccgggtcagggtctgtgtgg
 cccggcgactggcagggggctcgggttttggacacccccctaatgggggc
 cacagcactgtgaccatctcacagctggggccgaggagtcgaggtcacc
 gtctcctcaggtgagtcctcgtcagccctctctcactctctggggggtt
 tgctgcatttgtgggggaaagaggatgcctgggtctcaggtctaaaggt
 ctaggggcagcgccggggcccagggaaggggccgaggggccaggctcggct
 cgggcaggagcagagctccagacatctgcctcctggcggtgcagtca
 ggccttggccgggggggtctcagcaccaccaggcctcttggctcccag
 gtccccggccccggctgcctcaccaggcaccgtgcgcgggtggggccgggc
 tcttggtcggccacccttcttaactgggatccgggcttagttgtcgcaa
 tgtgacaacgggtcgaagctggggccaggggaccctagtctacgacgc
 ctgggtgggtgtccgcacccctccccactttcacggcactcggcgaga
 cctggggagtcaggtgtggggacactttggaggtcaggaacgggagctg
 gggagagggctctgtcagcgggtccagagatggggcccccctcaaggac
 gccctgcgcggggacaagggtcttggcctggcctggccgcttacttg
 ggcgtcagggggggcttccggggcaggcggtcagtcgaggcggttga
 attctgagcttgggtcgggttcgggttcggcctcatgaacagacag
 cccaggcggggcgttgttggcccctgggggcctggttgaatgcgaggt
 ctggggaagttagggggagcctggccagcagagggtcccagccctgcg
 gccgagggacctggagacgggcagggcattggcgtcgcagggccaggcc
 acaccccccaGGTTTTTGTggggcgagcctggagattgcacCACTGTGAT
 TACTATGCTATGGATCTCTGGGGCCCAGGCGTTGAAGTCGTCGTGTCCTC
 AGgtaagaacggccctccaggggcctttaattctgtctcgtctgtgggc
 tttctgactctgatctcgggaggcgtctgtgcccccccggggatgag
 gccggcttgcaggagggtcagggaccaggagcctgtgggaagttctga
 cgggggctgcaggcgggaaggggccccaccggggggcgagccccaggccgc
 tgggcggcaggagaccgtgagagtgccctgaggagggtgtctgcgga
 accacgaacggccggggaagggttgcctgcaatgcggtcttcagacgg
 gaggcgtcttgcctcaccgtcttcaagccctgtgggtctgaaaga
 gccatgtcggagagagaaggacaggcctgtcccagctggccgagagcg
 ggcagccccgggggagagcggggcgatcgccctgggctctgtgaggccag
 gtccaaggaggacgtgtgtcctcgtgacaggtgcacttgcgaaacctt
 agaagacggggtatgttgaagcggctcctgatgtttaagaaaagggaga
 ctgtaaagttagcagagtcctcaagtgtgttaagggtttaaggtcaaag
 tgttttaaaccttctgactgcagttagcaagcgtcggggagtgaaatgg
 ggtgccagggtggccgagaggcagtagggccgtgccgtcctctaatt
 cagggttagtttgcagaataaagtcggcctgtttctaaaagcattgg

tggtgctgagctggtggaggaggccgcgggcagccctggccacctgcagc
aggtggcaggaagcaggtcggccaagaggctattttaggaagccagaaaa
cacggctgatgaattatagcttctggttccaggagggtggtgggcatg
gcttgcgcagcggccacagaaccgaaagtcccactgagaaaaacaact
cctgcftaatttgcatttttctaaaagaagaacagaggctgacggaaac
tggaaggttctgttttaactactcgaattgagtttctggtcttagctta
tcaactgctcacttagattcatttcaaagtaaactttaagagccgagg
cattcctatcctcttctaaggcggtattcttgagggtcattcacgcca
gcacctccgctgcctgcaggcattgctgcaccgtcaccgtgacggcgcg
cacgattttaggttggcccgttccctcgtgattaggacagacggggc
actctggcccagccgttggctcagtatctgcaggcgtccgtctcggga
cggagctcaggggaagagcgtgactccagttgaacgtgatagtcggtgcg
ttgagaggagaccagtcgggtgtcagtcagaagggggccggggccga
ggccctgggcaggacggccgtgcccgtcatcacgggcccagcgtcctag
aggcaggactctggtggagagtgtgagggtgctggggcccctccggagc
tggggccgtgcgggtgcaggttgggtctcggcgcggttgggtgttct
gcgggatttggaggaattcttccagtgatgggagtcgccagtaccgggc
accaggctggtgaagaggaggccgcccgtcgtggccagagcagctgggagg
gttcggtaaaaggctcggccgttctttaaagaggacttttctggagg
gcatttagctagtcgggaccgttctgactcgggaagagggtgctgggag
gagggcattgcccaggagccgaaggcgccgcggggagaagcccagggt
ctcctgtccccacagaggcgacgccactgccgcagacagacagggcctt
ccctctgatgacggcaagggcgctcgttcttgcgggggtgctggggggg
agtcgccccgaagccgtcaccagaggcctgaggggtgagactgaccga
tgcccttggccgggctggggcccggaccgagggggactccgtggaggca
gggcgatggtggtgcgggaggggaaccgacctggggccgagcccggctg
gcgattcccgggcgagggccctcagccgaggcgagtgggtccggcggaac
cacccttctggccagcggccacagggctctcgggactgtccggggcgacg
ctgggctgcccgtggcaggccTGGGCTGACCTGGACTTCACCAGACAGAA
CAGGGCTTTCAGGGCTGAGCTGAGCCAGGTTTAGCGAGGCCAAGTGGGGC
TGAACCAGGCTCAACTGGCCTGAGCTGGGTTGAGCTGGGCTGACCTGGGC
TGAGCTGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGAC
TGGCTGAGCTGAGCTGGGTTGAGCTGAGCTGAGCTGGCCTGGGTTGAGCT
GGGCTGGGTTGAGCTGAGCTGGGTTGAGCTGGGTTGAGCTGGGTTGATCT
GAGCTGAGCTGGGCTGAGCTGAGCTAGGCTGGGGTGAGCTGGGCTGAGCT
GGTTTGAGTTGGGTTGAGCTGAGCTGAGCTGGGCTGTGCTGGCTGAGCTA
GGCTGAGCTAGGCTAGGTTGAGCTGGGCTGGGCTGAGCTGAGCTAGGCTG
GGCTGATTGGGCTGAGCTGAGCTGAGCTAGGCTGCGTTGAGCTGGCTGG
GCTGGATTGAGCTGGCTGAGCTGGCTGAGCTGGGCTGAGCTGGCCTGGGT
TGAGCTGAGCTGGACTGGTTTGAGCTGGGTCGATCTGGGTTGAGCTGTCC
TGGGTTGAGCTGGGCTGGGTTGAGCTGAGCTGGGTTGAGCTGGGCTCAGC
AGAGCTGGGTTGGGCTGAGCTGGGTTGAGCTGAGCTGGGCTGAGCTGGCC
TGGGTTGAGCTGGGCTGAGCTGAGCTGGGCTGAGCTGGCCTGTGTTGAGC
TGGGCTGGGTTGAGCTGGGCTGAGCTGGATTGAGCTGGGTTGAGCTGAGC
TGGGCTGGGCTGTGCTGACTGAGCTGGGCTGAGCTAGGCTGGGGTGAGCT
GGGCTGAGCTGATCCGAGCTAGGCTGGGCTGGTTTGGGCTGAGCTGAGCT
GAGCTAGGCTGGATTGATCTGGCTGAGCTGGGTTGAGCTGAGCTGGGCTG
AGCTGGTCTGAGCTGGCCTGGGTCGAGCTGAGCTGGACTGGTTTGAGCTG
GGTCGATCTGGGCTGAGCTGGCCTGGGTTGAGCTGGGCTGGGTTGAGCTG
AGCTGGGTTGAGCTGGGCTGAGCTGAGGGCTGGGGTGAGCTGGGCTGAAC
TAGCCTAGCTAGGTTGGGCTGAGCTGGGCTGGTTTGGGCTGAGCTGAGCT
GAGCTAGGCTGCATTGAGCAGGCTGAGCTGGGCTGAGCAGGCCTGGGGTG
AGCTGGGCTAGGTGGAGCTGAGCTGGGTCGAGCTGAGTTGGGCTGAGCTG
GCCTGGGTTGAGGTAGGCTGAGCTGAGCTGAGCTAGGCTGGGTTGAGCTG
GCTGGGCTGGTTTGCCTGGGTCAAGCTGGGCCGAGCTGGCCTGGGTTGA
GCTGGGCTCGGTTGAGCTGGGCTGAGCTGAGCCGACCTAGGCTGGGATGA

GCTGGGCTGATTTGGGCTGAGCTGAGCTGAGCTAGGCTGCATTGAGCAGG
 CTGAGCTGGGCCTGGAGCCTGGCCTGGGGTGAGCTGGGCTGAGCTGCGCT
 GAGCTAGGCTGGGTTGAGCTGGCTGGGCTGGTTTGCGCTGGGTCAAGCTG
 GGCCGAGCTGGCCTGGGATGAGCTGGGCCGGTTTGGGCTGAGCTGAGCTG
 AGCTAGGCTGCATTGAGCAGGCTGAGCTGGGCTGAGCTGGCCTGGGGTGA
 GCTGGGCTGAGCTAAGCTGAGCTGGGCTGGTTTGGGCTGAGCTGGCTGAG
 CTGGGTCTCTGCTGAGCTGGGCTGAGCTGACCAGGGGTGAGCTGGGCTGAG
 TTAGGCTGGGCTCAGCTAGGCTGGGTTGATCTGGCAGGGCTGGTTTGCGC
 TGGGTCAAGCTCCCGGGAGATGGCCTGGGATGAGCTGGGCTGGTTTGGGC
 TGAGCTGAGCTGAGCTGAGCTAGGCTGCATTGAGCAGGCTGAGCTGGGCT
 GAGCTGGCCTGGGGTGAGCTGGGCTGGGTGGAGCTGAGCTGGGCTGAACT
 GGGCTAAGCTGGCTGAGCTGGATCGAGCTGAGCTGGGCTGAGCTGGCCTG
 GGGTTAGCTGGGCTGAGCTGAGCTGAGCTAGGCTGGGTTGAGCTGGCTGG
 GCTGGTTTGCGCTGGGTCAAGCTGGGCCGAGCTGGCCTGGGTTGAGCTGG
 GCTGGGCTGAGCTGAGCTAGGCTGGGTTGAGCTGGGCTGGGCTGAGCTGA
 GCTAGGCTGCATTGAGCTGGCTGGGATGGATTGAGCTGGCTGAGCTGGCT
 GAGCTGGCTGAGCTGGGCTGAGCTGGCCTGGGTTGAGCTGGGCTGGGTTG
 AGCTGAGCTGGGCTGAGCTGGGCTCAGCAGAGCTGGGTTGAGCTGAGCTG
 GGTGAGCTGGGGTGAGCTGGGCTGAGCAGAGCTGGGTTGAGCTGAGCTG
 GGTGAGCTGGGCTCGAGCAGAGCTGGGTTGAGCTGAGCTGGGTTGAGCT
 GGGCTCAGCAGAGCTGGGTTGAGCTGAGCTGGGTTGAGCTGGGCTGAGCT
 AGCTGGGCTCAGCTAGGCTGGGTTGAGCTGAGCTGGGCTGAACTGGGCTG
 AGCTGGGCTGAACTGGGCTGAGCTGGGCTGAGCTGGGCTGAGCAGAGCTG
 GGCTGAGCAGAGCTGGGTTGGTCTGAGCTGGGTTGAGCTGGGCTGAGCTG
 GGCTGAGCAGAGTTGGGTTGAGCTGAGCTGGGTTGAGCTGGGCTGAGCTA
 GGCTGGGTTGAGCTGGGTTGAGTTGGGCTGAGCTGGGCTGGGTTGAGCGG
 AGCTGGGCTGAACTGGGCTGAGCTGGGCTGAGCGGAACTGGGTTGATCTG
 AATTGAGCTGGGCTGAGCCGGGCTGAGCCGGGCTGAGCTGGGCTAGGTTG
 AGCTTGGGTGAGCTTGCCTCAGCTGGTCTGAGCTAGGTTGGGTGGAGCTA
 GGCTGGATTGAGCTGGGCTGAGCTGAGCTGATCTGGCCTCAGCTGGGCTG
 AGGTAGGCTGAACTGGGCTGTGCTGGGCTGAGCTGAGCTGAGCCAGTTTG
 AGCTGGGTTGAGCTGGGCTGAGCTGGGCTGTGTTGATCTTCTCCTGAACTG
 GGCTGAGCTGGGCTGAGCTGGCCTAGCTGGATTGAACGGGGGTAAGCTGG
 GCCAGGCTGGACTGGGCTGAGCTGAGCTAGGCTGAGCTGAGTTGAATTGG
 GTTAAGCTGGGCTGAGATGGGCTGAGCTGGGCTGAGCTGGGTTGAGCCAG
 GTCGGACTGGGTTACCCTGGGCCACACTGGGCTGAGCTGGGCCGAGCTCG
 attaacctggcaggctgagtcgggtccagcagacatgcgctggccaggc
 tggcttgacctggacacgttcgatgagctgccttgggatgggtcacctca
 gctgagccaggctggctccagctgggctgagctggtagacctgggtgacct
 cggtagaccaggctgtcctgagtcgggccaagccgaggctgcatcagact
 cgccagaccaagggctgggccccggctggcaagccaggggcggtgaagg
 ctgggctggcaggactgtcccggaaggaggtgcacgtggagccgcccga
 ccccgaccggcaggacctggaaagacgcctctcactcccccttctctct
 gtccctctcgggtcctcagAGAGCCAGTCTGCCCCGAATCTCTACCCCC
 TCGTCTCCTGCGTCAGCCCCCGTCCGATGAGAGCCTGGTGGCCCTGGGC
 TGCTGGCCCCGGGACTTCCTGCCCAGCTCCGTACCTTCTCCTGGAACCTA
 CAAGAACAGCAGCAAGGTCAGCAGCCAGAACATCCAGGACTTCCCGTCC
 G
 TCCTGAGAGGGCGGCAAGTACTTGGCCTCCTCCCGGGTGCTCCTACCCTCT
 GTGAGCATCCCCCAGGACCCAGAGGCCTTCCTGGTGTGCGAGGTCCAGCA
 CCCCAGTGGCACCAAGTCCGTGTCCATCTCTGGGCCAGgtgagctgggct
 cccctgtggctgtggcggggcgggcggggtgcccgccgacagtgac
 gccccgttcctgcctgcagTCGTAGAGGAGCAGCCCCCGTCTTGAACAT
 CTTTCGTCCCCACCCGGGAGTCTTCTCCAGTACTCCCCAGCGCACGTCCA
 AGCTCATCTGCCAGGCCTCAGACTTCAGCCCCAAGCAGATCTCCATGGCC
 TGGTTCCGTGATGGGAAACGGGTGGTGTCTGGCGTCAGCACAGGCCCCGT

	GGAGACCCTACAGTCCAGTCCGGTGACCTACAGGCTCCACAGCATGCTGA CCGTACAGGAGTCCGAGTGGCTCAGCCAGAGCGTCTTCACCTGCCAGGTG GAGCACAAGGGCTGAACACGAGAAGAACGCGTCCTCTCTGTGCACCTC CAgtgagtgcagccctcgggcccggcgggcgggcgggagccacaca cacaccagctgctccctgagccttggttcgggagtgccaaggcgggg aggggctgtgcagggcagctggaggcactgtcagctggggcccagcacc ccctaccccggcaggggcccgggctccgaggggccccgcagtcgcaggcc ctgctcttgggggaagccctacttgccccctcaggcgctgacgtccc cccacccacccccgctagATCCCAACTCTCCCATCACCGTCTTCGCCAT CGCCCCCTCCTTCGCTGGCATCTTCCTCACCAAGTCGGCCAAGCTTTCT GCCTGGTCACGGGCCTCGTCACCAGGGAGAGCCTCAACATCTCCTGGACC CGCCAGGACGGCGAGGTTCTGAAGACCAGTATCGTCTTCTCTGAGATCTA CGCCAACGGCACCTTCGGCGCCAGGGGCGAAGCCTCCGTCTGCGTGGAGG ACTGGGAGTCGGGCGACAGGTTACGTGCACGGTGACCCACACGGACCTG CCCTCGCCGCTGAAGCAGAGCGTCTCCAAGCCCAGAGtaggcccctgcc tgcccctgctccgcccggcctgtgccttgccgcccggggcgggagccga gcctggccgaggagcgccctcggcccccgggctcccgaccacacccct cctgctctctccccagGGATCGCCAGGCACATGCCGTCCGTGTACGTGC TGCCGCGCGCCCCGGAGGAGCTGAGCCTGCAGGAGTGGGCCTCGGTACCC TGCTGGTGAAGGGCTTCTCCCCGGCGGACGTGTTCTGTGAGTGGCTGCA GAAGGGGGAGCCCGTGTCCGCCGACAAGTACGTGACCAGCGCGCCGGTG C CCGAGCCCGAGCCCAAGGCCCCCGCCTCCTACTTCGTGCAGAGCGTCCTG ACGGTGAGCGCCAAGGACTGGAGCGACGGGGAGACCTACACCTGCGTCG T GGGCCACGAGGCCCTGCCCCACACGGTGACCGAGAGGACCGTGGACAAG T CCACCGGTAAACCCACCCTGTACAACGTCTCCCTGGTCCTGTCCGACACG GCCAGCACCTGCTACTGACCCCTGGCTGCCCGCCGCGGCCGGGGCCAGA GCCCCCGGGCGACCATCGCTCTGTGTGGGCCTGTGTGCAACCCGACCCTG TCGGGGTGAGCGGTTCGATTTCTGAAAATTAGAAataaaAGATCTCGTGC CG
Seq ID No.1	TCTAgAAGACGCTGGAGAGAGGGCCagACTTCCTCGGAACAGCTCAAAGAG CTCTGTCAAAGCCAGATCCCATCACACGTGGGCACCAATAGGCCATGCCA GCCTCCAAGGGCCGAACCTGGGTTCTCCACGGCGCACATGAAGCCTGCAGC CTGGCTTATCCTCTTCCGTGGTGAAGAGGCAGGCCCCGGGACTGGACGAGG GGCTAGCAGGGTGTGGTAGGCACCTTGCGCCCCCACCCTGGCAGGAACC AGAGACCCTGGGGCTGAGAGTGAGCCTCCAAACAGGATGCCCCACCCTTC AGGCCACCTTTCAATCCAGCTACACTCCACCTGCCATTCTCCTCTGGGCA CAGGGCCCAGCCCCTGGATCTTGGCCTTGGCTCGACTTGACCCACGCGC ACACACACACTTCCTAACGTGCTGTCCGCTCACCCCTCCCCAGCGTGGTC CATGGGCAGCACGGCAGTGCGCGTCCGGCGGTAGTGAGTGACAGAGGTCCC TTCCCCTCCCCAGGAGCCCCAGGGGTGTGTGACAGATCTGGGGGCTCCTG TCCCTTACACCTTCATGCCCTCCCCTCATACCCACCCTCCAGGCGGGAG GCAGCGAGACCTTTGCCAGGGACTCAGCCAACGGGCACACGGGAGGCC A GCCCTCAGCAGCTGGG
Seq ID No.4	GGCCAGACTTCCTCGGAACAGCTCAAAGAGCTCTGTCAAAGCCAGATCCC ATCACACGTGGGCACCAATAGGCCATGCCAGCCTCCAAGGGCCGAACCTGG GTTCTCCACGGCGCACATGAAGCCTGCAGCCTGGCTTATCCTCTTCCGTG GTGAAGAGGCAGGCCCCGGGACTGGACGAGGGGCTAGCAGGGTGTGGTAG GCACCTTGCGCCCCCACCCTGGCAGGAACCAGAGACCCTGGGGCTGAGA G

TGAGCCTCCAAACAGGATGCCCCACCCTTCAGGCCACCTTTCAATCCAGC
TACACTCCACCTGCCATTCTCCTCTGGGCACAGGGCCCAGCCCCTGGATC
TTGGCCTTGGCTCGACTTGCACCACGCGCACACACACTTCCTAACGT
GCTGTCCGCTCACCCCTCCCCAGCGTGGTCCATGGGCAGCACGGCAGTGC
GCGTCCGGCGGTAGTGAGTGACAGAGGTCCCTTCCCCTCCCCCAGGAGCCC
CAGGGGTGTGTGCAGATCTGGGGGCTCCTGTCCCTTACACCTTCATGCCC
CTCCCCTCATACCCACCCTCCAGGCGGGAGGCAGCGAGACCTTTGCCCAG
GGACTCAGCCAACGGGCACACGGGAGGCCAGCCCTCAGCAGCTGGCTCCC
AAAGAGGAGGTGGGAGGTAGGTCCACAGCTGCCACAGAGAGAAACCCTG
ACGGACCCACAGGGGCCACGCCAGCCGGAACCAGCTCCCTCGTGGGTGA
GCAATGGCCAGGGCCCCCGCCGGCCACCAGGCTGGCCTTGCGCCAGCTGA
G
AACTCACGTCCAGTGCAGGGAGACTCAAGACAGCCTGTGCACACAGCCTC
GGATCTGCTCCCATTTCAAGCAGAAAAAGGAAACCGTGCAGGCAGCCCTC
AGCATTTCAAGGATTGTAGCAGCGGCCAACTATTCTGTCGGCAGTGGCCGA
TTAGAATGACCGTGGAGAAGGGCGGAAGGGTCTCTCGTGGGCTCTGCGGC
CAACAGGCCCTGGCTCCACCTGCCCCGCTGCCAGCCCGAGGGGCTTGGGCC
GAGCCAGGAACCACAGTGCTCACCGGGACCACAGTGACTGACCAAATCC
CGGCCAGAGCAGCCCCAGGCCAGCCGGGCTCTCGCCCTGGAGGACTCACC
ATCAGATGCACAAGGGGGCGAGTGTGGAAGAGACGTGTCGCCCCGGGCCA
T
TTGGGAAGGCGAAGGGACCTTCCAGGTGGACAGGAGGTGGGACGCACTC
C
AGGCAAGGGACTGGGTCCCCAAGGCCTGGGGAAGGGGTACTGGCTTGGG
G
GTTAGCCTGGCCAGGGAACGGGGAGCGGGGCGGGGGGCTGAGCAGGGAG
G
ACCTGACCTCGTGGGAGCGAGGCAAGTCAGGCTTCAGGCAGCAGCCGCAC
ATCCCAGACCAGGAGGCTGAGGCAGGAGGGGCTTGCAGCGGGGCGGGGG
C
CTGCCTGGCTCCGGGGGCTCCTGGGGGACGCTGGCTCTTGTTCCTGTGC
CCGCAGCACAGGGCCAGCTCGCTGGGCCTATGCTTACCTTGATGTCTGGG
GCCGGGGCGTCAAGGTCTGCTCCTCAGGGGAGAGTCCCCTGAGGCTA
CGCTGGGG*GGGGACTATGGCAGCTCCACCAGGGGCCTGGGGACCAGGG
G
CCTGGACCAGGCTGCAGCCCGGAGGACGGGCAGGGCTCTGGCTCTCCAGC
ATCTGGCCCTCGGAAATGGCAGAACCCTGGCGGGTGAGCGAGCTGAGA
G
CGGGTCAGACAGACAGGGGGCCGGCCGGAAGGAGAAGTTGGGGGCAGAG
C
CCGCCAGGGGCCAGGCCCAAGGTTCTGTGTGCCAGGGCCTGGGTGGGCAC
ATTGGTGTGGCCATGGCTACTTAGATTCGTGGGGCCAGGGCATCCTGGTC
ACCGTCTCCTCAGGTGAGCCTGGTGTCTGATGTCCAGCTAGGCGCTGGTG
GGCCGCGGGTGGGCCTGTCTCAGGCTAGGGCAGGGGCTGGGATGTGTATT
TGTCAGGAGGGGCAACAGGGTGCAGACTGTGCCCTGGAACTTGACCA
CTGGGGCAGGGGCGTCCTGGTACGTCTCCTCAGGTAAGACGGCCCTGTG
CCCCTCTCTCGCGGGACTGGAAAAGGAATTTTCCAAGATTCTTGGTCTG
TGTGGGGCCCTCTGGGGCCCCCGGGGGTGGCTCCCCTCCTGCCCAGATGG
GGCCTCGGCCTGTGGAGCACGGGCTGGGCACACAGCTCGAGTCTAGGGCC
ACAGAGGCCCGGGCTCAGGGCTCTGTGTGGCCCGGCGACTGGCAGGGGG
C
TCGGGTTTTTGGACACCCCCTAATGGGGGCCACAGCACTGTGACCATCTT
CACAGCTGGGGCCGAGGAGTCGAGGTACCGTCTCCTCAGGTGAGTCCTC
GTCAGCCCTCTCTACTCTCTGGGGGGTTTTGCTGCATTTTGTGGGGGAA
AGAGGATGCCTGGGTCTCAGGTCTAAAGGTCTAGGGCCAGCGCCGGGGCC
CAGGAAGGGGCCGAGGGGCCAGGCTCGGCTCGGCCAGGAGCAGAGCTTC

C
AGACATCTCGCCTCCTGGCGGCTGCAGTCAGGCCTTTGGCCGGGGGGGTC
TCAGCACCAACAGGCCTCTTGGCTCCCGAGGTCCCCGGCCCCGGCTGCCT
CACCAGGCACCGTGCGCGGTGGGCCCCGGGCTCTTGGTCGGCCACCCTTTC
TTAACTGGGATCCGGGCTTAGTTGTGCAATGTGACAACGGGCTCGAAAG
CTGGGGCCAGGGGACCCTAGT*TACGACGCCTCGGGTGGGTGTCCCGCAC
CCCTCCCCACTTTCACGGCACTCGGCGAGACCTGGGGAGTCAGGTGTTGG
GGACACTTTGGAGGTCAGGAACGGGAGCTGGGGAGAGGGCTCTGTCAGC
G
GGGTCCAGAGATGGGCCGCCCTCCAAGGACGCCCTGCGCGGGGACAAGG
G
CTTCTTGGCCTGGCCTGGCCGCTTCACTTGGGCGTCAGGGGGGGCTTCCC
GGGGCAGGCGGTCAGTGAGGCGGGTTGGAATTCTGAGTCTGGGTTCGGG
GTTTCGGGGTTCGGCCTTCATGAACAGACAGCCCAGGCGGGCCGTTGTTG
GCCCCTGGGGGCCTGGTTGGAATGCGAGGTCTCGGGAAGTCAGGAGGGA
G
CCTGGCCAGCAGAGGGTTCCAGCCCTGCGGCCGAGGGACCTGGAGACG
G
GCAGGGCATTGGCCGTCGCAAGGGCCAGGCCACACCCCCCAGGTTTTTGTG
GGGCGAGCCTGGAGATTGCACCACTGTGATTACTATGCTATGGATCTCTG
GGGCCAGGCGTTGAAGTCGTGCTGTCCTCAGGTAAGAACGGCCCTCCAG
GGCCTTTAATTTCTGCTCTCGTCTGTGGGCTTTTCTGACTCTGATCCTCG
GGAGGCGTCTGTGCCCCCCCCGGGGATGAGGCCGGCTTGCCAGGAGGGGT
CAGGGACCAGGAGCCTGTGGGAAGTTCTGACGGGGGCTGCAGGCGGGAA
G
GGCCCCACCGGGGGGCGAGCCCCAGGCCGCTGGGCGGCAGGAGACCCGT
G
AGAGTGCGCCTTGAGGAGGGTGTCTGCGGAACCACGAACGCCCGCCGGG
A
AGGGCTTGCTGCAATGCGGTCTTCAGACGGGAGGCGTCTTCTGCCCTCAC
CGTCTTTCAAGCCCTTGTGGGTCTGAAAGAGCCATGTTCGGAGAGAGAAGG
GACAGGCCTGTCCCGACCTGGCCGAGAGCGGGCAGCCCCGGGGGAGAGC
G
GGGCGATCGGCCTGGGCTCTGTGAGGCCAGGTCCAAGGGAGGACGTGTG
G
TCCTCGTGACAGGTGCACTTGCGAAACCTTAGAAGACGGGGTATGTTGGA
AGCGGCTCCTGATGTTTAAAGAAAAGGGAGACTGTAAAGTGAGCAGAGTCC
TCAAGTGTGTTAAGGTTTTAAAGGTCAAAGTGTTTTAAACCTTTGTGACT
GCAGTTAGCAAGCGTGCGGGGAGTGAATGGGGTGCCAGGGTGGCCGAGA
G
GCAGTACGAGGGCCGTGCCGTCCTCTAATTCAGGGCTTAGTTTTGCAGAA
TAAAGTCGGCCTGTTTTCTAAAAGCATTGGTGGTGCTGAGCTGGTGGAGG
AGGCCGCGGGCAGCCCTGGCCACCTGCAGCAGGTGGCAGGAAGCAGGTC
G
GCCAAGAGGCTATTTTAGGAAGCCAGAAAACACGGTCGATGAATTTATAG
CTTCTGGTTTCCAGGAGGTGGTTGGGCATGGCTTTGCGCAGCGCCACAGA
ACCGAAAGTGCCCACTGAGAAAAACAACCTCCTGCTTAATTTGCATTTTT
CTAAAAGAAGAAACAGAGGCTGACGGAAACTGGAAAGTTCCTGTTTTAAC
TACTCGAATTGAGTTTTCGGTCTTAGCTTATCAACTGCTCACTTAGATTC
ATTTTCAAAGTAAACGTTTAAAGAGCCGAGGCATTCCTATCCTCTTCTAAG
GCGTTATTCCTGGAGGCTCATTCACCGCCAGCACCTCCGCTGCCTGCAGG
CATTGCTGTACCGTACCGTGACGGCGCGCACGATTTTCAGTTGGCCCG
CTTCCCCTCGTGATTAGGACAGACGCGGGCACTCTGGCCCAGCCGTCTTG
GCTCAGTATCTGCAGGCGTCCGTCTCGGGACGGAGCTCAGGGGAAGAGCG
TGACTCCAGTTGAACGTGATAGTCGGTGCGTTGAGAGGAGACCCAGTCGG
GTGTGAGTCAGAAAGGGGCCCGGGGCCCGAGGCCCTGGGCAGGACGGCC

C
GTGCCCTGCATCACGGGCCCAGCGTCCTAGAGGCAGGACTCTGGTGGAGA
GTGTGAGGGTGCCTGGGGCCCCCTCCGGAGCTGGGGCCGTGCGGTGCAGGT
TGGGCTCTCGGCGCGGTGTTGGCTGTTTCTGCGGGATTTGGAGGAATTCT
TCCAGTGATGGGAGTCGCCAGTGACCGGGCACCAGGCTGGTAAGAGGGA
G
GCCGCCGTCTGTGGCCAGAGCAGCTGGGAGGGTTTCGGTAAAAGGCTCGCCC
GTTTCCTTTAATGAGGACTTTTCCTGGAGGGCATTAGTCTAGTCGGGAC
CGTTTTCGACTCGGGAAGAGGGATGCGGAGGAGGGCATGTGCCCAGGAG
C
CGAAGGCGCCGCGGGGAGAAGCCCAGGGCTCTCCTGTCCCCACAGAGGC
G
ACGCCACTGCCGCAGACAGACAGGGCCTTTCCCTCTGATGACGGCAAAGG
CGCCTCGGCTCTTGCGGGGTGCTGGGGGGGAGTCGCCCCGAAGCCGCTCA
CCCAGAGGCCTGAGGGGTGAGACTGACCGATGCCTCTTGGCCGGGCCTGG
GGCCGACCGAGGGGGACTCCGTGGAGGCAGGGCGATGGTGGCTGCGGG
A
GGGAACCGACCCTGGGCCGAGCCCGGCTTGGCGATTCCCGGGCGAGGGCC
CTCAGCCGAGGCGAGTGGGTCCGGCGGAACCACCCTTTCTGGCCAGCGCC
ACAGGGCTCTCGGGACTGTCCGGGGCGACGCTGGGCTGCCCGTGGCAGGC
CTGGGCTGACCTGGACTTCACCAGACAGAACAGGGCTTTCAGGGCTGAGC
TGAGCCAGGTTTAGCGAGGCCAAGTGGGGCTGAACCAGGCTCAACTGGCC
TGAGCTGGGTTGAGCTGGGCTGACCTGGGCTGAGCTGAGCTGGGCTGGGC
TGGGCTGGGCTGGGCTGGGCTGGGCTGGACTGGCTGAGCTGAGCTGGGTT
GAGCTGAGCTGAGCTGGCCTGGGTTGAGCTGGGCTGGGTTGAGCTGAGCT
GGGTTGAGCTGGGTTGAGCTGGGTTGATCTGAGCTGAGCTGGGCTGAGCT
GAGCTAGGCTGGGGTGAGCTGGGCTGAGCTGGTTTGAGTTGGGTTGAGCT
GAGCTGAGCTGGGCTGTGCTGGCTGAGCTAGGCTGAGCTAGGCTAGGTTG
AGCTGGGCTGGGCTGAGCTGAGCTAGGCTGGGCTGATTGGGCTGAGCTG
AGCTGAGCTAGGCTGCGTTGAGCTGGCTGGGCTGGATTGAGCTGGCTGAG
CTGGCTGAGCTGGGCTGAGCTGGCCTGGGTTGAGCTGAGCTGGACTGGTT
TGAGCTGGGTCGATCTGGGTTGAGCTGTCCTGGGTTGAGCTGGGCTGGGT
TGAGCTGAGCTGGGTTGAGCTGGGCTCAGCAGAGCTGGGTTGGGCTGAGC
TGGGTTGAGCTGAGCTGGGCTGAGCTGGCCTGGGTTGAGCTGGGCTGAGC
TGAGCTGGGCTGAGCTGGCCTGTGTTGAGCTGGGCTGGGTTGAGCTGGGC
TGAGCTGGATTGAGCTGGGTTGAGCTGAGCTGGGCTGGGCTGTGCTGACT
GAGCTGGGCTGAGCTAGGCTGGGGTGAGCTGGGCTGAGCTGATCCGAGCT
AGGCTGGGCTGGTTTGGGCTGAGCTGAGCTGAGCTAGGCTGGATTGATCT
GGCTGAGCTGGGTTGAGCTGAGCTGGGCTGAGCTGGTCTGAGCTGGCCTG
GGTCGAGCTGAGCTGGACTGGTTTGAAGCTGGGTCGATCTGGGCTGAGCTG
GCCTGGGTTGAGCTGGGCTGGGTTGAGCTGAGCTGGGTTGAGCTGGGCTG
AGCTGAGGGCTGGGGTGAGCTGGGCTGAACTAGCCTAGCTAGGTTGGGCT
GAGCTGGGCTGGTTTGGGCTGAGCTGAGCTGAGCTAGGCTGCATTGAGCA
GGCTGAGCTGGGCTGAGCAGGCCTGGGGTGAGCTGGGCTAGGTGGAGCT
G
AGCTGGGTCGAGCTGAGTTGGGCTGAGCTGGCCTGGGTTGAGGTAGGCTG
AGCTGAGCTGAGCTAGGCTGGGTTGAGCTGGCTGGGCTGGTTTGCCTGG
GTCAAGCTGGGCCGAGCTGGCCTGGGTTGAGCTGGGCTCGGTTGAGCTGG
GCTGAGCTGAGCCGACCTAGGCTGGGATGAGCTGGGCTGATTTGGGCTGA
GCTGAGCTGAGCTAGGCTGCATTGAGCAGGCTGAGCTGGGCCTGGAGCCT
GGCCTGGGGTGAGCTGGGCTGAGCTGCGCTGAGCTAGGCTGGGTTGAGCT
GGCTGGGCTGGTTTGCCTGGGTTCAAGCTGGGCCGAGCTGGCCTGGGATG
AGCTGGGCCGGTTTGGGCTGAGCTGAGCTGAGCTAGGCTGCATTGAGCAG
GCTGAGCTGGGCTGAGCTGGCCTGGGGTGAGCTGGGCTGAGCTAAGCTGA
GCTGGGCTGGTTTGGGCTGAGCTGGCTGAGCTGGGTCCTGCTGAGCTGGG
CTGAGCTGACCAGGGGTGAGCTGGGCTGAGTTAGGCTGGGCTCAGCTAGG

	CTGGGTTGATCTGGCAGGGCTGGTTTGGCGCTGGGTCAAGCTCCCGGGAGA TGGCCTGGGATGAGCTGGGCTGGTTTGGGCTGAGCTGAGCTGAGCTGAGC TAGGCTGCATTGAGCAGGCTGAGCTGGGCTGAGCTGGCCTGGGGTGAGCT GGGCTGGGTGGAGCTGAGCTGGGCTGAACTGGGCTAAGCTGGCTGAGCTG GATCGAGCTGAGCTGGGCTGAGCTGGCCTGGGGTTAGCTGGGCTGAGCTG AGCTGAGCTAGGCTGGGTTGAGCTGGCTGGGCTGGTTTGGCGCTGGGTCAA GCTGGGCCGAGCTGGCCTGGGTTGAGCTGGGCTGGGCTGAGCTGAGCTAG GCTGGGTTGAGCTGGGCTGGGCTGAGCTGAGCTAGGCTGCATTGAGCTGG CTGGGATGGATTGAGCTGGCTGAGCTGGCTGAGCTGGCTGAGCTGGGCTG AGCTGGCCTGGGTTGAGCTGGGCTGGGTTGAGCTGAGCTGGGCTGAGCTG GGCTCAGCAGAGCTGGGTTGAGCTGAGCTGGGTTGAGCTGGGGTGAGCTG GGCTGAGCAGAGCTGGGTTGAGCTGAGCTGGGTTGAGCTGGGCTCGAGCA GAGCTGGGTTGAGCTGAGCTGGGTTGAGCTGGGCTCAGCAGAGCTGGGTT GAGCTGAGCTGGGTTGAGCTGGGCTGAGCTAGCTGGGCTCAGCTAGGCTG GGTTGAGCTGAGCTGGGCTGAACTGGGCTGAGCTGGGCTGAACTGGGCTG AGCTGGGCTGAGCTGGGCTGAGCAGAGCTGGGCTGAGCAGAGCTGGGTT G GTCTGAGCTGGGTTGAGCTGGGCTGAGCTGGGCTGAGCAGAGTTGGGTTG AGCTGAGCTGGGTTGAGCTGGGCTGAGCTAGGCTGGGTTGAGCTGGGTTG AGTTGGGCTGAGCTGGGCTGGGTTGAGCGGAGCTGGGCTGAACTGGGCTG AGCTGGGCTGAGCGGAACTGGGTTGATCTGAATTGAGCTGGGCTGAGCCG GGCTGAGCCGGGCTGAGCTGGGCTAGGTTGAGCTTGGGTGAGCTTGCCTC AGCTGGTCTGAGCTAGGTTGGGTGGAGCTAGGCTGGATTGAGCTGGGCTG AGCTGAGCTGATCTGGCCTCAGCTGGGCTGAGGTAGGCTGAACTGGGCTG TGCTGGGCTGAGCTGAGCTGAGCCAGTTTGAAGCTGGGTTGAGCTGGGCTG AGCTGGGCTGTGTTGATCTTTCCTGAACTGGGCTGAGCTGGGCTGAGCTG GCCTAGCTGGATTGAACGGGGGTAAGCTGGGCCAGGCTGGACTGGGCTGA GCTGAGCTAGGCTGAGCTGAGTTGAATTGGGTTAAGCTGGGCTGAGATGG GCTGAGCTGGGCTGAGCTGGGTTGAGCCAGGTCGGACTGGGTTACCTGG GCCACACTGGGCTGAGCTGGGCGGAGCTCGATTAACCTGGTCAGGCTGAG TCGGGTCCAGCAGACATGCGCTGGCCAGGCTGGCTTGACCTGGACACGTT CGATGAGCTGCCTTGGGATGGTTACCTCAGCTGAGCCAGGTGGCTCCAG CTGGGCTGAGCTGGTGACCCTGGGTGACCTCGGTGACCAGGTTGTCCTGA GTCCGGGCCAAGCCGAGGCTGCATCAGACTCGCCAGACCCAAGGCCTGGG CCCCGGCTGGCAAGCCAGGGGCGGTGAAGGCTGGGCTGGCAGGACTGTC CCGGAAGGAGGTGCACGTGGAGCCGCCCGGACCCCGACCGGCAGGACCT GGAAAGACGCCTCTCACTCCCCCTTCTCTTCTGTCCCCTCTCGGGTCCTCA GAGAGCCAGTCTGCCCCGAATCTCTACCCCTCGTCTCCTGCGTCAGCCCC CCGTCCGATGAGAGCCTGGTGGCCCTGGGCTGCCTGGCCCGGGACTTCCT GCCCAGCTCCGTCACCTTCTCCTGGAA
--	---

Porcine Kappa Light Chain

In another embodiment, novel genomic sequences encoding the kappa light chain locus of ungulate immunoglobulin are provided. The present invention provides the first reported genomic sequence of ungulate kappa light chain regions. In one embodiment, nucleic acid sequence is provided that encodes the porcine kappa light chain locus. In another embodiment, the nucleic acid sequence can contain at least one joining region, one constant region and/or one enhancer region of kappa light chain. In a further embodiment, the nucleotide sequence can

include at least five joining regions, one constant region and one enhancer region, for example, as represented in Seq ID No. 30. In a further embodiment, an isolated nucleotide sequence is provided that contains at least one, at least two, at least three, at least four or five joining regions and 3' flanking sequence to the joining region of porcine genomic kappa light chain, for example, as represented in Seq ID No 12. In another embodiment, an isolated nucleotide sequence of porcine genomic kappa light chain is provided that contains 5' flanking sequence to the first joining region, for example, as represented in Seq ID No. 25. In a further embodiment, an isolated nucleotide sequence is provided that contains 3' flanking sequence to the constant region and, optionally, the 5' portion of the enhancer region, of porcine genomic kappa light chain, for example, as represented in Seq ID Nos. 15, 16 and/or 19.

In further embodiments, isolated nucleotide sequences as depicted in Seq ID Nos 30, 12, 25, 15, 16 or 19 are provided. Nucleic acid sequences at least 80, 85, 90, 95, 98 or 99% homologous to Seq ID Nos 30, 12, 25, 15, 16 or 19 are also provided. In addition, nucleotide sequences that contain at least 10, 15, 17, 20, 25 or 30 contiguous nucleotides of Seq ID Nos 30, 12, 25, 15, 16 or 19 are provided. In addition, nucleotide sequences that contain at least 10, 15, 17, 20, 25 or 30 contiguous nucleotides of Seq ID Nos 1, 4 or 29 are provided. In other embodiments, nucleotide sequences that contain at least 50, 100, 1,000, 2,500, 5,000, 7,000, 8,000, 8,500, 9,000, 10,000 or 15,000 contiguous nucleotides of Seq ID No. 30 are provided. Further provided are nucleotide sequences that hybridizes, optionally under stringent conditions, to Seq ID Nos 30, 12, 25, 15, 16 or 19, as well as, nucleotides homologous thereto.

In one embodiment, an isolated nucleotide sequence encoding kappa light chain is provided that includes at least five joining regions, one constant region and one enhancer region, for example, as represented in Seq ID No. 30. In Seq ID No. 30, the coding region of kappa light chain is represented, for example by residues 1-549 and 10026-10549, whereas the intronic sequence is represented, for example, by residues 550-10025, the Joining region of kappa light chain is represented, for example, by residues 5822- 7207 (for example, J1:5822-5859, J2:6180-6218, J3:6486-6523, J4:6826-6863, J5:7170-7207), the Constant Region is represented by the following residues: 10026- 10549 (C exon) and 10026-10354 (C coding), 10524-10529 (Poly(A) signal) and 11160-11264 (SINE element).

Seq ID No 30	GCGTCCGAAGTCAAAAATATCTGCAGCCTTCATGTATTTCATAGAAACAAG GAATGTCTACATTTTCCAAAGTGGGACCAGAATCTTGGGTCATGTCTAAG GCATGTGCATTTGCACATGGTAGGCAAAGGACTTTGCTTCTCCCAGCACA TCTTTCTGCAGAGATCCATGGAAACAAGACTCAACTCCAAAGCAGCAAAG AAGCAGCAAGTTCTCAAGTGATCTCCTCTGACTCCCTCCTCCCAGGCTAA TGAAGCCATGTTGCCCCTGGGGGATTAAGGGCAGGTGTCCATTGTGGCAC CCAGCCCGAAGACAAGCAATTTGATCAGGTTCTGAGCACTCCTGAATGTG GACTCTGGAATTTTCTCCTCACCTTGTGGCATATCAGCTTAAGTCAAGTA CAAGTGACAAACAACATAATCCTAAGAAGAGAGGAATCAAGCTGAAGTC A AAGGATCACTGCCTTGGATTCTACTGTGAATGATGACCTGGAAAATATCC TGAACAACAGCTTCAGGGTGATCATCAGAGACAAAAGTTCCAGAGCCAGg tagggaaacctcaagccttgcaaagagcaaaatcatgccattgggtct taacctgctgagtattactatatgttactgtggaggcaaaagcgtca aatagcctgggtaagtatgtcaataaaaagcaaaagtgggtttctga aatgttagacctgaggaaggaatattgataactaccaataatttcaga atgattatagatgtgacttagtcagtgtctccacccgcacctgac aagcagtttagaatttattctaagaatctaggttctgggggctacatg ggaatcagcttcagtgaagagttgttggaatgattcactaaatttcta ttccagcataaatccaagaacctctcagactagttattgacactgctt ttctccataatccatctcatctccgtccatcatggacactttagaat gacaggtcctggcagagactcacagatgcttctgaaacatccttgcctt caaagaatgaacagcacacatactaaggatctcagtcatccacaaattag ttttgccacaatggttcttatgataaaagctttcattaacagcaaatt gtttataatagttgttctgctttataataattgcatgcttcaatttctt ttctttctttttttttttttgcttttagtgccgcagggtgcagca tatgaaattcccaggctaggggtcaaatcagaactacacctaactggcct acgccacagccacagcaactcaggatctaagccatgtcggtgacctacac tacagctcatggcaatgccagatccttaaccaatgagcgaggccaggga tcgaacctatgctctcatggatactagtcaggctcattatccgctgagcc ataacaggaactcccaggttgccttttatcaaaattggtacagccttat tgttctgaaaaccacaaaatgaatgtattcacataattttaaaaggta aataatttatgataacaagacaatagaaagagaaaacgtcattgcctct ttctccacgacaacagcctccttaattgattgaagaaataactactg agcatggtttagtgacttcttcagcaattagcctgtattcatagccat acatttcaattaaaatgagatcatgatcacacaatacataccataca gcctatagggattttacaatcatctccacatgactacataaaaaccta cctaaaaaaaaaaaaaacctacttcatcctctattggctgcttgtgc tccattaaaaagctctatcataattaggttatgatgaggatttcatttt ctaccttcaagcaacattcaatgcacagtcttatatacacatttgagc ctaatttcttttcttttttttgggttttttttttttttttttt ggctttttgtcttttaaggctgcatatggagggtcccaggctagctg tctaatacagaactatagctgctggcctacgccacatccacagcaatacaa gatctgagccatgtctgcaacttacaccacagctcacagcaacggtggat ccttaaacactgagcaaggccagggatcaaaccataactcatggctc ctagttggatttgaaccactgagccatgatggcaactcctgagcctac tttttaatacattccaacctaggacactttttaagtttcattttct ccccccacccctgtttctgaagtgtgttcttccactgggtgacttc actcccaggatctcatctgcaggatactgcagctaagtgtatgagctctg aatttgaatcccaactctgccactcaaagggataggagttccgatgtgg cccaatgggatcagtggtcatctctgcagtgcaggacgcaggttccatcc ctggcccagcacagtgggttaagaatctggcattgctgcagctgaggcat agatttcaattgtcctcagatctgatccttggcccaaggactgcatatg cctcagggcaacaaaaaagagaaaaggggggtgatagcattagttcta gatttgggggataaataaaagtatccatgtacaatgtatggcattt tgtaaatgctcaaaaattcaactattatggagtcccatcatggctca
--------------	--

gtggaagggaatctgattagcatccatgaggacacaggtccaaccccgac
 ctgtctcagtgggcattgctgtgagctgtggcatgggttacagacgaagc
 tcggatctggcattgctgtggctgtggtgaagccagcaactacagctct
 cattcagcccctagcctgggaacctccatagcctaaaagacaaaaata
 aaatttaataaaaaataaagaatgttaactattatgattggtactgct
 tgcattactgcaaagaaagtcactttctatactctttaatacttagtg
 actgtgtgctcagtgaaactatttggacacttaattccactctctcta
 tctccaacttgacaactctcttctctctcttggtgagatccactgctg
 actttgctcttaaggcaactagaaaagtgtcagtgacaaaatcaaaga
 aagttacctaatacttcagaattacaatacttaagtctctgtaaagct
 actatttcagtggttagtattattccttggtcccttacaactatcagct
 ctgactattgctgatttcaactatttattgttgaggtttttctttt
 ttccctgttctctgcaaatgttctgagcattgtcaagtgaagata
 ctggactgggccttccaaatataagacaatgaacatcgggagttctcat
 tatggtgcagcagaaacgaatccaactaggaatgtgaggttcaggttc
 gatccctgcccttgcagtggttaaggatccagcattaccgtgagctg
 tgggttaggttcagacgtggctcagatcctgcgttgctgtggctgtggc
 ataggctggcagctctagctctgattcgaccgtagcctgggaacctcca
 tgcgccccgagtgagcccttaaaaagcaaaaaaaaaaagaaagaagaaa
 aagacaatgaaacatcaaacagctaacaatccagtagggtagaagaatc
 tggcaacagataagagcgattaaatgttctaggtccagtgacctgcctc
 tgtgctctacacagtcgtgccacttgctgagggaaggtctctcttgag
 ttgagtcctgaaagacattagttgtcacaactaatgccagtgaagtga
 ggtgtttccaagcagagggagagtttggtaaaaagctggaagtcacagaa
 agactctaaagagtttaggatggtgggagcaacatacgctgagatggggc
 tggaaaggttaagagggaaacaactatagtaagtgaagctggactcacagc
 aaagtgaggacctcagcctcctgatggggttaccatggaacaccaagg
 cacaccttgatttccaaaacagcaggcacctgattcagcccaatgtgaca
 tgggtgggtacccctctagctctacctgttctgtgacaactgacaaccaac
 gaagttaagtcgtgatttctactctgctgacctgtttttgttcaca
 cgtcatctatagcttcatgcaaaaatagagtcaaggttaagacgcgggcc
 ttggttgatatacatgtagctatcttgtttgagacaatatggtggcaa
 ggaagaggttcaaacaggaaaatactctctaattatgattaactgagaaa
 agctaaagagctccataatgacactgaatgaagttcatcatttgcaaaag
 cttccccccccccaggagactataaaaaagtgaatttttaaatgaa
 ctattttacaaaacagaaatagactcacagacataggaaacgaacagatg
 gttaccaagggtgaaagggagtaggagggataaataaggagtctggggtt
 agcagatacaccccagtgacacaaaataaacaacagggacctactatat
 agcacaggggaactatagcagtagcttacaataacctataatggaaaaga
 atgtgaaaaagaatatatgtatgcgtgtgtgtgaactgaatcactttgc
 tgaacctgaatctaacaataacattgaaatcaactacagtttttttt
 ttttaagtgcagggtttgtgtttttttttttcattttgtttgtt
 ttgtttttgttttagggccacacccagacatatgggggttcccagg
 ctagggttctaatagagctacagttgccggttgaccacagccacagc
 aacatcagatccgagccgcacttgcgacttacaccacagctcatggcaat
 accagatccttaacctgagcaaggccagggatcgtacccgcaacct
 catggttccatgtagcatttctgctgcgtacaatgggaactcaa
 gtgcagtttttgaatgtgcttcttctttgttaattcatattcatcc
 tacttcccaataaataaataacataaataaataacataccattgtaa
 atcaactacaatttttttaaatgcagggtttttgtttttgtttttgt
 ttgtcttttgccttttctagggccgctcccatggcatatggaggttcc
 caggctaggggtcgaatcgagctgtagccaccggcctacgccagagcca
 cagcaacgcgggatccgagccgctgtgcaacctacaccacagctcacgg
 caacgccgatcgttaacctgagcaagggcaggatcgaacctgcaa
 cctcatggttctagtcagattcgttaactactgagccacaacggaaact
 cctaaagtgcagtttttaaatgtcttcttctttgttaattacactc

aacctacttccaataaataaataaataaacaataatcatagacatgg
ttgaattctaaggaaggaccatcaggccttagacagaaatacgtcatc
ttctagtattttaaacacactaaagaagacaaacatgctctgccagaga
agcccagggcctccacagctgcttgcaaaggagtaggcttcagtagct
gacccaaggctctgttctcttcagggaagggtttttgttcagtga
cagcagacagctgtcactgtGTGGACGTTTCGGCCAAGGAACCAAGCTGG
AACTCAAACgtaagtcaatccaaacgttcttcttggtgtctgtgtct
tacggtctctgtggtctgaaatgattcatgtgtgactctctgaaacca
gactgacattctccagggaactaaagcctgtcatcaactggaaaaac
tgagggcacatttttgggcagaactaagagtcaggcactgggtgaggaa
aaactgttagaatgtagtttcagaaacttactgggaagcaaagcccat
gttctgaacagagctctgtcaagggtcaggagggaaccagttttgt
caggagggaagttagacgaacccctgtTATATGGTTTCGGCGCGGGGA
CCAAGCTGGAGCTCAAACgtaagtggcttttccgactgattcttctg
tttctaattgttggtggttttctcattttcaggttttcatcgaa
ttagtgtcagggaccaaacaattgccttccagattaggtaccaggga
ggggacattgtgcatgggagaccagagggtggctaattttaacgttc
caagccaaaataactggggaagggggcttctgtcctgtgagggtaggt
ttatagaagtgaagttaagggaatcgctatGTTCACTTTTGGCTC
GGGGACCAAAGTGGAGCCCCAAAAtgagtacatttccatcaattattg
tgagattttgtcctgtgtgtcatttgcagttttgacatttgg
tgaatgagccattccagggaaccaaaggatgagaccgaaaagtagaaa
agagccaacttttaagctgagcagacagaccgaattgtgagttgtgag
gagagtagggtttagaggagaaagggaacagatcgctggcttttctc
tgaattagcctttctcatgggactggcttcagaggggtttttagagg
gaagtgtctagagccttaactgtGGTTGTGTTTCGGTAGCGGGACCAAG
CTGGAAATCAAACgtaagtgcacttttctactccttttcttctatac
gggtgtgaaattggggactttcatgtttggagtatgagttaggtcagt
tctgaagagagtgggactcatccaaaatctgaggagtaagggtcagaac
agagttgtctcatggaagaacaagacctagttagttgatgaggcagcta
aatgagtcagttgacttgggatccaaatggccagacttctgtgtaacca
acaatctaagatgtagcagcaaaaagagatttccattgaggggaaag
taaaattgttaattgtGATCACCTTTGGTGAAGGGACATCCGTGGAG
ATTGAACgtaagtatttttctactaccttctgaaattgtctaaatg
ccagtgtgacttttagaggcttaagtgtcagttttgtgaaaaatgggta
aacaagagcatttcatatttattatcagttcaaaagttaactcagctc
caaaaatgaattttagacaaaaagattaatttaagccaaattgaatgat
tcaaaggaaaaaaaattagttagatgaaaagggaattcttacagctcc
aaagagcaaaagcgaattaattttcttgaacttgcctaaatctgtaaa
tgattttgttcttacaatttaaaaagggttagagaaatgtatttcttag
tctgtttctctctctgtctgataaattattatagataaaaatgaa
aattaataggatgtgctaaaaatcagtaagaagttagaaaaatatgt
ttatgttaaagttgccacttaattgagaatcagaagcaatgtattttta
aagtctaaaatgagagataaactgtcaatacttaattctgcagagattc
tatacttgacagatatctccttttcaaaaatccaatttctatggtaga
ctaaattgaaatgatcttctcataatggagggaagatggactgacc
ccaaaagctcagatttaagaaatctgtttaagtgaagaaaaataaaaga
actgcatttttaagggcccatgaattttagaaaaataggaaatattt
aataagtgtattcttttatttctgttatttctgtggtgttttata
ccgccaaggaggccgtggcaccgtcagtgatctgtagaccccatggcg
gcctttttcgcgattgaatgaccttggcggtgggtcccagggtcttg
tggcagcgcaccagccgctaaaagccgctaaaactgccgctaaaggcca
cagcaaccccgacccgctcaactgtgtgacacagtatacagat
aatgtcgctaacagaggagaatagaatatgacgggcacacgctaattgt
gggaaaaaggaggagaagcctgattttatttttagagattctagagata
aaattcccagtattatactttaaataaaaaatttctattaggagatta

taaagaattaaagctatTTTTAAGTGGGTGTAATCTTCAGTAGT
 CTCTTGCAAATGGATTAAGTAATAGAGGCTTAATCCAAATGAGAGAAA
 TAGACGCATAACCTTCAAGGCAAAAGCTACAAGAGCAAAAATGAACA
 CAGCAGCCAGCCATCTAGCCACTCAGATTTGATCAGTTTACTGAGTT
 GAAGTAAATATCATGAAGGTATAATTGCTGATAAAAAATAAGATACAGG
 TGTGACACATCTTAAGTTTCAGAAATTAATGGCTCAGTAGGATTATA
 TTTCACGTATACAAAGTATCTAAGCAGATAAAAAATGCCATTATGGAAAC
 TTAATAGAAATATATTTAAATCTCTCATTCTGTGACAGAAATTTCT
 AATCTGGGTCTTTAATCACTACCTTTGAAAGAGTTAGTAATTTGCT
 ATTGCCATCGCTGTTACTCCAGCTAATTTCAAAGTGACTTGAGAA
 AGATTATTTTGGTTGCAACCACCTGGCAGGACTATTTAGGGCCATT
 TAAACTCTTTCAACTAAGTATTTAACTGTTCTAAACCATTAGGG
 CCTTTAAAAATCTTTCATGAATTTCAACTTCGTTAAAGTTATTAAG
 GTGCTGGCAAGAACTCTTATCAAAATAGCTAATAGTTAATCTGTTA
 ATGCAGGATATAAAATTAAGTGATCAAGGCTGACCCAAACAGGAGTAT
 CTTCATAGCATATTTCCCTCTTTTCTAGAAATCATATGATTTGC
 TGCCAAGGCTATTTATATAATCTCTGGAAAAAATAGTAATGAAGGT
 AAAAGAGAAGAAAATATCAGAACATTAAGAATTCGGTATTTACTAAGT
 CTGGTTAACATGAAGTTTATTTATTAAGTTTCTATCTTTATAAA
 AATCTGTTCCCTTTCTGCTGATTTCTCAAGCAAAAGATTCTTGATTG
 TTTTAACTCTTACTCTCCACCAAGGGCTGAATGCCACAAAGGGG
 ACTCCAGGAGGCCATCTGGCAGCTGCTCACCCTCAGAAAGTGAAGCCAGC
 CAGTCTCTCTGGCAGGTGGCCAAATACAGTTGACCCCTCTGCTCT
 GGCTGAACCTTGCCCATATGGTGACAGCCATCTGGCCAGGGCCAGGTC
 TCCCTCTGAAGCCTTTGGGAGGAGAGGGAGAGTGGCTGGCCGATCACAG
 ATGCGGAAGGGGTGACTCTCAACCGGGTGCAGACTCTGCAGGGTGGG
 TCTGGGCCAACACACCCAAAGCACGCCAGGAAGGAAGGCAGCTGGT
 ATCACTGCCAGAGCTAGGAGAGGCACCGGGAAATGATCTGTCCAAGAC
 CCGTCTGTCTTAACTCCGAGGGGGTCAAGTGAAGTGGTTTGTTC
 TTGGCCTGAAGCATCTGTTCCCTGCAAGAAGCGGGGAACACAGAGGAAG
 GAGAGAAAAGATGAAGTGAACAAAGCATGAAGGCAAAAAGGCCTTAGG
 ATGGCTGCAGGAAGTAGTTCTCTGCAATGGCTCTTACTGGCTGCTG
 ATCGCCACAAACAACGCACCCAGTGGAGAATCTCCCTGTACTTAACA
 CCATTCTCTGTGCTTCTCTCAGGGGCTGATGCCAAGCCATCCGTCTT
 CATCTTCCCGCCATCGAAGGAGCAGTTAGCGACCCCAACTGTCTCTGTGG
 TGTGCTTGATCAATAACTTCTTCCCCAGAGAAATCAGTGTCAAGTGGA
 GTGGATGGGGTGGTCCAAAGCAGTGGTTCATCCGGATAGTGTACAGAGCA
 GGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCCTCTCGCTGCCCA
 CGTCACAGTACCTAAGTCATAATTTATATTCCTGTGAGGTCACCCACAAG
 ACCCTGGCCTCCCTCTGGTCAACAAGCTTCAACAGGAACGAGTGTGAGGC
 TtagAGGCCACAGGCCCTGGCCTGCCCCAGCCCCAGCCCCCTCCCC
 ACCTCAAGCCTCAGGCCCTTGCCCCAGAGGATCCTTGGCAATCCCCCAGC
 CCTCTTCCCTCCTCATCCCCCTCCCCCTCTTTGGCTTTAACCGTGTTAAT
 ACTGGGGGGGTGGGGGAATGAATAAATAAAGTGAACCTTTGCACCTGTGAT
 TTCTCTCTCTGTGATTTAAGGTGTTAAATGTTGTTTCCCCATTA
 TAGTTAATCTTTAAGGAACATACACTGAGTTGCTAAAACTACACCAT
 CACTTAAAAATTCAGCCTTCTCAGTTCTCCCTCCCTCTCTGCTCTC
 GTAAGACAGGCTCTGTGAACCCATAAGCACTTCTTTACACCTCTC
 CTGGGCCGGGTAGGAGACTTTTGTGCTCCCTCTCAGCAAGCCTCAG
 AACCATTTGAGGGGACAGTTCTTACAGTCACAT*TCCTGTGATCTAAT
 GACTTAGTTACCGAAAAGCAGTCTCTCAAAAAGAGGGAACGGCTAGA
 AACCAAGTCATAGAAATATATATGTATAAAATATATATATCCATAT
 GTAAAATAACAAATAATGATAACAGCATAGGTCAACAGGCAACAGGGAA
 TGTGAAGTCCATTCTGGCACTTCAATTAAGGGAATAGGATGCCTCAT
 TACATTTAAATACAATACATGGAGAGCTCTCTGCTGCAAGACCA
 TCTGAATGCTTCCACACTCACTACAAGGTAAAGCATTCAATCAAT

	ggtgatcaggaggttcccgttgtggctcagcaggtaagaacgtgactgg tatccaggaggatgcgggttgggtcccagcctcgtcagtggaataagg atccagtgttgcgaagatcacgggctcagatcccgtgttctatggcta tggtgtaggctggtagctgcagccctaattgacccctagcctggg aactgccatagccacatgtgaggcccttaaaacctaagaaaaaaa gaaaagaaatattacaccaatttatagataagagagaagctaagggtg gcaggcccaggatcaaagccctacctgcctatcttgacacctgatacaa ttctgttcttagggtttcaacactgcatagaacagagggtcaaacatg ctaccctcccaggactcctcccttcaaatgacataaattttgtgcca tctctgggggcaaaactcaacaatcaatggcatctctagtaccaagcaag gctcttctcatgaagcaaaactctgaagccagatccatcatgaccaagg aagtaagacagggtgttactgggtgaactgtatcctcaattcaatatgc tcaattccaactcccagtcctcgtaaatacaaccccccttgggaagaga gtccttgcatgtagccacgttaaaaagagattatacagaaaggctagt gaggatgcagtgaaacgggatcttcatactgtggtggaaatgtaa atgctgcaggcactctagaaaataattgccagttttgaaaagctaaa caaaatagtttagttgcattctgggttattatccccagaaattaaaa ttatgtccgcacaaaaacgtgtacataatcattcataacagcctgtac
--	--

Seq ID No.12	caaggaaacagctggaactcaaacgtaagtcaatccaaacgttcttcc ttggctgtctgtctctacgggtctgtggtctgaaatgattcatgtgc tgactctctgaaaccagactgacattctccaggggcaaaactaaagcctgt catcaaactggaaaactgagggcacatttctgggcagaactaagagtca ggcactgggtgaggaaaaactgttagaatgatagttcagaaacttact gggaagcaaaagcccatgttctgaacagagctctgctcaagggtcaggagg ggaaccagttttgtacaggagggaagttgagacgaacccctgtgtatat ggtttcggcgcggggaccaagctggagctcaaacgtaagtggcttttcc gactgattcttctgttcttaattgttgggttggcttttgcattttt cagtgtttcatcgaattagtgtcagggaacaaacaaattgccttccca gattaggtaccaggaggaggacattgctgcatgggagaccagagggtggc taattttaacgtttccaagccaaaataactggggaagggggctgtgt cctgtgagggttaggttttatagaagtgaagtaaggggaaatcgctat gggtcacttttggctcggggaccaagtgagccaaaattgagtacatt ttccatcaattattgtgagattttgtcctgtgtgtcattgtgcaag ttttgacatttgggtgaatgagccattccagggaacccaaaaggatga gaccgaaaagtagaaaagagccaaactttaagctgagcagacagaccgaa ttgttgagttgtgaggagagtaggggtttagggagaaaggggaacaga tcgttggttttctctgaattagccttctcatgggactggcttcagag ggggttttgatgagggaagtggtctagagccttaactgtgggtgtgtt cggtagcgggaccaagctggaaatcaaacgtaagtgcacttttctactcc ttttctttctataggggtgtgaaattggggacttttcatgtttggagt atgagttgaggtcagttctgaagagagtgggactcatccaaaatctgag gagtaagggtcagaacagagttgtctcatggaagaacaaagacctagtta gtgatgaggcagctaaatgagtcagttgacttgggatccaaatggccag acttcgtctgtaaccaacaatctaatagatgtagcagcaaaaagagatt tccattgaggggaaagttaaattgttaattgttgatcaccttgggtga agggacatccgtggagattgaacgtaagtatttttctactaccttct gaaatttctaaatgccaggttgacttttagaggcttaagtgtcagtt ttgtgaaaatgggtaaacaagagcatttcataattattatcagtttcaa aagttaaactcagctccaaaatgaattgtagacaaaaagattaattta agccaaattgaatgattcaaaggaaaaaaaattagttagatgaaaag gaattctacagctccaaagagcaaaagcgaattaattttcttgaactt
--------------	---

	tgccaaatcttgtaaagatgtttgttctttacaatttaaaaagggtaga gaaatgtatttcttagtctgtttctctctctgtctgataaattattat atgagataaaaatgaaaaataataggatgtgctaaaaaatcagtaagaag ttgaaaaatataatgtttatgttaaagttgccacttaattgagaatcaga agcaatgttatttttaagtcataaatgagagataaactgtcaatactta aattctgcagagattctatatcttgacagatatctccttttcaaaaatc caatttctatggtagactaaattgaaatgatcttctcataatggaggg aaaagatggactgaccccaaaaagctcagattt*aagaaaacctgtttaag *gaaagaaaaataaagaactgcatttttaaggcccatgaattttaga aaaaaggaaatattttaataagtgattctttatttctgttattac ttgatgggtgtttataccgccaaggaggccgtggcaccgtcagtgat ctgtagaccccatggcgcccttttctcgatgaatgacctggcggtg ggtccccagggtctgtgtggcagcgcaccagccgctaaaagccgctaaaa actgccgctaaaggccacagcaaccccgccagccgctcaactgtgct gacacagtatacagataatgtcgtaacagaggagaatagaaatagac gggcacacgctaattgtgggaaaagaggagaaagcctgattttatttt tagagattctagagataaaaattcccagattatactctttaataaaaaa ttctattaggagattataaagaatttaagctatttttaagtgggt gtaattcttcagtagtctctgtcaaatggatttaagtaatagaggctt aatccaaatgagagaaatagacgcataaccccttcaaggcaaaagctaca agagcaaaaattgaacacagcagccagccatctagccactcagatttga tcagtttactgagttgaagtaaatcatgaaggataattgctgata aaaaaataagatacaggtgtgacacatcttaagttcagaaattaatg gcttcagtaggattatattcacgtatacaaaagtatctaagcagataaaa atgccattaatggaaacttaatagaatatattttaaatcttcattc tgtgacagaaatttctaactctgggtcttttaacacctacccttgaaa gagtttagtaatttctatttgcctcgtgttactccagctaatttca aaagtatacttgagaaagattattttggttgcaaccacctggcagga ctattttagggccattttaaaactctttcaaaactaagtattttaactg ttctaaaccatttagggccttttaaaaactctttcatgaatttcaactt cgttaaaagttattaagggtgtctggcaagaacttcttatcaaatatgct aatagtttaactgttaatgcaggatataaaatgaagtgatcaaggctt gacccaaacaggagatcttcatagcataattccctccttttttctag aattcatatgatttctgccaaggctattttatataatctctggaaaaa aaatagtaatgaagggttaaaagagaagaaaatacagaacattaagaatt cggatttttactaactgcttggttaacatgaagggtttttattttaag gtttctatctttataaaaactgttcccttttctgtgatttctccaagc aaaagattcttgatttgttttaactcttacttcccacccaagggcct gaatgcccacaaaaggggacttccaggaggccatctggcagctgctaccg tcagaagtgaaggccagccagttcctctgggcaggtggccaaaattacag ttgacccctcctggctggtgaaccttggcccatatggtagacagccatc tggccagggcccagggtctccctctgaagccttgggaggagaggagagt ggctggcccgatcacagatgcggaagggtgactcctcaaccggggtgc agactctgcagggtgggtctgggccaacacacccaaagcacgccagga aggaaaggcagcttggtatcactgccagagctaggagaggcaccgggaa aatgatctgtccaagaccgttcttcttaactccgaggggggtcaga tgaagtgtttgtttcttgccctgaagcatcgtgttccctgcaagaagc ggggaacacagaggaaggagagaaaagatgaactgaacaaagcatgcaag gcaaaaaaggccttaggatggctgcaggaagtagttcttctgattggc tcttactggctcgtgatcgcccaacaacgcacccagtgaggagaact tccctgttacttaaacaccattctgtgttcttctcaggggctgat gccaaagcatccgtcttcatcttccgccatcgaaggagcagttagcgac cccaactgtctctgtgggtgtgcttgatca
Seq ID No.15	gatgccaaagccatccgtcttcatcttcccgccatcgaaggagcagtttagc

	gacccaactgtctgtgtgtgtgtgatcaataacttcttcccagag aatcagtgtaagtggaaagtggatgggtgtccaaagcagtggtcat cggataggtcacagagcaggacagcaaggacagcacctacagcctcag cagcacctctcgtgcccacgtcacagtacctaagtcataatttatatt cctgtgaggtcacccacaagaccctggcctcccctctgtgtcacAAGCTTC AACAGGAACGAGTGTGAGGCTTAGAGGCCACAGGCCCTGGCCTGCCCC CAGCCCCAGCCCCCTCCCCACCTCAAGCCTCAGGCCCTTGCCCCAGAGG ATCCTTGGCAATCCCCCAGCCCCCTCTTCCCTCCTCATCCCCCTCCCCCTCT TTGGCTTTAACCGTGTTAATACTGGGGGGTGGGGGAATGAATAAATAAAG TGAACCTTTGCACCTGTGATTTCTCTCTCCTGTCTGATTTTAAGGTTGTT AAATGTTGTTTTCCCCATTATAGTTAATCTTTTAAGGAACTACATACTGA GTTGCTAAAAACTACACCATCACTTATAAAATTCa_gCCTTCTCAGTTCT CCCCCTCCCCCTCCTGTCTCCGTAAAGACAGGCCCTCCGTGAAACCCATAAGC ACTTCTCTTTACACCCTCTCCTGGGCCGGGGTAGGAGACTTTTTGATGTC CCCTcTTCAGCAAGCCTCAGAACCATTTTGAGGGGGACAGTTCTTACAGT CACAT*TCCtGtGATCTAATGACTTTAGTTaCCGAAAAGCCAGTCTCTCA AAAAGAAGGGAACGGCTAGAAACCAAGTCATAGAAATATATATGTATAAA ATATATATATATCCATATATGTAAAATAACAAAATAATGATAACAGCATA GGTCAACAGGCAACAGGGAATGTTGAAGTCCATTCTGGCACTTCAATTTA AGGGAATAGGATGCCTTCATTACATTTTAAATACAATACACATGGAGAGC TTCCTATCTGCCAAAGACCATCCTGAATGCCTTCCACACTCACTACAAGG TTAAAAGCATTCACTACAATGTTGATCGAGGAGTTCCCGTTGTGGCTCAG CAGGTTAAGAACGTGACTGGTATCCAGGAGGATGCGGGTTTGGTCCCCAG CCTCGCTCAGTGGATTAAGGATCCAGTGTGCTGCAAGATCACGGGCTCA GATCCCGTGTCTATGGCTATGGTGTAGGCTGGTAGCTGCATGCAGCCCT AATTTGACCCCTAGCCTGGGAAGTCCATAtGCCACATGTGAGGCCCTTA AAACCTAAAAGAAAAAaAAAGAAAAGAAATATCTTACACCCAATTTATAG ATAAGAGAGAAGCTAAGGTGGCAGGCCCAGGATCAAAGCCCTACCTGCCT ATCTTGACACCTGAtACAAATTCTGTCTTCTAGGGtTTCCAACACTGCAT AGAACAGAGGGTCAAACATGCTACCCTCCCAGGGACTCCTCCCTTCAAAT GACATAAATTTTGTGCCCCATCTCTGGGGGGCAAACTCAACAATCAATGG CATCTCTAGTACCAAGCAAGGCTCTTCTCATGAAGCAAACTCTGAAGCC AGATCCATCATGACCCAAGGAAGTAAAGACAGGTGTTACTGGTTGAACTG TATCCTTCAATTCAATATGCTCAATTTCCAACCTCCCAGTCCCCGTAAATA CAACCCCTTTGGGAAGAGAGTCTTGCAGATGTAGCCACGTAAAAAAGA GATTATACAGAAAGGCTAGTGAGGATGCAGTGAAACGGGATCTTTCATAC ATTGCTGGTGGAAATGTAAAATGCTGCAGGCACTCTAGAAAATAATTTGC CAGTTTTTTGAAAAGCTAAACAAAATAGTTTAGTTGCATTCTGGGTTATT TATCCCCCAGAAATTAAAAATTATGTCCGCACAAAACGTGTACATAATC ATTCATAACAGCCTTGTACGAAAAGCTT
Seq ID No.16	GGATCCTTAACCCACTAATCGAGGATCAAACACGCATCCTCATGGACAAT ATGTTGGGTTCTTAGCCTGCTGAGACACAACAGGAACTCCCCTGGCACCA CTTTAGAGGCCAGAGAAACAGCACAGATAAAATTCCCTGCCCTCATGAAG CTTATAGTCTAGCTGGGGAGATATCATAGGCAAGATAAACACATACAAAT ACATCATCTTAGGTAATAATATACTAAGGAGAAAATTACAGGGGAGAA AGAGGACAGGAATTGCTAGGGTAGGATTATAAGTTCAGATAGTTTCATCAG GAACACTGTTGCTGAGAAGATAACATTTAGGTAAAGACCGAAGTAGTAAG GAAATGGACCGTGTGCCTAAGTGGGTAAGACCATTCTAGGCAGCAGGAAC AGCGATGAAAGCACTGAGGTGGGTGTTCACTGCACAGAGTTGTTCACTGC ACAGAGTTGTGTGGGGAGGGGTAGGTCTTGCAGGCTCTTATGGTCACAGG AAGAATTGTTTTACTCCCACCGAGATGAAGGTTGGTGGATTTTGAGCAGA AGAATAATTCTGCCTGGTTTATATATAACAGGATTTCCCTGGGTGCTCTG ATGAGAATAATCTGTCAGGGGTGGGATAGGGAGAGATATGGCAATAGGAG CCTTGGCTAGGAGCCCACGACAATAATTCCAAGTGAGAGGTGGTGTGCTGCA

TTGAAAGCAGGACTAACAAGACCTGCTGACAGTGTGGATGTAGAAAAAGA
 TAGAGGAGACGAAGGTGCATCTAGGGTTTTCTGCCTGAGGAATTAGAAAG
 ATAAAGCTAAAGCTTATAGAAGATGCAGCGCTCTGGGGAGAAAGACCAGC
 AGCTCAGTTTTGATCCATCTGGAATTAATTTTGGCATAAAGTATGAGGTA
 TGTGGGTAAACATTATTTGTTTTTTTTTTTCCATGTAGCTATCCAACCTG
 TCCCAGCATCATTTATTTTAAAAGACTTTCCTTTCCCCTATTGGATTGTT
 TTGGCACCTTCACTGAAGATCAACTGAGCATAAAATTGGGTCTATTTCTA
 AGCTCTTGATTCCATTCCATGACCTATTTGTTTCATCTTTACCCCAGTAGA
 CACTGCCTTGATGATTAAAGCCCCTGTTACCATGTCTGTTTTGGACATGG
 TAAATCTGAGATGCCTATTAGCCAACCAAGCAAGCACGGCCCTTAGAGAG
 CTAGATATGAGAGCCTGGAATTCAGACGAGAAAGGTCAGTCCTAGAGACA
 TACATGTAGTGCCATCACCATGCGGATGGTGTAAAAGCCATCAGACTGC
 AACAGACTGTGAGAGGGTACCAAGCTAGAGAGCATGGATAGAGAAACCCA
 AGCACTGAGCTGGGAGGTGCTCCTACATTAAGAGATTAGTGAGATGAAGG
 ACTGAGAAGATTGATCAGAGAAGAAGGAaAATCAGGAAAATGGTGCTGTC
 cTGAaAATCCAAGGGAAGAGATGTTCCAAAGAGGAGAAaAACTGATCAGTT
 GTCAGCTAGCGTCAATTGGGATGAAAATGGACCATTGGACAGAGGGGATGT
 AGTGGGTCATGGGTGAATAGATAAGAGCAGCTTCTATAGAATGGCAGGGG
 CAAAATTCTCATCTGATCGGCATGGGTTcTAAAGAAAACGGGAAGAAAAA
 ATTGAGTGCATGACCAGTCCCTTCAAGTAGAGAGGTgGAAAAGGGAAGGA
 GGAAAATGAGGCCACGACAACATGAGAGAAATGACAGCATTTTTTAAAAAT
 TTTTTATTTTATTTtATTTATTTATTTTTGCTTTTTTAGGGCTGCCCCTGC
 AAcatatggaggttcccaggttaggggtctaatacagagctatagctgcca
 gcctacaccacagccatagcaatgccagatctacatgacctacaccacag
 ctcacagcaacgccggtacctaaccactgagtgaggccagagatcaaa
 cccatatccttatggatactagtcaggttcattaccactgagccaaaatg
 ggaaATCCTGAGTAATGACAGCATTTTTTAATGTGCCAGGAAGCAAACT
 TGCCACCCCGAAATGTCTCTCAGGCATGTGGATTATTTTGAGCTGAAAAC
 GATTAAGGCCCAAAAAACACAAGAAGAAATGTGGACCTTCCCCCAACAGC
 CTAAAAAATTTAGATTGAGGGCCTGTTCCCAGAATAGAGCTATTGCCAGA
 CTTGTCTACAGAGGCTAAGGGCTAGGTGTGGTGGGGAAACCCTCAGAGAT
 CAGAGGGACGTTTATGTACCAAGCATTGACATTTCCATCTCCATGCGAAT
 GGCCTTCTTCCCCTCTGTAGCCCCAAACCACCACCCCAAAATCTTCTTC
 TGTCTTTAGCTGAAGATGGTGTGAAGGTGATAGTTTCAGCCACTTTGGC
 GAGTTCCTCAGTTGTTCTGGGTCTTTCCCTCCGGATCCACATTATTCGACT
 GTGTTTGATTTTCTCCTGTTTATCTGTCTCATTGGCACCCATTTCAATTCT
 TAGACCAGCCCAAGAACCTAGAAGAGTGAAGGAAAATTTCTTCCACCCT
 GACAAATGCTAAATGAGAATCACCGCAGTAGAGGAAAATGATCTGGTgCT
 GCGGGAGATAGAAGAGAAAATcGCTGGAGAGATGTCAGTGAAGTGAAG
 ATGGGAAAGGGGGGGGCACAGGTGGAGGTGTTGCCCTCAGCTAGGAAGACA
 GACAGTTcacagaagagaagcgggtgtccgtGGACATCTTGCCCTCATGGA
 TGAGGAAACCGAGGCTAAGAAAGACTGCAAAAGAAAGGTAAGGATTGCAG
 AGAGGTGATCCATGACTAAAATCACAGTAACCAACCCCAAAACCACCATG
 TTTTCTCCTAGTCTGGCACGTGGCAGGTACTGTGTAGGTTTTCAATATTA
 TTGGTTTGTAACAGTACCTATTAGGCCTCCATCcCCTCCTCTAATACTAA
 CAAAAGTGTGAGACTGGTCAGTGAAAAATGGTCTTCTTTCTCTATGCAAT
 CTTTCTCAAGAAGATACATAACTTTTTATTTTATCATaGGCTTGAAGAGC
 AAATGAGAAACAgCCTCCAACCTATGACACCGTAACAAAGTGTATTATGAT
 CAGTGAAGGGCAAGAAACAAAACATACACaGTAAAGACCCTCCATAATAT
 TGtGGGCTGGCCCAaCACAGGCCAGGTTGTAAAAGCTTTTTATTCTTTGA
 TAGAGGAATGGATAGTAATGTTTCAACCTGGACAGAGAT*CATGTTCACT
 GAATCCTTCCAAAAATTTCATGGGTAGTTTGAAtTATAAGGAAAATAAGAC
 TTAGGATAAATACTTTgTCCA*GATCCCAGAGTTAATgCCAAAATCAGTT
 TTCAGACTCCAGGCAGCCTGATCAAGAGCCTAAACTTTAAAGACACAGTC
 CCTTAATAACTACTATTACAGTTGCACTTTCAgGGCGCAAAGACTCATT
 GAATCCTACAATAGAATGAGTTTAGATATCAAATCTCTCAGTAATAGATG

	AGGAGACTAAATAGCGGGCATGACCTGGTCACTTAAAGACAGAATTGAGA TTCAAGGCTAGTGTTCTTTCTACCTGTTTTGTTTCTACAAGATGTAGCAA TGCGCTAATTACAGACCTCTCAGGGAAGGAATTCACAACCCTCAGCAAAA ACCAAAGACAAATCTAAGACAATAAGAGTGTTGGTTTAATTTGGAAAAA TAACTCACTAACCAAACGCCCTCTTAGCACCCCAATGTCTTCCACCATC ACAGTGCTCAGGCCTCAACCATGCCCAATCACCCAGCCCCAGACTGGT TATTACCAAGTTTCATGATGACTGGCCTGAGAAGATCAAAAAAGCAATGA CATCTTACAGGGGACTACCCCGAGGACCAAGATAGCAACTGTCATAGCAA CCGTCACACTGCTTTGGTCA
Seq ID No.19	ggatcaaacacgcacatcctcatggacaatatgttgggttcttagcctgctg agacacaacaggaactccctggcaccactttagaggccagagaaacagc acagataaaattccctgccctcatgaagcttatagtctagctggggagat atcataggcaagataaacacatacaatacatcatcttaggtaataatat atactaaggagaaaattacaggggagaaaggagacaggaattgctagggt aggattataagttcagatagttcatcaggaacactgttgctgagaagata acatttaggtaaagaccgaagtagtaaggaaatggaccgtgtgcctaagt gggtaagaccattttaggcagcaggaacagcgatgaaagcactgaggtgg gtgttactgcacagagttgttactgcacagagttgtgtggggaggggt aggtctgcaggtcttatggtcacaggaagaattgtttactcccaccg agatgaaggttggtggattttgagcagaagaataattctgcctggtttat atataacaggatttccctgggtgctctgatgagaataatctgtcaggggt gggataggagagatatggcaataggagccttggttaggagcccacgaca ataattccaagttagaggttggtgctgcattgaaagcaggactaacaagac ctgctgacagtggtgtagtaaaaagatagaggagacgaaggtgcatct agggttttctgcctgaggaattagaaagataaagctaaagcttatagaag atgcagcgctctggggagaaagaccagcagctcagtttgatccatctgg aattaattttggcataaagtatgaggtatgtgggttaacattatttgtt tttttttccatgtatccaactgtcccagcatcatttattttaaa agactttcctttccctatttgattgtttggcaccttactgaagatca actgagcataaaattgggtctatttctaagctcttgattccattccatga cctattttgctatcttaccagtagacactgccttgatgattaaagcc cctgttaccatgtctgttttgacatggtaaatctgagatgcctatttagc caaccaagcaagcacggcccttagagagctagatatgagagcctggaatt cagacgagaaaaggtcagtccttagagacatacatgtagtccatcaccatg cggatggtgttaaaagccatcagactgcaacagactgtgagagggtagcca agctagagagcatggatagagaaaccaagcactgagctgggaggtgctc ctacattaagagattagttagatgaaggactgagaagattgatcagagaa gaaggaaaatcaggaatgtgtgctgtcctgaaaatccaagggaagagat gttccaaaaggagaaaaactgatcagttgtcagctagcgtcaattgggat gaaaatggaccattggacagagggatgtagtgggtcatgggtgaatagat aagagcagcttctatagaatggcaggggcaaaattctcatctgatcgga tgggttctaaagaaaacgggaagaaaaaattgagtgcatgaccagtcct tcaagtagagaggtggaaaagggaaggaggaatgaggccacgacaaca tgagagaaatgacagcatttttaaaatttttattttatttattt tattttgcttttagggctgcccctgcaacatatggaggttcccagggt aggggtctaatcagagctatagctgccagctacaccacagccatagcaa tgccagatctacatgacctacaccacagctcacagcaacgccggatcctt aaccactgagttagggccagagatcaaaccatactcttatggatactag tcaggttattaccactgagccaaaatgggaaatctgagtaatgacagc atttttaattgtccaggaagcaaaacttgccaccccgaaatgtctctca ggcatgtgattttttagctgaaaacgattaaggcccaaaaaacacaa gaagaaatgtggaccttccccaacagcctaaaaaatttagattgagggc ctgttcccagaatagagctattgccagactgtctacagaggctaagggc taggtgtggtggggaaccctcagagatcagagggacgtttatgtaccaa

	gcattgacatttccatccatgcgaatggccttcttcccctctgtagcc ccaaaccaccacccccaaaatcttcttctgtcttagctgaagatgggtg tgaaggatagatttcagccacttggcgagttcctcagttgttctgggt cttctctccTgatccacattattcgactgtgtttgattttctctgttta tctgtctcattggcaccatttctttagaccagcccaaagaacctag aagagtgaaggaaaatttcttccaccctgacaaatgctaaatgagaatca ccgcagtagaggaaaatgatctggtgctgcgggagatagaagagaaaatc gctggagagatgtcactgagtaggtgagatgggaaaggggtgacacaggt ggaggtgttgcctcagctaggaagacagacagttcacagaagagaagcg ggtgtccgtggacatcttgcctcatggatgaggaaaccgaggctaagaaa gactgcaaaagaaaggaaggattgcagagaggtcgatccatgactaaaa tcacagtaaccaacccccaaaccaccatgtttctcctagtctggcagtg gcaggtactgtgtaggtttcaatattattggttgaacagtacctatt aggcctccatccccctcctctataactaactaaaagtgtgagactggtcagt gaaaaatggtcttcttctctatgaatcttctcaagaagatacataact ttttatttatcataggcttgaagagcaaatgagaaacagcctccaacct atgacaccgtaaaaaatgtttatgatcagtgaagggaagaaacaaaac atacacagtaaaagaccctccataatattgtgggtggcccaacacaggcca ggttgtaaaagctttttattctttgatagaggaaatggatagtaattgttc aacctggacagagatcatgttccactgaatccttccaaaaattcatgggta gttgtaattataaggaaaataagacttaggataaatactttgtccaagat cccagagttaatgccaaaatcagtttcagactccaggcagcctgatcaa gagcctaaactttaagacacagtccttaataactactattcacagtgtg cactttcaggggcgcaagactcattgaatcctacaatagaatgagtttag atatcaaatctctcagtaataatagatgaggagactaaatagcgggcatgacc tggtcacttaagacagaattgagattcaaggctagtgttcttctacct gtttgtttctacaagatgtagcaatgcgctaattacagacctctcaggg aagggaattcacaacctcagcaaaaaccaaagacaaatctaagacaacta agagtgttggttaatttgaaaaataactcactaaccaaacgcccctct tagcaccatgttctccaccatcacagtgtcaggcctcaacatgcc ccaatcacc
Seq ID No.25	GCACATGGTAGGCAAAGGACTTTGCTTCTCCCAGCACATCTTTCTGCAGA GATCCATGGAAACAAGACTCAACTCCAAAGCAGCAAAGAAGCAGCAAGTT CTCAAGTGATCTCCTCTGACTCCCTCCTCCCAGGCTAATGAAGCCATGTT GCCCCCTGGGGGATTAAGGGCAGGTGTCCATTGTGGCACCCAGCCCGAAGA CAAGCAATTTGATCAGGTTCTGAGCACTCCTGAATGTGGACTCTGGAATT TTCTCCTCACCTTGTGGCATATCAGCTTAAGTCAAGTACAAGTGACAAAC AACATAATCCTAAGAAGAGAGGAATCAAGCTGAAGTCAAAGGATCACTGC CTTGGATTCTACTGTGAATGATGACCTGGAAAATATCCTGAACAACAGCT TCAGGGTGATCATCAGAGACAAAAGTTCCAGAGCCAGGTAGGGAAACCCT CAAGCCTTGCAAAGAGCAAAATCATGCCATTGGGTTCTTAACCTGCTGAG TGATTTACTATATGTTACTGTGGGAGGCAAAGCGCTCAAATAGCCTGGGT AAGTATGTCAAATAAAAAGCAAAAGTGGTGTTCCTTGAAATGTTAGACCT GAGGAAGGAATATTGATAACTTACCAATAATTTTCAGAATGATTTATAGA TGTGCACTTAGTCAGTGTCTCTCCACCCCGCACCTGACAAGCAGTTTAGA ATTTATTCTAAGAATCTAGGTTTGCTGGGGGCTACATGGGAATCAGCTTC AGTGAAGAGTTTGTTGGAATGATTCATAAATTTTCTATTTCCAGCATAA ATCCAAGAACCTCTCAGACTAGTTTATTGACACTGCTTTTCCTCCATAAT CCATCTCATCTCCGTCCATCATGGACACTTTGTAGAATGACAGGTCTCTGG CAGAGACTCaCAGATGCTTCTGAAACATCCTTTGCCTTCAAAGAATGAAC AGCACACATACTAAGGATCTCAGTGATCCACAAATTAGTTTTTGCCACAA TGGTTCTTATGATAAAAGTCTTTCATTAACAGCAAATTGTTTTATAATAG TTGTTCTGCTTTATAATAATTGCATGCTTCACTTTCTTTTCTTTTCTTTT TTTTTCTTTTGTCTTTTAGTGCCGAGGTgcagcatatgaaatttcc caggctaggggtcaaatcagaactacacctactggcctacgccacagcca cagcaactcaggatctaagccatgtcggtgacctacactacagctcatgg

caatgccagatccttaaccaatgagcgaggccaggatcgaacccatgt
 cctcatggatactagtcaggctcattatccgctgagccataacaggaact
 cccGAGTTTGCTTTTTATCAAAATTGGTACAGCCTTATTGTTTCTGAAAA
 CCACAAAATGAATGTATTACATAATTTAAAAGGTTAAATAATTTATGA
 TATACAAGACAATAGAAAGAGAAAACGTCATTGCCTCTTTCTCCACGAC
 AACACGCCTCCTTAATTGATTTGAAGAAATAACTACTGAGCATGGTTTAG
 TGTACTTCTTTTTCAGCAATTAGCCTGTATTCATAGCCATACATATTCAATT
 AAAATGAGATCATGATATCACACAATACATACCATACAGCCTATAGGGAT
 TTTTACAATCATCTTCCACATGACTACATAAAAAACCTACCTAAAAA
 AAAAACCTACTTTCATCCTCCTATTGGCTGCTTTGTGCTCCATTAAAAAG
 CTCTATCATAATTAGGTTATGATGAGGATTTCCATTTTCTACCTTTCAAG
 CAACATTTCAATGCACAGTCTTATATACACATTTGAGCCTACTTTTCTTT
 TTCTTTCTTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTGGTCTTTTGTGTC
 TTTTCTAAGgctgcatatggagggtccaggctagctgtctaatacagaac
 tatactgctggcctacgccacatccacagcaataacagatctgagccat
 gctgcaacttacaccacagctcacagcaacggtggatccttaaacact
 gagcaaggccaggatcaaaccatAACTTCATGGCTCCTAGTTGGATTT
 GTTAACCACTGAGCCATGATGGCAACTCCTGAGCCTACTTTTCTAATCAT
 TTCCAACCCTAGGACACTTTTTTAAGTTTCATTTTCTCCCCCACC
 TGTTTTCTGAAGtGTGTTTGCTTCCACTGGGTGACTTCACtCCAGGATC
 TCATCTGCAGGATACTGCAGCTAAGTGTATGAGCTCTGAATTTGAATCCC
 AACTCTGCCACTCAAAGGGATAGGAGTTTCCGATGTGGCCCAATGGGATC
 AGTGGCATCTCTGCAGTGCCAGGACGCaggttccatccctggcccagcac
 agtgggtaagaatctggCATTGCTGCAGCTGAGGCATAGATTTCAATTG
 TGCCTCAgATCTGATCCTTGGCCCAAGGACTGCATATGCCTCAGGGCAAC
 CAAAAAAGAGAAAAGGGGGGTGATAGCATTAGTTTCTAGATTTGGGGGAT
 AATTAAATAAAGTGATCCATGTACAATGTATGGCATTTTGTAAATGCTCA
 ACAAAATTTCAACTATTATggagttccatcatggctcagtggaagggaat
 ctgattagcatccatgaggacacaggtCCAACCCCGACCTTGCTCAGTGG
 GCATTGCTGTGAGCTGTGGCATGGGTACAGACGAAGCTCGGATCTGGCA
 TTGCTGTGGCTGTGGTGTAAAGCCAgCAActacagctctcattcagccct
 agcctgggaacctccatagccTAAAAGACAAAAAATAAAATTTAAATTA
 AAAATAAAGAAATGTAACTATTATGATTGgTACTGCTTGCATTACTGCA
 AAGAAAGTCACTTTCTATACTCTTTAATATCTTAGTTGACTGTGTGCTCA
 GTGAACATATTTTGGACACTTAATTTCCACTCTCTTCTATCTCCAACCTGA
 CAACTCTCTTTCTCTCTTCTGGTGAGATCCACTGCTGACTTTGCTCTTT
 AAGGCAACTAGAAAAGTGCTCAGTGACAAAATCAAAGAAAGTTACCTTAA
 TCTTCAGAATTACAATCTTAAGTTCTCTTGTAAGCTTACTATTTCAAGT
 GTTAGTATTATTCCTTGGTCCCTTACAACCTATCAGCTCTGATCTATTGC
 TGATTTTCAACTATTTATTGTTGGAGTTTTTTCCTTTTTTCCCTGTTTCA
 TCTGCAAATGTTTGCTGAGCATTGTGCAAGTGAAGATACTGGACTGGGCC
 TTCCAAATATAAGACAATGAAACATCGGGAGTTCTCATTATGGTGCAGCA
 GAaacgaatccaactaggaaatgtgaggtgcaggttcgatccctgcct
 tgctcagtggttaaggatccagcattaccgtgagctgtgggtgtaggtg
 cagacgtggctcagatcctgcgttgctgtggctgtggcataggctggcag
 ctctagctctgattcgaccgtagcctgggaacctccatGCGCCCCGAGT
 GCAGCCCTTAAAAAGCAAAAAAAGAAAGAAAGAAAAAGACAATGAAA
 CATCAAACAGCTAACAATCCAGTAGGGTAGAAAGAATCTGGCAACAGATA
 AGAGCGATTAAATGTTCTAGGTCCAGTGACCTTGCCTCTGTGCTCTACAC
 AGTCGTGCCACTTGCTGAGGGAGAAGGTCTCTCTTGAGTTGAGTCCTGAA
 AGACATTAGTTGTTCAAACTAATGCCAGTGAGTGAAGGTGTTTCCAAG
 CAGAGGGAGAGTTTGGTAAAAAGCTGGAAGTCACAGAAAGACTCTAAAGA
 GTTTAGGATGGTGGGAGCAACATACGCTGAGATGGGGCTGGAAGGTAAAG
 AGGGAAACAACCTATAGTAAGTGAAGCTGGACTCACAGCAAAGTGAGGACC
 TCAGCATCCTTGATGGGGTTACCATGGAAACACCAAGGCACACCTTGATT
 TCCAAAACAGCAGGCACCTGATTCAGCCCAATGTGACATGGTGGGTACCC

CTCTAGCTCTACCTGTTCTGTGACAACTGACAACCAACGAAGTTAAGTCT
 GGATTTTCTACTCTGCTGATCCTTGTTTTTGTTCACACGTCATCTATAG
 CTTTCATGCCAAAATAGAGTTCAAGGTAAGACGCGGGCCTTGGTTTGATAT
 ACATGTAGTCTATCTTGTTTGAGACAATATGGTGGCAAGGAAGAGGTTCA
 AACAGGAAAATACTCTCTAATTATGATTAAGTGAAGAAAGCTAAAGAGTC
 CCATAATGACACTGAATGAAGTTCATCATTTTGCAAAAGCCTTCCCCCCCC
 CCCAGGAGACTATAAAAAAGTGCAATTTTTTAAATGAACTTATTTACAAA
 ACAGAAATAGACTCACAGACATAGGAAACGAACAGATGGTTACCAAGGGT
 GAAAGGGAGTAGGAGGGGATAAATAAGGAGTCTGGGGTTAGCAGATACACC
 CCAGTGTACACAAAATAAACACAGGGGACCTACTATATAGCACAGGGAAC
 TATATGCAGTAGCTTACAATAACCTATAATGGAAAAGAATGTGAAAAAGA
 ATATATGTATGCGTGTGTGTGTAAGTGAATCACTTTGCTGTAACTGAAT
 CTAACATAACATTGTAAATCAACTACAGTTTTTTTTTTTTTTAAGTGCAG
 GGTTTTGGTGTTTTTTTTTTTCATTTTGTGTGTGTGTGTGTGTGTGTGTGT
 TTTTAGGGCCACACCCAGACATATGGGGGTCCCAGGctAGGGGTcTAA
 TTAGAGcTACAGtGCCGGCTTGCAccacagccacagcaacatcagatcc
 gagccgcacttgcgacttacaccacagctcatggcaataccagatcctta
 acccactgagcaaggcccaggatcgtaaccgcaacctcatggttcctag
 tcagattcattCTGCTGCGCTACAATGGGAACTCCAAGTGCAGTTTTTT
 GTAATGTGCTtGTCTTTCTTTGTAAATTCATATTCATCCTACTTCCCAATA
 AATAAATAAATACATAAATAAATAAACATACCATTGTAAATCAACTACAAT
 TTTTTTAAATGCAGGGTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
 CCTTTTCTAgggccgctcccatggcatatggaggtcccaggctaggggt
 cgaatcggagctgtagccaccggcctacgccagagccacagcaacgcggg
 atccgagccgctgtgcaacctacaccacagctcacggcaacgcggatc
 gttacccactgagcaagggcagggatcgaacctgcaacctcatggttc
 tagtcagattcgttaactactgagccacaacggaaacTCCTAAAGTGCAG
 TTTTAAATGTGCTTGTCTTTCTTTGTAAATTTACACTCAACCTACTTCCC
 AATAAATAAATAAATAAACAATAAATCATAGACATGGTTGAATTCTAAA
 GGAAGGGACCATCAGGCCTTAGACAGAAATACGTCATCTTCTAGTATTTT
 AAAACACACTAAAGAAGACAAACATGCTCTGCCAGAGAAGCCCAGGGCCT
 CCACAGCTGCTTGCAAAGGGAGTTAGGCTTCAGTAGCTGACCCAAGGCTC
 TGTTCTCTTCAGGGAAAAGGGTTTTTGTTCAGTGAGACAGCAGACAGCT
 GTCAGTGTGgtggacgttcggccaaggaaccaagctggaactcaaacGTA
 AGTCAATCCAAACGTTCTTCTTGGCTGTCTGTGTCTTACGGTCTCTGT
 GGCTCTGAAATGATTCATGTGCTGACTCTCTGAAACCAGACTGACATTCT
 CCAGGGCAAACTAAAGCCTGTCATCAAACcGGAAACTGAGGGCACATT
 TTCTGGGCAGAACTAAGAGTCAGGCACTGGGTGAGGAAAAACTTGTTAGA
 ATGATAGTTTTCAGAACTTACTGGGAAGCAAAGCCCATGTTCTGAACAGA
 GCTCTGCTCAAGGGTCAGGAGGGGAACCAGTTTTTGTACAGGAGGGAAGT
 TGAGACGAACCCCTGTGTAtatggttcggcgcggggaccaagctggagc
 tcaaacGTAAGTGGCTTTTTCCGACTGATTCTTTGCTGTTTCTAATTGTT
 GGTTGGCTTTTTGTCCATTTTTCAGTGTTCATCGAATTAGTTGTCAGG
 GACCAAAACAAATTGCCTTCCCAGATTAGGTACCAGGGAGGGGACATTGCT
 GCATGGGAGACCAGAGGGTGGCTAATTTTAAACGTTTCCAAGCCAAAATA
 ACTGGGGAAGGGGGCTTGCTGTCTGTGAGGGTAGGTTTTTATAGAAGTG
 GAAGTTAAGGGGAAATCGCTATGGTtcacttttggtcggggaccacaaagt
 ggagcccaaaattgaGTACATTTTCCATCAATTATTTGTGAGATTTTTGT
 CCTGTTGTGTCATTTGTGCAAGTTTTTGACATTTTGGTTGAATGAGCCAT
 TCCCAGGGACCCAAAAGGATGAGACCGAAAAGTAGAAAAGAGCCAACCTT
 TAAGCTGAGCAGACAGACCGAATTGTTGAGTTTGTGAGGAGAGTAGGGTT
 TGTAGGGAGAAAGGGGAACAGATCGCTGGCTTTTTCTCTGAATTAGCCTT
 TCTCATGGGACTGGCTTCAGAGGGGGTTTTTGATGAGGGAAGTGTCTAG
 AGCCTTAAGTGTGGgttggttcggttagcgggaccaagctggaaatcaaa
 CGTAAGTGCACCTTTCTACTCC

Porcine Lambda Light Chain

In another embodiment, novel genomic sequences encoding the lambda light chain locus of ungulate immunoglobulin are provided. The present invention provides the first reported genomic sequence of ungulate lambda light chain regions. In one embodiment, the porcine lambda light chain nucleotides include a concatamer of J to C units. In a specific embodiment, an isolated porcine lambda nucleotide sequence is provided, such as that depicted in Seq ID No. 28.

In one embodiment, nucleotide sequence is provided that includes 5' flanking sequence to the first lambda J/C region of the porcine lambda light chain genomic sequence, for example, as represented by Seq ID No 32. Still further, nucleotide sequence is provided that includes 3' flanking sequence to the J/C cluster region of the porcine lambda light chain genomic sequence, for example, approximately 200 base pairs downstream of lambda J/C, such as that represented by Seq ID No 33. Alternatively, nucleotide sequence is provided that includes 3' flanking sequence to the J/C cluster region of the porcine lambda light chain genomic sequence, for example, approximately 11.8 kb downstream of the J/C cluster, near the enhancer (such as that represented by Seq ID No. 34), approximately 12 Kb downstream of lambda, including the enhancer region (such as that represented by Seq ID No. 35), approximately 17.6 Kb downstream of lambda (such as that represented by Seq ID No. 36), approximately 19.1 Kb downstream of lambda (such as that represented by Seq ID No. 37), approximately 21.3 Kb downstream of lambda (such as that represented by Seq ID No. 38), and/or approximately 27 Kb downstream of lambda (such as that represented by Seq ID No. 39).

In still further embodiments, isolated nucleotide sequences as depicted in Seq ID Nos 28, 31, 32, 33, 34, 35, 36, 37, 38, or 39 are provided. Nucleic acid sequences at least 80, 85, 90, 95, 98 or 99% homologous to Seq ID Nos 28, 31, 32, 33, 34, 35, 36, 37, 38, or 39 are also provided. In addition, nucleotide sequences that contain at least 10, 15, 17, 20, 25, 30, 40, 50, 75, 100, 150, 200, 250, 500 or 1,000 contiguous nucleotides of Seq ID Nos 28, 31, 32, 33, 34, 35, 36, 37, 38, or 39 are provided. Further provided are nucleotide sequences that hybridizes, optionally under stringent conditions, to Seq ID Nos 28, 31, 32, 33, 34, 35, 36, 37, 38, or 39, as well as, nucleotides homologous thereto.

Seq ID No.28	CCTTCCTCCTGCACCTGTCAACTCCCAATAAACCGTCCTCCTTGTCATTC AGAAATCATGCTCTCCGCTCACTTGTGTCTACCCATTTTCGGGGCTTGCAT GGGGTCATCCTCGAAGGTGGAGAGAGTCCCCCTTGGCCTTGGGGGAAGTCG AGGGGGGCGGGGGGAGGCCTGAGGCATGTGCCAGCGAGGGGGGTACCTC CACGCCCCTGAGGACCTTCTAGAACCAGGGGCGTGGGGCCACCGCCTGAG TGGAAGGCTGTCCACTTTTCCCCCGGGCCCCCAGGCTCCCTCCTCCGTGT GGACCTTGTCCACCTCTGACTGGCCCAGCCACTCATGCATTGTTTCCCCG AAACCCACAGGACGATAGCTCAGCACGCGACAGTGTCCCCCTCTGAGGGCC TCTGTCCATTTTCAGGACGACCCGCATGTACAGCGTGACCACTCTGCTCAC GCCCCTCACCACGTCTAGAGCCCCACCCCCAGCCCCATCCTTAGGGGGC ACAGCCAGcTCCGACCGCCCCGGGGACACCACCTCTGCCCTTcCCCAG GCCCTCCCTGTACACGCACCAACAGGGGCCCTCCGTCCCGAGACCCTGCTC CCTCATCCCTCGGTCCCCCTCAGGTAGCCTTCCACCCGCGTGTGTCCCGAG GTCCCAGATGCAGCAAGGCCCTGGGACAACGCCAGATCTCTGCTCTcCC CGACCCCTCAGAAGCCAGCCACGCCTGGCCCCACCACCACTGCCTAACg TCCAAGTGTCCATAGGCCTCGGGACCTCCAAGTCCAGGTTCTGCCTCTGG GATTCCGCCATGGGTCTGCCTGGGAAATGATGCACTTGGAGGAGCTCAGC ATGGGATGCGGGACCTTGTCTCTAGGCGCTcCCTCAGGATCCCACAGCTG CCCTGTGAGACACACACACACACACACACACACACACACACACACACA CACACAAACACGCATGCACGCACGCCGGCACACACGCTATTGCAGAGATG GCCACGGTAGCTGTGCCTCGAGGGCCGAGTGGAGTGTCTAGAACTCTCGGG GGTCCCCTCTGCAGACGACACTGCTCCATCCCCCCCCGTGCCCTGAAGGGC TCCTCACTCTCCCATCAGGATCTCTCCAAGCTGCTGACCTGGAGAGGAAG GGGCCTGGGACAGGCGGGGACACTCAGACCTCCCTGCTGCCCTCCTCTG CCTGGGCTTGGACGGCTCCCCCTTCCCACGGGTGAAGGTGCAGGTGGGG AGAGGGCACCCCCCTCAGCCTCCCAGACCCAGACCAGCCCCCGTGGCAGG GGCAGCCTGTGAGCCTCCAGCCAGATGCAGGTGGCCTGGGGTGGGGGGTG GAGGGGGCGGGAGGTTTATGTTTGAGGCTGTATCACTGTGTAATATTTTC GGCGGTGGGACCCATCTGACCGTCTCGGTGAGTCTCCCCCTTTCTCTCC TCCTTGGGGATCCGAGTGAAATCTGGGTGATCTTCTCTCCGTTCTCCTC CGACTGGGGCTGAGGTCTGAACCTCGGTGGGGTCCGAAGAGGAGGCCCT AGGCCAGGCTCCTCAGCCCCCTCAGCCCCGACcgGCCCTCTTGACACAGGG TCCAGCTAAGGGCAGACATGGAGGCTGCTAGTCCAGGGCCAGGCTCTGAG ACCCAAGGGCGCTGCCCAAGGAACCTTGCCCCAGGGACCTGGGAGCAA AGCTCCTCACTCAGAGCCTGCAGCCCTGGGGTCTGAGGACAAGGAGGGAC TGAGGACTGGGCGTGGGGAGTTCAGGCGGGGACACCAGGTCCAGGGAGGT GACAAAGGCGCTGGGAGGGGGCGGACGGTGCCGGGGACTCCTCCTGGGCC CTGTGGGCTCGGGGTCTTGTGAGGACCTGAGGGACTGAGGGGGCCCTG GGCCTAGGGACTTGCAgTgAGGGAGGCAGGGAGTGTCCCTTGAGAACGTG GCCTCCGCGGGCTGGGTCCCCCTCGTGCTCCCAGCC*GGGAGGACACCCC AGAGCAAGCGCCCCAGGTGGGCGGGGAGGGTCTCCTCACAGGGGCGAGCTG ACAGATAGAGGCCCCCGCCAGGCAGATGCTTGATCCTGGCAgTTATACTG GGTTC**GCACAACTTTCCCTGAACAAGGGGCCCTCCGAACAGACACAGA CGCAACCCAGTCGACCcaggCTCAGCACAgAAAATGCACTGACACCCAAA ACCCTCATCTggggGCCTGGCCGGcAtCCCGCCCCAGGACCCAAGGCCCC TGCCCCCTGGCAGCCCTGGACACGGTCTCTGTGGGCGGTGGGGTCgGGG CTGTGGTGACGGTGGCATCGGGGAGCCTGTGCCCCCTCCCTGAAAGGGCG GAGAGGCTCAAGAGGGGACAGAAATGTCCTCCCCTAGGAAGACCTCGGAC GGGGGCGGGGGGGTGGTCTCCGACAGACAGATGCCCGGGACCGACAGACC TGCCGAGGGAAGAGGGCACCTCGGTGCGGTAGGCTCCAGGCAGCACGAG GGAGCGAGGCTGGGAGGGTGAGGACATGGGAGCCTGAGGAGGAGCTGGAG ACTTCAGCAGGCCCCCAGCTCCGGGCTTCGGGCTCTGAGATGCTCGGACG CAAGGTGAGTGACCCACCTGTGGCTGACCTGACCTCAgGGgGACAAGGC
--------------	---

TCAGCCTGGGACTCTGTGTCCCCATCGCCTGcACAGGGGATTCCCCTGAT
GGACACTGAGCCAACGACCTCCCGTCTCTCCCCGACCCCCAGGTCAGCCC
AAgGCCaCTCCACGGTCAACCTCTTCCCGCCCTCCTCTGAGGAGCTCGG
CACCAACAAGGCCACCCTGGTGTGTCTAATAAGTGACTTCTACCCGGGCG
CCGTGACGGTGACCTGGAAGGCAGGCGGCACCACCGTCACCCAGGGCGTG
GAGACCACCAAGCCCTCGAAACAGAGCAACAACAAGTACGCGGCCAGCAG
CTACCTGGCCCTGTCCGCCAGTGACTGGAAATCTTCCAGCGGCTTCACCT
GCCAGGTCACCCACGAGGGGACCATTGTGGAGAAGACAGTGACGCCCTCC
GAGTGCGCCTAGGTCCCTGGGCCCCCACCCTCAGGGGCCTGGAGCCACAG
GACCCCCGCGAGGGTCTCCCCGCGACCCTGGTCCAGCCCAGCCCTTCCTC
CTGCACCTGTCAACTCCCAATAAACCGTCCTCCTTGTCAATTCAGAAATCA
TGCTCTCCGCTCACTTGTGTCTACCCATTTTCGGGCTTGCATGGGGTCAT
CCTCGAAGGTGGAGAGAGTCCCCCTTGGCCTTGGGgAAATCGAGGGGGGC
GGGGGGAGGCCTGAGGCATGTGCCAGCGAGGGGGGTACCTCCACGCCCC
TGAGGACCTTCTAGAACCAGGGGCGTGGGGCCACCGCCAGAGTGGAAGGC
TGTCCACTTTTCCCCCGGGCCCCCAGGCTCCCTCCTCCGTGTGGACCTTG
TCCACCTCTGACTGGCCCAGCCACTCATGCATTGTTTCCCCGAAACCCCA
GGACGATAGCTCAGCACGCGACAGTGTCCTCCCTCTGAGGGCCTCTGTCCA
TTTCAGGACGACCCGCATGTACAGCGTGACCACTCTGCTCACGCCCACTC
ACCACGTCTAGAGCCCCACCCCCAGCCCCATCCTTAGGGGCACAGCCAG
CTCCGACCGCCCCGGGGACACCACCCTCTGCCCTTCCCCAGGCCCTCCC
TGTCACACGCACCACAGGGCCCTCCGTCCCGAGACCCTGCTCCCTCATCC
CTCGGTCCCCTCAGGTAGCCTTCCACCCGCGTGTGTCCCGAGGTCCCAGA
TGCAGCAAGGCCCTGGGACAACGCCAGATCTCTGCTCTCCCCGACCCTC
AGAAGCCAGCCCACGCCTGGCCCACCACCACTGCCTAACGTCCAAGTGTC
CATAGGCTCGGGAcCTCcAaGTCCAGGTTCTGCCTCTGGGATTCCGCCAT
GGGTCTGCCTGGAATGATGCACTTGGAGgAgCTCAGcATGGGATGcGGAA
CTTGTCTAGcGCTCCTCAGATCCAcAGcTGCTGTGAgAcacacacac
acacacacacacAAAcAGcATGCACGCACGCCGGCACACACGCTATTA
CAGAGATGGCCACGGTAGCTGTGCCTCGAGGCCGAGTGGAGTGTCTAGAA
CTCTCGGGGGTCCCCTCTGCAGACGACACTGCTCCATCCCCCCCCGTGCC
TGAAGGGCTCCTCACTCTCCCATCAGGATCTCTCCAAGCTGCTGACCTGG
AGAGGAAGGGGCCTGGGACAGGCGGGGACACTCAGACCTCCCTGCTGCCC
CTCCTCTGCCTGGGCTTGGACGGCTCCCCCTTCCCACGGGTGAAGGTGC
AGGTGGGGAGAGGGCACCCCCCTCACCTCCCAGACCCAGACCAGCCCCC
GTGGCAGGGGCAGCCTGTGAGCCTCCAGCCAGATGCAGGTGGCCTGGGGT
GGGGGGTGGAGGGGGCGGGAGGTTTATGTTTGAGGCTGTATTCATCTGTG
TAATATtTCGGCGGTGGGACCCATCTGACCGTCCTCGGTGAGTCTCCCC
TtttcttctcttgggatccgagtgaATcTGGGTCGATCTTCTCTC
CGTTCTCCTCCGACTGGGGCTGAGGTCTGAACCTCGGTgGGGTCCGAAGA
GGAGGCCCTAGGCC*GGCTCcTCAGCCCCCTCCAGCCCGACCCGCCCTCT
TGACACAGGGTCCAGCTAAGGGCAGACAT***GGCTGCTAGTCCAGGGCC
AGGCTcTGAGACCCAAGGGCGCTGCCCAAGGAACCCTTGCCCCAGGGACC
CTGGGAGCAAAGCTCCTCACTCAGAGCCTGCAGCCCTGGgGTCTGAGGAC
AAGGAGGGACTGAGGACTGGGCGTGGGGAGTTTCAGGCgGGGACACCGGGT
CCAGGGAGGTGACAAAGGCGCTGGGAGGGGGCGGACGGTGCCGGAGACTC
CTCCTGGGCCCTGTGGGCTCGTGGTCCTTGTGAGGACCCTGAGGG*CTGA
GGGGCCCCCTGGGCCTAGGGACTTGCAGTGAGGGAGGCAGGGAGTGTCCCT
TGAGAACGTGGCCTCCGCGGGCTGGGTCCCCCTCGTGCTCCCAGCAGGGA
GGACACCCAGAGCAAGCGCCCCAGGTGGGCGGGGAGGGTCTCCTCACAG
GGGCAGCTGACAGATAGAC*GgccCCCGCCAGACAGATGCTTGATCCTGG
TCag***TACTGGGTTTCGcACTTCCCTGAACAGGGGCCCTCCGAACAGA
CACAGACGCAGACCaggCTCAGCACAgAAAATGCACTGACACCCAAAACC
CTCATCTGgGGCCTGGCCGGCATCCCGCCCCAGGACCCAAGGCCCTGC
CCCCTGGCAGCCCTGGACACGGTCCTCTGTGGGCGGTGGGGTCgGGGCTG
TGGTGACGGTGGCATCGGGGAGCCTGTGCCCCCTCCCTGAAAGGGCGGAG

	AGGCTCAAGAGGGGACAGAAATGTCCTCCCCTAGGAAGACCTCGGACGGG GGCGGGGGGGTGGTCTCCGACAGACAGATGCCCGGGACCGACAGACCTGC CGAGGGAAGAGGGGCACCTCGGTTCGGGTAGGCTCCAGGCAGCACGAGGGA GCGAGGCTGGGAGGGTGAGGACATGGGAGCCTGAGGAGGAGCTGGAGACT TCAGCAGGCCCCCAGCTCCGGGCTTCGGGCTCTGAGATGCTCGGACGCAA GGTGAGTGACCCACCTGTGGCTGACCTGACCTGACCTCAGGGGGACAAG GCTCAGCCTGGGACTCTGTGTCCCCATCGCCTGCACAGGGGATTCCCCTG ATGGACACTGAGCCAACGACCTCCCGTCTCTCCCCGACCCCCAGGTCAGC CCAAGGCCACTCCCACGGTCAACCTCTTCCCGCCCTCCTCTGAGGAGCTC GGCACCAACAAGGCCACCCTGGTGTGTCTA
--	--

Seq ID No.32	
	GCCACGCCCCACTCCATCATGCGGGGAGGGGATGGGCAGACCCTCCAGAAA GAAGCTCCCTGGGGTGCAGGTAAACAGCTTTCCAGACACAGCCAGTACT AGAGTGAGGTGAATAAGACATCCTCCTTGCTTGTAATTTAGGAAGTGC CCCCAACATCAGTCATTAAGATAAATAATATTGAATGCACTTTTTTTT TTTATTTTTTTTTTTTGGCTTTTTAGGGCCTAATCTGCAGCatatggaagt tcccaggctacaagtcgaaccagagctgcagctgccagcctacatcacag ccacagcaacaccagatccgagccacatctgtgactaacactgcagtca cagcaacgccagatccttaacccattgagtgaggccaggatcaaaccca catcctcatggatactagtctgttcgtaaacactgagccaCAAGGGGA ACTCCTGAATGCAATATTTTTGAAAATTGAAATTAAATCTGTCACTCTTT CACTTAAGAGTCCCCTTAGATTGGGGAAAATTTAAATATCTGTCATCTTA GTGCATCTTTGCTCATATGATGTGAATAAAATCCCAAAATCCATATGAAT GAAGCATCAAAATGTACATGAAGTCAGCCTGACCCTGCACTGCCCTCACT TGCCTCATGTACCCCCACCTCAAAGGAAGATGCAGAAAGGAGTCCAGCC CCTACACCGCCACCTGCCCCCACCCTGGAGCCCCCTCAGGTCTCCCACCT CCTTTTCTGAGCTTCAGTCTTCCTGTGGCATTGCCTACCTCTACAGCTGC CCCCTACTAGGCCCTCCCCCTGGGGCTGAGCTCCAGGCACTGGACTGGGA AAGTTAGAGGTTAAAGCATGGAAAATTCCCAAAGCCACCAGTTCCAGGCT GCCCCCACCACCGCCACGTCCAAAAGGGGCATCTTCCCAGATCTCT GGCTGGTATTGGTAGGACCCAGGACATAGTCTTTATACCAATTCTGCTGT GTGTCTTAGGAAAGAAactctccctctctgtgcttcagtttctcatcaa taaaAGGAGCAGGCCAGGTTGGAGGGTCTGTGACGTCTGCTGAAGCAGCA GGATTCTCTCTCCTTTTGCTGGAGGAGAACTGATCCTTCACCCCCAGGAT CAACAGAGAAGCCAAGGTCTTCAGCCTTCCTGGGGACCCCTCAGAGGGAA CTCAGGGCCACAGAGCCAGACCCTGATGCCAGAACCTTTGTCATATGCCC AGACGGAGACTTCATCCCCCTCCTCCTCAGACCCTCCAGGCCCAACAGT GAGATGCTGAAGATATTAAGAGAAGGGCAAGTCAGcTTAAGTTTGGGGGT AGAGGGGAACAGGGAGTGAGGAGATCTGGCCTGAGAGATAGGAGCCCTGG TGGCCACAGGAGGACTCTTTGGGTCTGTGCGGATGGACACAGGGGCGGCC GGGGGCATGTTGGAGCCCGGCTGGTTCTTACCAGAGGCAGGGGGCACCT CTGACACGGGAGCAGGGCATGTTCCATACATGACACACCCCTCTGCTCCA GGGCAGGTGGGTGGCGGCACAGAGGAGCCAGGGACTCTGAGCAAGGGGTC CACCAGTGGGGCAGTTGGATCCAGACTTCTCTGGGCCAGCGAGAGTCTAG CCCTCAGCCGTTCTCTGTCCAGGAGGGGGGTGGGGCAGGCCTGGGCGGCC AGAGCTCATCCCTCAAGGGTTCCAGGGTCTGCCAGACCCAGATTTCG ACCGCAGCCACCACAAGAGGATGTGGTCTGCTGTGGCAGCTGCCAAGACC TTGCAGCAGGTGCAGGGTGGGGGGGTGGGGGCACCTGGGGGCAGCTGGGG TCACTGAGTTCAGGGAAAACCCCTTTTTTCCCTAAACCTGGGGCCATCC CTAGGGGAACCAACAACCTTCTGAGCCCTGGGCAGTGGCTGCTGGGAGGGA AGAGCTTCATCCTGGACCCTGGGGGGGAACCCAGCTCCAAGGTGCAAGG GGCCAGGTCCAAGGCTAGAGTGGGCCAAGCACCGCAATGGCCAGGGAGT GGGGGAGGTGGAGCTGGACTGGATCAGGGCCTCCTTGGGACTCCCTACAC

CCTGTGTGACATGTTAGGGTACCCACACCCCATCACCAGTCAGGGCCTGG
 CCCATCTCCAGGGCCAGGGATGTGCATGTAAGTGTGTGTGAGTGTGTGTG
 TGTGGTGTAGTACACCCCTTGGCATCCGGTTCCGAGGCCTTGGGTTCTC
 CAAAGTTGCTCTCTGAATTAGGTCAAACCTGTGAGGTCCTGATCGCCATCA
 TCAACTTCGTTCTCCCCACCTCCCATCATTATCAAGAGCTGGGGAGGGTC
 TGGGATTTCTTCCCACCCACAAGCCAAAAGATAAGCCTGCTGGTGTATGGC
 AGAAGACACAGGATCCTGGGTGAGAGACAAAGGCCAGTGTGTACAGCGA
 GAGAGGCAGCCGACTATCAGCTGTACAGAGAGGCTTAGTCCGCTGAA
 CTCAGGCCCCAGTGACTCCTGTTCCACTGGGCACTGGCCCCCTCCACAG
 CGCCCCCAGGCCCCAGGGAGAGGCGTCACAGCTTAGAGATGGCCCTGCTG
 AACAGGGAACAAGAACAGGTGTGCCCCATCCAGCGCCCCAGGGGTGGGAC
 AGGTGGGCTGGATTTGGTGTGAAGCCCTTGAGCCCTGgAACCCAAcCACA
 GCagGGCAGTTGGTAGATGCCATTTGGGGAGAGGCCCCAGGAGTAAGGGC
 CATGGGCCCTTGAGGGGGGCCAGGAGCTGAGGACAGGGACAGAGACGGCCC
 AGGCAGAGGACAGGGCCATGAGGGGTGCACTGAGATGGCCACTGCCAGCA
 GGGGCAGCTGCCAACCCGTCCAGGGAACCTTATTCAGCAGTCAGCTGGAGG
 TGCCATTGACCCTGAGGGCAGATGAAGCCCAGGCCAGGCTAGGTGGGCTG
 TGAAGACCCCAGGGGACAGAGCTCTGTCCCTGGGCAGCACTGGCCTCTCA
 TTCTGCAGGGCTTGACGGGATCCCAAGGCCTGCTGCCCTGATGGTAGTG
 GCAGTACCGCCCAGAGCAGGACCCCAGCATGGAAACCCCAACGGGACGCA
 GCCTGCGGAGCCCAAAAACAGTAAGGAGCCGAAGCAGTCATGGCACGG
 GGAGTGTGGACTTCCCTTTGATGGGGCCAGGCATGAAGGACAGAATGGG
 ACAGCGGCCATGAGCAGAAAATCAGCCGGAGGGGATGGGCCTAGGCAGAC
 GCTGGCTTTATTTGAAGTGTGGCATTITGTCTGGTGTGTATTGTTGGTA
 TTGATTTTATTTTAGTATGTCAGTGACATACTGACATATTATGTAACGAC
 ATATTATTATGTGTTTTAAGAAGCACTCCAAGGGAACAGGCTGTCTGTAA
 TGTGTCCAGAGAAGAGAGCAAGAGCTTGGCTCAGTCTCCCCAAGGAGGT
 CAGTTCCTCAACAGGGGTCCTAAATGTTTCTGGAGCCAGGCCTGAATCA
 AGGGGgTCATATCTACACGTGGGGCAGACCCATGGACCATTITTCGGAGCA
 ATAAGATGGCAGGGAGGATACCAAGCTGGTCTTACAGATCCAGGGCTTTG
 ACCTGTGACGCGGGCGCTCCTCCAGGCAAAGGGGAGAAGCCAGCAGGAAGC
 TTTCAGAACTGGGGAGAACAGGGTGCAGACCTCCAGGGTCTTGTACAACG
 CACCCTTTATCCTGGGGTCCAGGAGGGGTCACTGAGGGATTTAAGTGGGG
 GACCATCAGAACCAGGTTTGTGTTTTGGAAAAATGGCTCCAAAGCAGAGA
 CCAGTGTGAGGCCAGATTAGATGATGAAGAAGAGGCAGTGGAAGTCGAT
 GGGTGGCCAGGTAGCAAGAGGGCCTATGGAGTTGGCAAGTGAATTTAAAG
 TGGTGGCACCAGAGGGCAGATGGGGAGGAGCAGGCACTGTCATGGACTGT
 CTATAGAAATCTAAAATGTATACCCTTTTTAGCAATATGCAGTGAGTCAT
 AAAAGAACACATATATATTTAAATTGTGTAATTCACCTTCTAAGGATTCA
 TCCCAAGGGGGGAAAATAATCAAAGATGTAACCAAAGGTTTACAAACAAG
 AACTCATCATTAAATCTTCCTTGTGTTATTTCAACGATATTATTATTATT
 ACTATTATTATTATTATTATTttgtcttttgcattttctagggccactc
 ccacggcatagagaggttcccaggctaggggtcaaatcgagctacagct
 gccggcctacgccagagccacagcaacgcaggatctgagccacagcaatg
 caggatctacaccacagctcatgtaacgctggatccttaaccaatgag
 tgaggccagggatcgaacctgtaacttcattggttcttagtcggattcatt
 aacctgagccacgacaggaactccAACATTATTAATGATGGGAGAAAA
 CTGGAAGTAACCTAAATATCCAGCAGAAAGGGTGTGGCCAAATACAGCAT
 GGAGTAGCCATCATAAGGAATCTTACACAAGCCTCCAAAATTGTGTTTCT
 GAAATTGGGTTTAAAGTACGTTTGCATTTTAAAAAGCCTGCCAGAAAATA
 CAGAAAAATGTCTGTGATATGTCTCTGGCTGATAGGATTTTGCTTAGTTT
 TAATTTTGGCTTTATAATTTTCTATAGTTATGAAAATGTTTACAAGAAGA
 TATATTTTCAATTTAGCTTCTAAAATAATTATAACACAGAAGTAATTTGTG
 CTTTAAAAAAATATTCAACACAGAAGTATATAAAGTAAAAATTGaggagt
 tcccatgctggctcagtgattaacaaaccaactagtatccatgaggata
 tggatttgatccctggccttgctcagtgagggttgaggatccagtggtgctg

tgagctgtggtgtaggtgcagacacagcactctggcgttgctgtgactc
tggcgtaggcccggcagctacagctccatttggacccttagcctgggaacc
tccatatgcctgagatacgcccTAAAAAGTCAAAAGCCAAAAAATAGT
AAAAATTGAGTGTTTCTACTTACCACCCCTGCCACATCTTATGCTAAAA
CCCGTTCTCCAGAGACAAACATCGTCAGGTGGGTCTATATATTTCCAGCC
CTCCTCCTGTGTGTGTATGTCCGTAAAAACACACACACACACACACG
CACACACACACACACGTATCTAATTAGCATTGGTATTAGTTTTTCAAAG
GGAGGTCATGCTCTACCTTTTAGGCGGCAAATAGATTATTTAAACAAATC
TGTTGACATTTTCTATATCAACCCATAAGATCTCCCATGTTCTTGAAAG
GCTTTGTAAGACATCAACATCTGGGTAAACCAGCATGGTTTTTAGGGGGT
TGTGTGGATTTTTTTCATATTTTTTAGGGGCACACCTGCAGcatatggagg
ttcccaggctaggggtgaatcagagctgtagctggcgccctacaccaca
gccacagcaacgccagatccttaaccactgagaaaggccagggttgaa
cctgcacctcatggATGCTGGTCAGATTTATTTCTGCTGAGCCACAACA
GGAACCTCCCTGAACCAGAATGCTTTTAACCATTCACCTTTGCATGGACAT
TTAGATTGTTTCCATTTAAAAATACAAATTACAaggagtcccgctgtagg
ctcagtggtaacgaattggactaggaacctagaggttcgggttcgatcc
ctggccttgctcggtgggttaaggatccagcatgatgtgagatatggtg
taggtcgcagacgtggctcggtacccacgttgctgtggctctggcgtagg
ccggcaacaacagctccgattcgacccttagccTGggaacctccatgtgc
cacaggagcagccctaGAAAAGGCCAAAAAGACAAAAAATAAAAAATTAA
AATGAAAAAATAAAATAAAAAATACAAATTACAAGAGACGGCTACAAGGAA
ATCCCCAAGTGTGTGCAAAATGCCATATATGTATAAAATGTACTAGTGTCT
CCTCGCGGGAAAGTTGCCTAAAAGTGGGTGGCTGGACAGAGAGGACAGG
CTTTGACATTCTCATAGGTAGTAGCAATGGGCTTCTCAAATGCTGTTCC
AGTTTACACTCACCATAGCAAATGACAGTGCCTCTTCTCTCCACCCTTG
CCAATAATGTGACAGGTGGATCTTTTTCTATTTTGTGTATCTGACAAGCA
AAAAATGAGAACAggagttcctgtcgtggtgcagtggagacaaatctgac
taggaacctgaaatttcgggtcaatccctggcctcactcagtaggtaa
aggatccagggttgacgtgagctgtgggttaggtcgcagacacagtgcaa
atttggccctgttggtgtggtgtaggcccggcagctatagctccaatt
ggacccttagcctgggaacctccttatgccgtgggtgaggccctAAAAAA
AAGAGTGCAAAAAAATAAAGAACAAAAATGATCATCGTTTAATTCT
TTATTTGATCATTGGTGAACTTATTTTCTTTTATATTTTATTGACTG
ATTTTATTTCTCCTATGAATTTACCGGTCATAGTTTTGCCTGGGTGTTTT
TACTCCGGTTTTAGTTTTGGTTGGTTGTATTTTCTTAGAGAGCTATAGAA
ACTCTTCATCTATTTGGAATAGTAATTCCTCATTAAGTATTTGTGCTGCA
AAAAATTTTCCCTGATCTGTTTTATGCTTTTGTGTGGGGTCTTTCACG
AGAAAGCCTTTTAGTTTTACACCTCAGCTTGGTTGTTTTCTTGATTG
TGTCTGTAATCTGCGGCCAACATAGGAAACACATTTTTACTTTAGTGTTT
TTTTCTATTTTCTTCAAGTACGTCCATTGTTTTGGTGTCTGATTTTACT
TTGCCTGGGGTTTTGTTTTTGTGTGGCAGGAATATAAACTTATGTATTTT
CAAATGGAGAGCCAATGGTTGTATATTTGTTGAATTCAAATGCAACTTTA
TCAAACACCAAATCATCGATTTATCACAACCTCTTCTCTGGTTTATTGATC
TAATGATCAATTCCTGTTCCACGCTGTTTTAATTATTTTAGCTTTGTGGA
TTTTGGTGCCTGGTAGAGAACAAAGCCTCCATTATTTTCATTCAAATAG
TCCCGTCTATTATCTGCCATTGTTGTAGTATTAGACTTTAAATCAATTT
ACTGATTTTCAAAGTTATTCCTTTGGTGATGTGGAATACTTTATACTTC
ATAAGGTACATGGATTCAATTGTGGGGAATTGATGTCTTTGCTATTGTGG
CCATTTGTCAAGTTGTGTAATATTTTACCCATGCCAACTTTGCATATTGT
ATGTGAGTTTATCCCAGGGTTTTTAATAGGATGTTTATTGAAGTTGTCA
GTGTTTCCACAATTTTCATCGCCTCAGTGCTTACTGTTTGCATAAAAGGAA
ACCTACTCACTTTTGCCTATTGCTCTTGTATTCAATCATTTTAGTTAACT
CTTGTGTTAATTTTGTAGAGTTTTTTCAGCTGACTGTCTGGGGTTTTCTTTA
ATAGACTAGCCCTTTGTCTGTAAAGAATAATTTTATCGAATTTTTCTTAA
CACTCACACTCTCCCCACCCCCACCCCCGCTCATCTCCTTTTCATTGGGTC

	AAATCTGTAGAATACAATAAAAGTAAGAGTGGGAACCTTAGCCTTTAAGT CGATTTTGCCTTTAAATGTGAATGTTGCTATGTTTCGGGACATTCTCTTT ATCAAGTTGCGGATGTTTCCTTAGATAATTAACCTTAATAAAAGACTGGAT GTTTGCTTTCTTCAAATCAGAATTGTGTTGAATTTATATTGCTATTCTGT TTAATTTTGTTCAAAAAATTTACATGCACACCTTAAAGATAACCATGAC CAAATAGTCCTCCTGCTGAGAGAAAATGTTGGCCCCAATGCCACAGGTTA CCTCCCGACTCAGATAAACTACAATGGGAGATAAAATCAGATTTGGCAAA GCCTGTGGATTCTTGCCATAACTCTCAGAGCATGACTTGGGTGTTTTTTC CTTTTCTAAGTATTTTAATGGTATTTTGTGTTACAATAGGAAATCTAGG ACACAGAGAGTGATTCAATGAGGGGAACGCATTCTGGGATGACTCTAGGC CTCTGGTTTGGGGAGAGCTCTATTGAAGTAAAGACAATGAGAGGAAGCAA GTTTGCAGGGAAGTGTGAGGAATTTAGATGGGGAATGTTGGGTTTGAGGT TTCTATAGGGCACGCAAGCAGAGATGCACTCAGGAGGAAGAAGGAGCATA AATCTAGAGGCAAAAAGAGAGGTCAGGACTGGAAATAGAGATGCGAGACA CCAGGGTGGCAGTCAGAGAGCACAGTGTGGGTGAGAAGACAGTGAAGAA CACAAGGGACAGAGAGGGATCTCCAACCTTCACTGGGATGAGGGCCTTGTT GGCCTTGACCTGAGAGATTTCCAGGAGTTGAGGGTGGGAAGGAGAGGGCT CCTGCACATGTCCTGACATGAAACGGTGCCCAGCATATGGGTGCTTGGA GACATTGTTGGACAGATGGATGGATGATGGATGATGGATGAATGGATGGA TGGAAGATGATGGATAAATGGATGATGGATGGATGGACAGAAGGACAAAG AGATGGACAGAAAGACAGTGATCTGAGAGAGCAGAGAAGGCTTCATGAAA GGACAGGAACTGAACTGTCTCAGTGGGTGGAGACAATGGTGTAGGGGGTT TCCACATGGAGGCACCAGGGGTGAGGAATAATCTAGTGTCCACAGGCCCA GGAAGGAAGCTGTCTGCAGGAAATTGTGGGGAAGAACCTCAGAGTCCTTA AATGAGGTCAGGAGTGGTCAGGAGGGTCTGATCAGGTAAGGACTCATGTC CATCATCACATGGTCACCTAAGGGCATGTAGCTCTCAGCATCTCCATCAG GACAGTCTCAGAATGGGGGCGGGGTACACACTGGGTGACTCAAGGCGTG GGTCATGCCTGCCTCGGACGTGGGCCTGGGCATGGGGACACCTCCAGACC ATGGGCCCCGCCAGGGCTGCACTGGcctctggtgggtagctaccgtec aagcaacacaggacacagccctacgtgcaaccctgtgcccgaacgc ccatctggttctgtctccagcccggccccaggaacaggactcaggtgct agcccaatggggtttgttcgagcctcagtcagcgtggTATTTCTCCGGC AGCGAGACTCAGTTCACCGCCTTAGGttaagtgggttctcatgaattcct agcagtcctgcactctgtatgccgggaaagtcactttgtcgtggggg ctgttccccgtgcccttgagaatcaaggattgcccactttctctgtg ggggaggtggctggtcttgggggtaccagcaggaaggggcccaaaagcag gagcagctgcctccagAATACAACTGTGCGCTACAGCTCAAACAGGAGGC CTGGACTGGGGTTTAACCACCAGGGCGGCACGAAGGAGCGAGGCTGGGAG GGTGAGGACATGGGAGCCTGAGGAGGAGCTGGAGACTTCAGCAGGCCCCC AGCTCCGGGCTTCGGGCTCTGAGATGCTCGGACGCAAGGTGAGTGACCCC ACCTGTGGCTGACCTGACCTCAGGGGGACAAGGCTCAGCCTGAGACTCTG TGTCCCCATCGCCTGCACAGgggattcccctgatggacactgagccaacg acctccgtctctccccgacccccaggtcagcccaaggccgccccacgg tcaacctcttccccgccctctctgaggagctcggcaccaacaaggccacc ctggtgtgttaataagtgaacttctaccgAAGGGCGAATTCCAGCACAC TGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACCAAGCTTGATGCATAG CTTGAGTATCTA
--	--

Seq ID No.33	agatctttaaccaccgagcaaggccaggatgaaccgcatcctcatg aatcctagtgggttcgtaaccgctgaaccacaatgggaactcctGTCT TTCACATTTAATTCAACCTCTCCAGGATTCTGGGGGTGGGTGGGGAAT CCTAGGTACCCACTGGGAAAGTAATCCAAGGGGAGAGGCTCACGGACTcT AGGGATCGGCGGAGGAGGGAAGGTATCTCCAGGAACTGGCCAGGACAC ATTGGTCCTCCGCCCTCCCCTTCCCTCCACTCCTCCTCCAGACAGGACTG

TGCCACCCCCTGCCACCTTTCTGGCCAGAACTGTCCATGGCAGGTGACC
 TTCACATGAGCCCTTCCTCCCTGCCTGCCCTAGTGGGACCCTCCATACCT
 CCCCCTGGACCCCGTTGTCCTTTCTTTCCAGTGTGGCCCTGAGCATAACT
 GATGCCATCATGGGCTGCTGACCCACCCGGGACTGTGTTGTGCAGTGAGT
 CACTTCTCTGTCATCAGGGCTTTGTAATTGATAGATAGTGTTTCATCATC
 ATTAGGACCGGGTGGCCTCTATGCTCTGTTAGTCTCCAAACACTGATGAA
 AACCTTCGTTGGCATAGTCCCAGCTTCCTGTTGCCCATCCATAAATCTTG
 ACTTAGGGATGCACATCCTGTCTCCAAGCAACCACCCCTCCCCTAGGCTA
 ACTATAAACTGTCCCAATGGCCCTTGTGTGGTGCAGAGTTCATGCTTCC
 AGATCATTTCTCTGCTAGATCCATATCTCACCTTGTAAGTCATCCTATAA
 TAACTGATCCATTGATTATTTGCTTCTGTTTTTCCATCTCAAAACAGC
 TTCTCAGTTCAGTTCGAATTTTTTATCCCTCCATCCACCCATACTTTCC
 TCAGCCTGGGGAACCCTTGCCCCCAGTCCCATGCCCTTCCTCCCTCTCTG
 CCCAGCTCAGCACCTGCCACCCCTCACCTTCCTGTCACTCCCTAGGACT
 GGACCATCCACTGGGGCCAGGACACTCCAGCAGCCTTGGCTTCATGGGCT
 CTGAAATCCATGGCCCATCTCTATTCCTCACTGGATGGCAGGTTCAGAGA
 TGTGAAAGGTCTAGGAGGAAGCCAGGAAGGAAACTGTTGCATGAAAGGCC
 GGCCTGATGGTTCAGTACTTAAATAATATGAGCTCTGAGCTCCCCAGGAA
 CCAAAGCATGGAGGGAGTATGTGCCTCAGAATCTCTCTGAGATTCAGCAA
 AGCCTTTGCTAGAGGGAAAATAGTGGCTCAACCTTGAGGGGCCAGCATCTT
 GCACCACAGTTAAAAGTGGGTATTTGTTTTACCTGAGGCCTCAGCATTAT
 GGGAAACCGGGCTCTGACACAAACACAGGTGCAGCCCGGCAGCCTCAGAAC
 ACAGCAACGACCACAAGCTGGGACAGCTGCCCCCTGAACGGGGAGTCCACC
 ATGCTTCTGTCTCGGGTACCACCAAGGTCAACATCCCTGGGGGAGGTAGTT
 CCATAGCAGTAGTCCCCTGATTTCCGCCCTCGGGCGTGTAGCCAGGCAAG
 CTCCTGCCTCTGGACCCAGGGTGGACCCCTTGCTCCCCACTACCCTGCACA
 TGCCAGACAGTCAAGACCACTCCACCTCTGTCTGAGGCCCCCTTGGGTG
 TCCCAGGGCCCCCGAGCTGTCCTCTACTCATGGTTCTTCCACCTGGGTAC
 AAAAGAGGCGAGGGACACTTTTCTCAGGTTTGCGGCTCAGAAAGGTACCT
 TCCTAGGGTTTGTCCACTGGGAGTCACCTCCCTTGCACTCTCAATGTCAGT
 GGGGAAAACCTGGGTCCCATGGGGGGATTAGTGCCACTGTGAGGCCCTGA
 AGTCTGGGGCCTCTAGACACTATGATGATGAGGGATGTGGTGAAAAACCC
 CACCCAGCCCTTCTTGCCGGGACCCTGGGCTGTGGCTCCCCCATTGCAC
 TTGGGGTCAGAGGGGTGGATGGTGGCTATGGTCAGGCATGTTTCCCATGA
 GCTGGGGGCACCCTGGGTGACTTTCTCCTGTGAATCCTGAATTAGCAGCT
 ATAACAAATTGCCCAAACCTTAGGCTTAAAACAACACACATTTATTCCT
 CTGGGTCCCAGGGTCAGAAGTCCAAAATGAGTCCTATAGGCTAAATTTGA
 GGTGTCTCTGGGTTGAGCTCCTCCTGGAAGCCTTTTCCAGCCTCTAGAGT
 CCCAAGTCCTTGCTCTGGGCCCCCTCCCTCAAGCTTCAAAGCCACAGAAG
 CTTCTAATCTCTCTCCCTTCCCCTCTGACCTCTGCTCCCATCCTCATACC
 CTGTCCCCTCACTCTGACCCTCCTGCCTCCCTCTTCCCTTATAAAGACC
 CTGCATGGGGCCACGGAGATAATCCAGGGTAATCGCCCCCTCTTCCAGCCC
 TTAATCCATCCCATCTGCAAAATCCCTGTCAACCCATAATGGACCTACT
 GATGGTCTGGGGGTTAGGACGTGGACAACCTTGGGGCCTTATTCATCTGAT
 CAAACTCCAGTTCCCAGACCCCCAGACCCCGGGCATTAGGGAAACTTC
 TCCCAGTTCCTCTCCCTCTGTGTCCTGCCAGTCTCCAGGATGGGCCACT
 CCCGAGGGCCCTTCAGCTCAGGCTCCCCCTCCTTTCTCCCTGGCCTCTTG
 TGGCCCCATCTCCTCCTCCGCTCACAGGGAGAGAACTTTGATTTAGCTT
 TGGCTCTGGGGCTTTGCTTCCCTTCTGGCCATTGGCTGAAGGGCGGGTTTC
 TCCAGGTCTTACCTGTCACTCATCAAACCGCCCTTGGAGGAAGACCCTAA
 TATGATCCTTACCCTACAGATGGAGACTCGAGGCCAGAGATCCTGAGTG
 ACCTGCTCACATTCACAGCAGGGACTGAACCCAGTCACCTACCCAACCTC
 CAGGGCTCAGCGCTTTTTTTTTTTTTTTTTCTTTTTgcctttcgagggcc
 gtccccgaacatatggagatttcaggctaggggtctaattggagcagt
 cgacactggcctaagccaaagccacagcaacaagggaagccgcttctgc
 agcctataccacagctcagggcaatgccggatccttaaccactgagcaa

	agccagggattgaacctgcaacctcatgttcttagtcaaattgttaac cactgacccatgacgggaactcccAGGGCTCAGCTCTTGACTCCAGGTTC GCAGCTGCCCTCAAAGCAATGCAACCCTGGCTGGCCCCGCCTCATGCATC CGGCCTCCTCCCCAAAGAGCTCTGAGCCCACCTGGGCCTAGGTCCTCCTC CCTGGGACTCATGGCCTAAGGGTACAGAGTTACTGGGGCTGATGAAGGGA CCAATGGGGACAGGGGCCTCAAATCAAAGTGGCTGTCTCTCTCATGTCCC TTCCTCTCCTCAGGGTCCAAAATCAGGGTCAGGGCCCCAGGGCAGGGGCT GAGAGGGCCTCTTTCTGAAGGCCCTGTCTCAGTGCAGGTTATGGGGGTCT GGGGGAGGGTCAATGCAGGGCTCACCTTCAGTGCCCCAAAGCCTAGAGA GTGAGTGCCTGCCAGTGGCTTCCCAGGCCCAATCCCTTGACTGCCTGGGA ATGCTCAAATGCAGGAAGTGTACAAACACCTTCAGTCAGGGGCTGCTCTG GGAGGAAAAACACTCAGAATTGGGGGTTCAGGGAAGGCCCAGTGCCAAGC ATAGCAGGAGCTCAGGTGGCTGCAGATGGTGTGAACCCAGGAGCAGGAT GGCCGGCACTCCCCCAGACCCTCCAGAGCCCCAGGTTGGCTGCCCTCTT CACTGCCGACACCCCTGGGTCCACTTCTGCCCTTTCCCACCTAAAACCTT TAGGGCTCCCACTTTCTCCCAAATGTGAGACATCACACGGCTCCCAGGG AGTGTCCAGAAGGGCATCTGGCTGAGAGGTCCTGACATCTGGGAGCCTCA GGCCCCACAATGGACAGACGCCCTGCCAGGATGCTGCTGCAGGGCTGTTA GCTAGGCGGGGTGGAGATGGGGTACTTTGCCTCTCAGAGGCCCCGGCCCC ACCATGAAACCTCAGTGACACCCCATTTCCCTGAGTTCACATACCTGTAT CCTACTCCAGTCACCTTCCCCACGAACCCCTGGGAGCCCAGGATGATGCT GGGGCTGGAGCCACGACCAGCCCACGAGTGATCCAGCTCTGCCAATCAGC AGTCATTTCCCAAGTGTTCAGCCCTGCCAGGTCCCCTACTACAGCAGTAAT GGAGGCCCCAGACACCAGTCCAGCAGTTAGAGGGCTGGACTAGCACCAGC TTTCAAGCCTCAGCATCTCAAGGTGAATGGCCAGTGCCCCTCCCCGTGGC CATCACAGGATCGCAGATATGACCCTAGGGGAAGAAATATCCTGGGAGTA AGGAAGTGCCATACTCAAGGATGGCCCCTCTGTGACCTAACCTGTCCCT GAGGATTGTACTTCCAGGCGTTAAACAGTAGAACGCCTGCCTGTGAACC CCCGCCAAGGGACTGCTTGGGGAGGGCCCCCTAAACCAGAACACAGGCACT CCAGCAGGACCTCTGAACTCTGACCACCCTCAGCAAGTGGCACCCCCCGC AGCTTCCAAGGCAC
--	--

Seq ID No.34	AACAAAGATGCTACCCACCAACAAAATTCACCGGAGAAGACAAGGACAGG GGGTTCTGGGGTCTGACAGGGTCACCAAAGAGGGTTCTGGGGCAGCAG CAACTCCAGCCGCCTCAGAACAGAGCCTGGAAGCTGTACCCTCAGAGCAG AGGCGGAGAGAGAAAGGGCCTCTTGGTGGGTGAGCAGGAGCAGAGGCTCA GAGGTGGGGGTGTCAGCCCCCTTCAACAGGCCAACACAGTGAAGCAGC TGACCCCTCCACCTTGGAGACCCAGACTCCTGTCTCCACGCCACCTTG GTTTTTAAGGTAATTTTATTTATATCAGAGTATGGTTGACTTACAATG TTGTGTTGGTTTCAGGTGTACAGCAGAGTGATTCACTTCTACATAGACTC ATATCTATTCTTTCTCAGATTCTTTTCCCATATAGGTTATTACAGAATAT TGAGTAGATCCCTGCTGATTACCCATTTTATAATTGTATATGTTAATCC CAAACCTCCTAATTTATCCCTCCCCAGACTATGATTCTTTATATCTCTATC TGTTTCCTAATCTGTCTCCTCTAAGTCACCCTAGGAGAGCAGAGGGGTCA CGTCTGTCTGTCTGGCCCAGCCACCTCTCTCCACCCAGGAATCCCTTG CATTTGGTGCCAAGGGCCCCGGCCCCGCCCTAAAGAGAAAGGAGAACGGGA TGTGGACAGGACACCGGGCAGAGAGGGACAAGCAGAGGATGCCAGGGTAG GGAGGTCTCCAGGGTGGATGGTGGTCTGTCCGCAGGCAGGATGAGGCAGG AAGGGTGTGGATGTACTCGGTGAGGCTGGCGCATGGCCTGGAGTGTCTG AGCCCTGGGAGGCCTCAGCCCTGGATCAGATCTGTGATTCCAAAGGGCCA CTGCATCCAGAGACCGTTGAGTGGCCCATTTGCTCTGAACCATTTATAGAA CACAGGACAAGCGGTACCTGACTAAGCTGCTCACAGATTCCATGAGGCTG ATGCCAGGGTTGTACCCCATCTCACAGGCAGGGAAACTGATGCATATAC TGCAGAGCCAGGCAGAGGCCCTCCCAGTGCCCCCTCCCAGCCTGTGGCCC

	CCCTCCAGTGGCTGGACACTGAGGCCACACTGGGGCACCCCTGTGGAGATC t
--	--

Seq ID No.35	AGATCTGGCCAGGCCAGAGAAGCCCATGTGGTGACCTCCCTCCATCACTC CACGCCCTGACCTGCCAGGGAGCAGAAAGTAGGCCAGGGTGGACCCGGT GGCCACCTGCCACCCCATGGCTGGGAGAAGGGAGGGCCTGGGCAAAGGGC CTGGGAAGCCTGTGGTGGGACCCAGACCCAGGGTGGACAGGGAGGGTC CCACACCCACAGCCATTTGCTTCCCTCTGTGGGTTCAGTGTCCTCATCTC ATCTGTGGGGAGGGGGCTGATAATGAATCTCCCCATTGGGGTGGGCTTG GGGATTAAAGGGCCAGTGTCTGTGATATGCCTGGACCATAGTGACCCTCA CCCTCCCCAGCCATTGCTGTCACCTTCCGGGCTCTTGCCCAGGCCTGCCT GACATGCTGTGTGACCCTGGGCAAGATGATCCCCCTTTCTGGGCCCCAGC CTTCTCTCTGCTCCGGAAGTGCTTCCTGGGGAAACCTGTGGGCTGGATC CTATAGGAAACCTGTCCAATTCCTGGATGCACAGAGGGGCAGGGAGGGCC TGGGCCTGGAGGGGCAGGGAGGCTCGAGGTGGGAGCAGGGTAGGGGCCAG TCCAGGGCAAGGAGGTGGGTGGGTAGGGTG
--------------	---

Seq ID No. 36 :	GATCTGTGTTCCATCTCAGAGCTATCTTAGCAGAGAGGTGCAGGGGCCTC CAGGGCCACCAAAGTCCAGGCTCAGCCAGAGGCAATGGGGTATCGATGAG CTACAGGACACAGGCGTCAGCCCAGTGTGAGGGAGAATCACCTTGTTTGT TTTCTGAGTTCCTCTTAAATAGAGTTAATTGGTCTTGCCCTTACGGTTT ACAATAACAACCTGCACCCTGTAAACAACGTGAAGAGTACAGAACAACAAA TGGGGGAAAACATATTTACCTGAAAGAGCCACCGCTCATATTTTGATGG ATTTCTTCTAGTTTAATCCTGTTTTAATTGTAAACTGTAAAACAAACA TAAATAAAGAAAATGCATCTGTAAAGTTTAAAGTCATATCTATGGTGAT GGTTGCAAAACACTGTGAATGTTCACTTTGAAATCGTGAACCTCTACGTGA TATGCATGTCCCGTTAATTAACCTCACAGGCTCAGAATGTGGTTCATTAT TTCTTTAATTTTCTTTAATTTTATGTCCTCTGTGTGTGCCCTTAAACCA ACTACTTTTCAGCTCTGCCTGTTTTTGACCTTCACATAGATGACATTTGT GAGTGTTTTCTTTCTCAACACTGGGTCTGATACCCACCCACGCTGTCTGC TGTCACTGCGGACGTGGAGGGCCACCACCCAGCTATGGCCCCAGCCAGGC CAACACTGGATGAATCTGCCCCCAGAGCAGGGCCACCAACACTGGAGGTG CAGAGAGGGTTTCTTCAGGGCCATCATTATCCAAGGCATTGTTTCTACTG TAAGCTTTCAAATGCTTCCCCTGATTATTAAGAAATAATAAGATGGG GGGAAAGTACAAGAAGGGAAGTTTCCAGCCCAGCCTGAAGATCGTGCTGG TTGTATCTGGAGCCTGTCTTCTGACAGGCCTCTATTCCCAGAGTTA
-----------------	---

Seq ID No. 37 :	GGATCCTAGGGAAGGGAGGGCGGGGGCCTGGACAAAGGGGGCCTAAAGGA CATTCTCACCTATCCCACCTGGACCcctgctgtgctgagggagggagca gagagggggtctgaggcctttccagCTCCTCTGAGTCCCTCCTCCGAG CACCTGGACGGAAGCCCCCTCCTCAGGGAGTCCTCAGACCCCTCCCCTCCA GCCAGGTTGGCCTGTGTGGAGTCCCCAGTAAGAATAGAATGCTCAGGGCT TCGAGCTGAGCCCTGGCTACTTGGGGGGGTGCTGGGGATTGGGGGTGCTG GGCGGGGAGCTGGGGTGTCACTAGATGCCAGTAGGCTGTGGGCTCGGGTC TGGGGGGTCTGCACATGTGCAGCTGTGGGAAGGCCCTATTGGTGGTACCC TCAGACACATATGGCCCCCTCAATTTCTGAGACCAGAGACCCAGTCTGGC CTTCCCAGAACAGCTGCCCCCTGGTGGGGGAGATGTAGGGGGGCCTTCAGC CCAGGACCCCAACGGCAGGGCCTGAGGCCCCCATCCCCTTGTCCTGGGC CCAGAGCCTCAGCTATCAGGCCTATCAGAGATCCTGGCTGCCAGCTCAG GTTCCCCAGGAGCCAGAGGGAGGCCAGGGGTACTAGGAAATCCGGAAAG
-----------------	--

	GGTCTTTGAGGCTGGGCCCCACCCTCTCAGCTTTCACAGGAGAAACAGAG GCCCACAGGGGGCAAAGGACTTGCCAGACTCACAATGAGCCCAGCAGCTG GACTCAAGGCCAGTGTTTCGGCCCCACAACAGCACTCACGTGCCCTTGAT CGTGAGGGGGCCCCCTCTCAGCCAGGCATTTCAGACCTGTGACCTGCATCTA AGATTCAGCATCAGCCATTCTGAGCTGAAGAGCCCTCAGGGTCTGCAGTC AAGGCCACAGGGCCAGACCTCCAACGGCCAGACATCCCAGCCAGATTCTT TTCTGGTCAATGGGCCCCAGTCTGGCTTGGCTCCTGCAGGCCAGTGCCG CCTTCTTCCCCTGGGCCTGTGGAGTCCAGCCTTTCAGTTTCCCACCCACA TCCTCAGCCACAATCCAGGCTCAGAGGCAATGTCCGTGGGCAGCCCCTGT GTGACCCCTCTGTGGGTGATCCTCAGTCCTACCCTTAGCAGACAGCGCAT GAGGGGCCCTCTTGAACCTGAGGGATACTCCATGTTCGGAGGGGAGAAAGCT GGCCTTCCCCACCCCCACTTCCAGGCCTTGGGGAGCAGAGAAAGACCCCA GACCTGGGTCCCTTCTAACAGGCCAGGCCCCAGCCAGCTCTCCACCAGC CCCAGGGGCCCTCGGGTCCACGCCTGGGGACTGGAGGGTGGGCCTGTCAGG CGCTGACCCAGAGGCAGGACAGCCAAGTTCAGGATCCCAGCCAGGTGGTC CCCGTGCACCATGCAGGGGTGTCACCCACACAGGGGTGTTGCCACCCTCA CCTGACTGTCCTCATGGGCCACATGGAGGTATCCTGGGTTCATTACTGGT CAACATAACCCGTGTCCCTGCAGTGCCCCCTCTGGGcgacgcgtgcacgcg cacacgcacacactcatacaGAGGCTCCAGCCAACAGTGCCCTCTAGTAG GCACTGCTGTCACTTCTCTAAAAGGTCGCAATCATACTTGTAAGACCCA AGATTGTTTCAGAAATCCCAGATGGAGAAGTCTGGAAAGATCTTTTCTCC TTTCACGGGCTGGGGAAATGTGACCTGGCCAAGGTACACAGCAAGTGGT GGAACCCTGGCCCCCTGATTCCAGCTCATTCCAGTTCCCAAGGCCCTGCCA GAGCCCAGAGGCTGGGCCCTCTGGGGCAGAGGAGCTGGGGTCCTCCCCC TACACAGAGCACACAGCCCCGCAAGAGAGAAGAGACACCTTGGGGAGAGG AATCTCCAGACCAGAGATCCCAGTATGGGTCTCCTCTATGCTGACGGGAT GGGATGTCAAGAGGGGAGGGGGCTGGGCTTTAGGGAAACACACAAAATC GCTGAGAACACTGACAGGTGCGACACACCCACCCCTAATGCTAACCTGTG GCCCATTACTIONAgatct
Seq ID No.38	GATCTTCTCCTAAGACCAAGGAAAACCTGGTCATACCAGGTCCACTTGTC CCTGTGGCCATTGTCCCTCCTTCCCCAGAAGAAACAAGCACTTTCCACTC CACAAGTAGCTCCTGATCAGCTTGGAAGCCCGGTGCTGCTCTGGGCCCTG GGGACACGGCAGGGGCATCAGAGACCAAAATCCTGGAACAAAGTTCCAGTG GGTGAGGCAGGCCGGACAAGCAACACGTTATACCATAATATGAGGCAAAA TATAATGTGAGTTCTTTATGAAAGGAAGGGGTGTCAGGTGCAACTGTTGG CTTAGGTGGATGGTCACCCCTGAATGGAGGAGGGGGTTCCAGGGCATGT GCCTGGGGAGAAGGGCTCCTGGCAGGAGGGACAGCAAGTGCAAGGGCCCT GTGATCAAATGTGCCTGGCAAGTTGCAGGAACAGCTAGAAGGCCAGCAAG GTTGGAACCAAGGAAGGGGTGAGGGGAGGGGCAGGGCCCTCAGGGCCTTG CCCAGCAGCCTGAGCATCTGGAGATTTGTCCAAAGTTTCAAATGTACCTG GGCAACCTCATGCCATATAACCATTCCTAACTTCTGCACTTAACATCTCT AGGACTGGGACCCAGCCAGTCAAGCGGGGGGACCCAGAGAGCTCCGGTGT GAACACCGAGGTGCTGGTGGGTCTGCGTGTGTGGACATAGGGCAGTCCCG GTCCTTCCTTCACTAACACGGCCCCGGGAAGCCCTGTGCCTCCCTGGTGCG CGGGTCGGCGCTTCCGGAGGGTACAGGCCACCTGGAGCCCGGGCACAGT GCATGCAAGTCGGGTTCACGGCAACCTGAGCTGGCTCTGCAGGGCAGTGG GACTCACAGCCAGGGGTACAGGGCAGACCGGTCTGCCTCTGCGCCCCCTC CCTGGCCTGTGGCCCCCTGGACGTGATCCCCAACAGTTAGCATGCCCCGCC GGTGCTGAGAACCTGGACGAGGTCCGCAGGCGTCACTGGGCGGTCACTGA GCCCCCCCCAGGCCCCCTCTGCCCCCTTCTGGGGTGACCGTGGACTCCTG GATGACCCTGGACCCTAGACTTCCCAGGGTGTCTCGCGGAGGTTCTCTCAG CCAGGATCTCTGCGTCTCCTCCTTCCATAGAGGGGACGGCGCCCCCTTGT GGCCAAGGAGGGGACGGTGGGTCCCGGAGCTGGGGCGGAGAACACAGGGA

GCCCCTCCCAGACCCCGCTCTGGGCAGAACCTGGGAAGGGATGTGGCCAT
CGGGGGATCCCTCCAGGCCATCTCCTCAGATGGGGGCTGGTCGACTAGCT
TCTGAGTCCTCCAAGGAACCGGGTCCTTCTAGTCATGACTCTGCCCAGAT
GAAGAAGGAGAGCACTTCTCTCCATCAGGAGGATCTGAGCTTCTCTTAAT
TAGAATCAGCTCCTTGGCTTCTACCCCTTAAAAAAGGTACAGAACTTT
GCACCTTGATCCAGTATCAGGGGAATTTATCAATCAATGTGGGAGAAATT
GGCATCTTTACCACACTGAATCTTTCAATCCATGAATATCCTCTCTCTCT
TCCATGCATAGGTTTTAATAATTCTCAATGGAGTTTAATGTAAGTTTTCC
TCATAGACAATTGCCTTTGGACATCTCTTTAGACTCATCTCTAGTAACT
GATATTCTTAATGCAATTATAAAATGTATCCTGCTTAATGTTATTTTCTA
TTCATTTGCTGTTATATAGAGATACAATGAGTTTCCACATTTGAAACTGG
ATCTGGTAAATTGGCTACCCTTTTTTTATAGATTCTATTAATTTTTTATAC
ATTCTGTGGGACTTGCTACATACTTAATCATGTACCTGTGAAGAATGAC
AATTTGGTTGCTACCCTCCCAATTCTTATATGTCTCATTTCTTTCCCTCT
GCTGGTACTCTGGCAGCAGCAGGGAAGATAATGGGCCTCCTTATCTTGTC
ACAAAAGGATGTTTTTAAAGATTTTCGTTATAAAACATAACGCTTTCTGGT
TTTCTTTAAAGATTCTCTCACCAGCTTAAGAAAATTTTCTTATACTCTGT
ATGATAAATGGGTTTTTGACAATCATTTGTTGCATTTTACCTAGTGTTTT
CTCTGCATCTTTATATGCTTTTTCTCCTTTAATCCTGAAAATTGTTTCGA
TTTTTCTAACATTGAACCAATCTTACATTCCCTGGAATGGATGGACCAGAC
TAGTCCACATGTTTATTCTGCCCAATGGCTAGATTTTGTGTTCaatattt
tggtcagaatgttgcatctatattcttGAGTGAGACAGAGCTGCCCTTG
TTAGGTTTTACAACCGAGGTTGTGTTAGCTTCATAAAATGAGACGTTTAT
TCTCTAAAAGAATTGTTTCGCTTCTCTGGATGAATTTGTGTAAGGTTAGA
ATTGCTTACCAGTGAagatctCGGGgCCAGTTCTTCTTTAGGGGAAGATT
TTCAACAATTAAGCTCAATGCCTTTAGAAGAACTGAGAGTTTCTATTATT
TCTTGAGTTAAATATATGTATTTAATTAGACTTTCTAGGAATAGTCTCAT
TTCATCTCAAATAATTGACATATGCTATTAAAGCAGATTCTCATGAACCA
TTGTAGGTATTCCAGGTCTAGAAAAATGTTCCCTTTGCATCCCTAATGT
GTTTAATTTTACCTTCTTTCTTTTGTCTTGAGAAATTCACCAAATCAT
TTTCAATTTCAGTCATATCCCAAAGCAACCAACTCTCTACCTTCTTGTTT
TATCATCCCTGCTGGATTTTTGTTATCTACTTCTTCAGTATTTGTTCTTC
CCTTTCTTCTATTCTCATTTCCATTTTCCCTTGTTTTCTAACTTTCTGA
GATATATGCTTAGTTCCCTTCATTTGAAGCCTTTTTATTTTCTTTTTTTT
TTTTGGTCTTTTTGTCTTTTGTGTTGTTGTTGTGCTATTCTTGGGCCG
CTCCCGCGGCATATGGAGGTTCCCAGGCTAGGAGTCGAATCGGAGCTGTA
GCCACCGGCCTACGCCAGAGCCACAGCAATGCGGGATCCGAGCCGCGTCT
GCAACCTACACCACAGCTCATGGCAACGCCGATCGTTAACCCACTGAGC
AAGGGCAGGAACCGAACCCGCAACCTCATGGTTCCTAGTCGGATTTCGTAA
CCACTGTGCCACAACAGGAACTCCGCCTTTTTATTTTCTATAAAAATTTCT
TATGTACATTTTAAGGTTATAGGTTTCTTCTATGTACCCCATTTGGCTGT
ATCCTCAGGGTTCTGTGGAGTGATTTCAATTATTGTTCAAGTTCAATATGT
CTTCTGATTTTCCAATTTGAATACCTCTCTAAATCAGTAGGTGAATATTT
CTTTTTCTTTTTCTTTTTCTTTTTCTTTTTTTTTTTTTCTTTTCAGCCAGGT
CCATGGCATGCAGAAATTCCCAGGCCAGGAATCAAACCTCTCACCATGGCA
GTGACAATGTTCGGATCCTTTACCCACTAGGCCACCAGGGAACCTCTGGGAG
CATATGTTTTTATTTCCCGACATCTGAGGATGCCTAGTATGTCTTCATTA
TTGATTTCTAGTTTGCCACTGATTTCTAGTATTTTGCTCATAGAGTGAT
GCTCAATGGTTTTTGGTCATTTGAAATGTATTTAGTCCTGCTTTATGACCC
AGTATGTGGTCAGTTTTGTCAATGTTCTTTCTGCTTGAAGAGAACCTA
CATGCTGTAACCTCTGGGTGCATGTTCTGTATATAAGTCTATAGGCTGAGC
CGGGGGAGCCTTCTAATCTGCCGTTATCTTCTTCGAGTTATTCTAGGTAC
TATTTCTTAGCCATAAACCTTTAAATTCTGATATCAATATAATGACCCCA
GCCCCGCTTAGGGTCGGCACTTCATGTTATCTTTTTCCATCCATTTAATCC
CTCCCCACTGTTTTGGCCACACCCGTTGGGATATGGGAGTTCTGGGCCAA
GGATCaGATCTGAGCCGCAGCTGCCACCTATGCCACAGCAgcagcaatga

	tggatctttaacccactgcaccacactggggattgaacccaagcctcagc agcaacccaagctactgcagagacaacaccagatccttaacctgctgtgc catagcgggaaTTTCCATCCATTTACTTTCAAGCCAGCTGAATAACCTAG CCCACCATGCCTGGACATGGGTGCTCTGCTTCAAATGATTTTGTTCAGTC AGCATCCATCTCTGAAATGTGTGCCAAGCATTATATGCATGCAAGAGTC ATGTTGGCACTTCTATCATTTCACAGTTCAGTAGCCTTTGTATCATGA CATTTCTTGGCCTTTTCTCTACAATATTTGAGGCTGAGCAGACTGGCCGT GCCCCTGTCCATGCTTCCAGAGCCTGTGTGCAGACTTCTGCTCTAGACAG AGACAGCTAACCATCCTGCAGTGCCAGAAAACCCAACCTCAAAGACCCTC AAGTAAGGAAGGATTTATTGGCTCACGTAATCTGGAATCCAGGCATGGGG TATTCAGGGCCACCTGAACCAGAGGCCCTGGCCCTGTTCTCTAAGCTTCT TCCTGCCCTGCCCTCGTTCTGGAAGTGACCCTGAAGGACAGCAATGAAGG GCAGCTCCCCCAGGGACAGATGACTGAGAGGTCCATTTCAAGTCCAACCTT GGCCTAGATTGAGAGGCAGCAAGAAATATGGACCTACAGTGAGTCACAGG ATTTACCAGTGGTTTGGCTGGGTTGTCAGTGTTACAGGCTAAACATTTGG GTCCTCCAAAATTAACATGTTGCCACTCTAACCACCAAAATCatggtat ttgggggtggggcccttgagagtaattaggttagaaAGAATGAAGAGGG GGCCCTTGTGATGGGACTAGTGCCTTTATAGAGAGAGAAGAGAGAGGG
--	---

Seq ID No.39	CACCTCATCCCCAACCACTGGATGGTGGCAAGTGGCAGGCTGAGAGGCT GCATATGAGCTCATCAAGAGGGTCCCCACCCACAGAGGCTGACCCAGCT GCCACTGCCACCTAGTGGCTGATCGGCCAAGAGCAGGAGCCCCAGGGGCA GCTCCATTCCCTGGGGCGGCCAGGGAACCACTGGTGGTAGGACAATTCC ATTGCACCTCATCCATCAGGAAAAGGTTTGCCTTCCCTGGCAGTAATGCA TCTTCCCATAACATGGTCCCTGGCCTCTTGAATGGCTTGGCCACCGTCA TGGCCTCACCCACAAAGCCTTGTGTCTCAGCAAGGAACTTATTCCACAGC AAAGGACTTGCAGCCTGGAATGAACTGGTCTGACTACATAACCCATTGCC CAGAAGTAGGTGGTCTATTGCAAAGTGGAGTGGCTTACCCAAGACTCAGT TGTGCCCAAGTTGAGAGATAGCATCCTAAATATGGGCTTATGTCTCACT GGCTGAGGTTTATTCTTTGAATCAAAGACAATTATATGGTGTGGTCCCCC CAGAGATAGAATACATGAGTCTGGGAATCAAGGGATAGAAGTAAGAAGAG ATTTTGTACCATTAATCCCAATAACTCGCCCAAAGAATATTTGCTTTCT GTCCTGGCAGCTCTGCTGCTTTGGCAATAACTTCTAGAAATATAATGTCT CCACCAGGGGACTCCACAACGGTTCCATTGATTTGAAGCCAATGGGCAGA GGAGGGGCTGCCTTACTGGTCGGACTGGTCAGCCCTGATTACTAAGGAGA AATCAGGCAACTTCAACAAAATAAGGCAGGGGGGACTTTGTCTAGAACC CAAAGCACTAAGCATCTTAGTACTTTTTAGTTCTCAGAGCCTCCAAGAAC AAAGATTTAGCCCCTCAGCACCACCAGGTAAAGAACAGGTAAATCCAGCT GAGGACAAGAGAAATATTGAATGGATAGAGGAAGAAAGAAATTATAGATA TCAACTATGGCCTCATGACTAGAGTCTCCAGATTAAGCGGAATAAAAATA CAGATGATTaGATCTGAACATCAGGCCAAACAACGAACAACAGTTTAAGT GCGACCTAGGCAATATTTGGGACATACTTATACTAAAATTTTTTCGCTAT TTGAGCATCCTGTATTTTATCTGGCAACTTTATTCATCCCTAGCGAAAAA GGAAGTGTGGTAACCTTAGTGTATTTTACTTTGCTCATTATTGTGTATAT ACCTACTTGTATTTATCAATCATATTTACTCTGTTCTCAGTATTACTTTA TATAGCAGTTGGTGGTGTGTTAGCAACATATTCAGTGGAAGTGTGACT GAATTTGAGGAGAAATTAACAGAGTTGGCTGTGGCTACAATAACCCCTTCG GGACATGTGTCCCCTCATTTTGGGGAGATGGTTAgatctCTGGGTAAATG TTAGGGCATCTGAGCCAGAAACCAAGATTTTGCCAGCTGGTGCAATGTCA GATTTTACCAGCAGAGGGTGCCAGAGGAATGCGGCAAAACCCGAGTGCCA GAAAGCACCTCCCTGTTTTCCAGCTTTTCTTCCTTTTTATTTATTTTATT TACGGCCCAGGAGTCCGTAATAGCGCTGAGGATGGCCCAGGCTCTTCTCA GCAGCCCTGACTGACTAGTTCAGCAATGCGCTCAGGCCCCATCTGGCCAC CGGGCAGCCTCTTCTGTGGTAGCTCCAGCCTCAGCCAGTGCAAAAGGCTA
--------------	---

	CCCTACACTGGCGCCACTTCTACAATCAGCACTGGCCACACCCTCCACGC CATCCGGCACGGAGCCAGGTGATCTGCCGGCCAGATTGCAGTTCGTGCTG CCTGAGTCCAGGTGATTACACTGGCTGCATCTTTTCTTTCTGGACCAATC attccatttttt

Bovine Lambda Light Chain

In a further embodiment, nucleic acid sequences are provided that encode bovine lambda light chain locus, which can include at least one joining region-constant region pair and/or at least one variable region, for example, as represented by Seq ID No. 31. In Seq ID No 31, bovine lambda C can be found at residues 993-1333, a J to C pair can be found at the complement of residues 33848-35628 where C is the complement of 33848-34328 and J is the complement of 35599-35628, V regions can be found at (or in the complement of) residues 10676-10728, 11092-11446, 15088-15381, 25239-25528, 29784-30228, and 51718-52357. Seq ID No. 31 can be found in Genbank ACCESSION No. AC117274. Further provided are vectors and/or targetting constructs that contain all or part of Seq ID No. 31, for example at least 100, 250, 500, 1000, 2000, 5000, 10000, 20000, 500000, 75000 or 100000 contiguous nucleotides of Seq ID No. 31, as well as cells and animals that contain a disrupted bovine lambda gene.

Seq ID No 31	1 tgggttctat gccaccagc ttgtctctg atggtcactt gagggcccca tctcatggca 61 aagagggaac tggattgcag atgaggacc gtgggcagac atcagaggga cacagaaccc 121 tcaaggctgg ggaccagagt cagaggcca ggaaggctg gggaccttg gtctagggat 181 ccgggtcagg gactcggcaa aggtggaggg ctcccaagg cctccatggg gcggacctgc 241 agatcctggg ccggccaggg acccaggga agtgaaggg gaagacgggg gaggagaagg 301 tgctgaactc agaactggg aaagagatag gaggtcagga tgcaggggac acggactcct 361 gactctgcag gacacactcc tcagaagcag gactccctga agaagcagag agacaggtag 421 cagggcagga aacctccaga cccaagaaga ctacagagg aacctgagct cagatctgcg 481 gatgggggga ccgaggacag gcagacaggc tccccctga ccagcacaga ggctccaagg 541 gacacagact tggagacaa cggacgcctt cgggcaaagg ctgaacaca catgtcagct 601 caaatatac ctggactgac tcacaggagg ccaggaggc cacatcatcc actcagggga 661 cagactgcca gcccaggca gacccatca accgtcagac gggcaggcaa ggagagttag 721 ggtcagatgt ctgtgtggga aaccaagaac caggagctt caggacagcg ctggcagggg 781 tccaggctca ggcttccca ggaagatggg gaggtgcctg agaaaacccc acccacttc 841 cctggcacag gccctctgga tcacagtgtt gctggactc ggggtctgc tgggtctca 901 aaggatcctg tgccccctg tgacacagac tcaggggctc ccatgacggg caccagacct 961 ctgattgtgg tcttctccc ctgcccact tgcaggta gcccaagtc acacctcg 1021 tcacctgtt cccgccctc aaggaggagc tcagaccaa caaggccacc ctggtgtgtc 1081 tcacagcga ctctacccg ggtagcgtga ccgtggtta gaaggcagac ggcagacca 1141 tcaccgcaa cgtggagacc acccgggct ccaaacagag caacagcaag tacgggcca 1201 gcagctacct gagcctgatg ggcagcact ggaatcgaa aggcagttac agctgcagg 1261 tcacgcagc ggggagcacc gtgacgaaga cagtgaagcc tcagagtgtt cttagggccc 1321 tgggccccca ccccggaag tctaccctc ccacctggt tccccctagc ccttctct
-----------------	---

1381 gcacacaatc agctcttaat aaaatgtcct cattgtcatt cagaaatgaa tgctctctgc
 1441 tcatttttgt tgatacattt ggtgccctga gctcagttat ctcaaagga aacaaatcct
 1501 cttagccttt gggaatcagg agagaggggtg gaagcttggg ggtttgggga gggatgattt
 1561 cactgtcatc cagaatcccc cagagaacat tctggaacag gggatggggc cactgcagga
 1621 gtggaagtct gtccaccctc cccatcagcc gccatgcttc ctctctgtg tggaccgtgt
 1681 ccagctctga tggtaacggc aacacactct ggttgccacg ggcccagggc agtatctcgg
 1741 ctccctccac tgggtgctca gcaatcacat ctggaagctg ctctgtctca agcggccctc
 1801 tgtccactta gatgatgacc cccctgaagt catgcgtgtt ttggtgaaa cccaccctg
 1861 gtgattccca gtctcacag ccaagactcc cccgactcg accttccaa gggcactacc
 1921 ctctgcccct ccccagggc tcccctcac agtcttcagg ggaccggcaa gcccacaacc
 1981 ctggctactc atctcacagt tccccaggt cggcctctc ccactgcat ggcaggaggg
 2041 tccagctga ctctgaggtc tctgaccagc ccagctctgc tctgagacc cttaaaactc
 2101 agcccaccac ggagcccagc accatctcag gtccaagtgg ccgttttgtt tgatgggttc
 2161 cgtgagctca agcccagaat caggttaggg aggtcgtggc gtggtcatct ctgacctgg
 2221 gtggtttctt aggagctcag aatgggagct gatacacgga taggctgtgc taggcactcc
 2281 cacgggacca cacgtgagca ccgttagaca cacacacaca cacacacaca cacacacaca
 2341 cacacacgag tcaactaaaa cacggccatg ttggttgac gcactctag gaccagaggc
 2401 gcttcagaa tccgcatgg cctcactctg cggagaccac agtccatcc cctccgggt
 2461 gaaaaccgtc tctcaccct cccaccgggg tgaccccaa agctgtctac gaggagcccc
 2521 cacctctcc aggagaagt ccttgggacc cgggtgtgaca cccagccgtc ctctctgcc
 2581 ctccccgcc tggagatggc cggcgccca ttcccaggg gtgaactcac aggacgggag
 2641 gggctcctcc cctcaccgc cgggagggc aaccagcccc ttgaccagg aggggggagg
 2701 acctggggct ccgagtcag ctgcaggcgg gccccgggg gtggcggggc tggcggcagg
 2761 gttatgctg gaggctgtg cactgtcgt gtttgcctg ttgagggacc cagctggcca
 2821 tccggggtga gtctccctt tccagcttc cggagtcagg agtgacaaat gggtagattc
 2881 ttgtgtttt ctaccatc tggggtgag gtctccgta ccctaggcct gtaacctcc
 2941 ccttttagc ctgtccctc tgggtctct cactttctt tgaggagacg ttactgtc
 3001 acccagcaaa gccagagaa tatccagatg gggcaggcaa tatgggacgg caagctagtc
 3061 caccctctta ccttgggtc cccgcccct ccgataatg tctgagctgc ctccctggat
 3121 gcttcactt ctgagactgt gaggcaagaa acccctccc caaaaggag gagaccgac
 3181 cccagtcag atgaactgc tctgagggga ccttgggagt aagtgggtc tggcggggac
 3241 cgtgatcatt gcagactgat gcccaggca ggtgagagg tcatggccgc cgacaccagc
 3301 agctgcagg agcacaggc gggggcaagt catgcagaca ggacaggacg tctgacctg
 3361 aagagtcaga gtgacacgc gggggggggc cggagctcc cgagattagg gcttgggtc
 3421 taacgggatc caggagggc caggggcca cccagccct ctccctgac ccaatcaact
 3481 tgcaataaaa cgtctctat tctttacaa aaacctgct ctctgtcat gttttctt
 3541 gccccgcat taatcgtcaa cctctccagg attctggaac tgggtgggg nnnnnnnnnn
 3601 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 3661 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn agcttatgtg gtgggcaggg gggtagtaag
 3721 atcaaaagt cttaaattaa taaagccggc atgatatac agtttgata aaaaatagat
 3781 ggaaaagtaa gaaaggacag gagggggggtg aggcggaaga aggggggaag aaggaaaaa
 3841 aaataagaga gaggaacaaa gaaaggagg gggccgggtg atgggggtg gatagaatat
 3901 aataattga gtaaagata gcgggtggt gttaattccg ggggggaata gagaaaaaa
 3961 aaaaaaatg tgcgggtgg cggtaatat ggagattta taaatattat gtgtggaata
 4021 atgagcgggg gtggacgggc aaggcgagag taaaagggg cgagagaaaa aaattaggat
 4081 ggaatatat gggtaattt taaatagagg gtgatatat ttagattgag caagatataa
 4141 atatagatg tgggggaaa gagacaagg tgagcgcaa aacgccctc cgtatcatt
 4201 gcttcttc cttaaccac tcttcaaac ttttttca gaacctgaa gcggtcaggc
 4261 ccggggctgg ggttgggata cccggggagg ggtgcgcct cctctttgc agagggggtc
 4321 gaggagtgg agctgaggca ggagactggc aggttgaga gatggctgt gacttctgc
 4381 ctgttgaac tcacagtcac agtccagac ccactgaatt gggctaaata ccatatttt
 4441 ctggggagag agttagagc gagcgactga ggcgagctca tctatctac agggccgcca
 4501 gctgcaggga ctttgtgtg tctgtctcg ttgctcagt gtgtccgact ctttatgact
 4561 tcatggactg taacctgcca ggctctctg tccgtggaat tctcaggca agaatactgg
 4621 agtgggtagc cattctatc tccgggggat ctctgacc caagaatcaa acctgagtct
 4681 cccgattgc aggcagctt ttttgtct gagccaccag ggaagccct taagtggagg

4741 atctaaatag agtgttagg agtataagag aaaggaagga cgtctataca agatccttcg
 4801 gttcctgtaa ctacgactcg agttaacaag cctgtgtga gtgagttgcc agtaattatt
 4861 gctaacctgt ttcttctact cactgagcca ggtatcctgt gagacggcat acttacctcc
 4921 tctctgcat tctcgggat ggagctgtgc ggtggcctct aggactacca catcgaccag
 4981 gtcagaccca gggacagagg attgctgaga tgcactgaga agttgtcag cctaggtctt
 5041 caccacaca gactgtgtg tegtctacca cgtaattctt cctgtccaaa gaactggta
 5101 aacgtcctg aagegtatc tggctgtctt caaaaagtgc ctcttctt tataagtcc
 5161 gccaatcctg gacttgtcc caggccagtc tactttatt gtgggaaagg ttttttgg
 5221 ctttttgg ttaaactctg cagaaatgc ttacacttt ggtgtgcaat ggctcactct
 5281 tacggtcta gctgtatca aagggggtgc tttcttgt ttttaaagct tttgaacgt
 5341 ggaccattt taaagtctt attaaacgc taacatcgt tctggttat tttcgttg
 5401 tctggcatg aggcctacgg gtcttagtc cctaccagg gtccaacca catccctgc
 5461 actggacggc aaggtctta ctttgaacc accagagagc ttctgaaagg ggctgtttt
 5521 ctccaatcct cttgtccc tgcctgtgg tagggattca gcaccctgc aatagccctg
 5581 tctgttcta ggggctcagt agccttctg cctgggtgtg gagctgggtg tgaagagag
 5641 ctcatgat ttggacaga cctacgact agaggtaaga ctccatcta gcgctgaat
 5701 gacctctt caacaaccac cccaccacc ctggaccact gatcaggaga gatattctc
 5761 tctctatca tcaacgtgt cagtccaaa ctgcaccg gcctgtcata gatgtagcag
 5821 gtaagcaata aatattgtt gaattgaa tgaattgaa taacataagt gaaaaagaa
 5881 aacttaaaa acatgtgtt ttataattac acagtaaaata tataatcatt gtagaaaaa
 5941 atcgaaagag tggcgggggc caagtgaata ccaccatccc tggatgtcc accgcccgg
 6001 gtagcccgag gtaagaggtg cggacagga tggccctgta gacacagaga cacacgtca
 6061 tatgtgggt cttgtctgt gacctctgg ggtatgtt atttcacga tgccatcaa
 6121 acctctacc acaccattt tagagggtc ttatcgtaa atcattcac tgccttgtt
 6181 tctgatttg aaagtgtac attctcgag aatgagaag gaacaggcgc gcataaggaa
 6241 gaaagtaac acgtggcctt gttccaggg ggcactcagc gtgtgtgt gcacgtggc
 6301 agtctttct ctgtacagt catggcctt tccaaagggt gggctcagat aagaccgct
 6361 cccatccct gtcctgtcc cgtcccta cgtggaacc caccacggc acgtctcga
 6421 ggcccttgg ggctgtggac gtaggctgt tggacatgc tctgtgtgg gaccagggc
 6481 tgggcagcac gttgtccctg ggtccgggc cagtaggag ctccaagga gcagggtgc
 6541 tgggcaaaag ggcagtgcg cccaggcca tggacaagg gataattc ctgtgaagg
 6601 gctggactgc gtctccctg ggcccttgg agtcatgggc agtggggagg cctctgtca
 6661 cccgttgc caccatggc tcatgtgca gccaggagc cctggggctg ggacgccgag
 6721 gccggagccc ctccctgtg tctgacggg ctgggtgacc ctccgccc ctccctggg
 6781 cctgtgtac cgcgggggc accccggcca gttctgagat tccctgggg tccagccctc
 6841 caggatcca ggaaccagga tggcaaggat gttaggagg cagctaggg gcagatcag
 6901 gccagaccg gggctggga ggggctgggc gcaggcgggt ggggggtct gcacncccc
 6961 acctgnagc tgnnnnnn ttgntnncg tctccctgn tctgtgtg tcccgcgg
 7021 ggggcccc ctggtctgt ttgtcccc tcccgctcc tccccctt ttccgtct
 7081 cctccctct tttatgcc cctgtgtg gtttttgc cgtccctt tgttttt
 7141 gtctttct tttcccc ctctccct gtctcttt tctctgtg gttttctg
 7201 tccctccct cccccccg ctttttcc ctgtgtgt tttgttct cctctctac
 7261 cccctgca gctatttt ttatatac cattcccc tagtattgg ccccgctta
 7321 ctctcccta attttatt tcttcttt aactaaaac accgtgtgt tataagttt
 7381 aacctttt gaccgcca caatgcaac ttacgcag cccccccg cagcctct
 7441 aaatactt gctactgc cctccttg tataataac cgtcacgtg tcaaccatta
 7501 tcacctcc accacctac cacatttcc tcnnnnnn nnnnnnnn nnnnnnnn
 7561 nnnnnnnn nnnnnnnn nnnnnnnn nnnnnnnn nnnnnnnn nnnnnnnn
 7621 nnnnnnnn nntgaaaa agaaaagggt gggcagggtt taatatggg ggttggagt
 7681 ggaatgaaa tgcattggag tgggtgcaac aatggaaag gtctcaggag cgtctctcc
 7741 ccatcaggag ctggaagaa gtggaagca agcaaggaat tctgtgat gccagaggtc
 7801 aggggcagg agctgcaa agtccggct gttgtgact gncgtctcc ggtgcattt
 7861 gtagcagg aggcattaca ctcatgtt gttgtgta ctattctta ctattgta
 7921 gttgcaagg catgtgtac tcttgcaac ccaggactg cagccgcca ggtctctg
 7981 tcatgggat ttgcaggca agaatactg agtggttag catttctc accatggat
 8041 ctcccgagc cagaaatga acccgagtc cctctgtg atggggtct ctgcctaaca

8101 ggcagatatt tgacgtctga gccaacaggg aggacagacg gtaattatac caaccattga
 8161 aagaggaatt acacactaat ctttatcaaa atctttcaaa cagtagagga gaaaggatac
 8221 tctctagttt attccataaa gttggaatta cgttatcaa taaagacatt acaagaaaag
 8281 aaagtgaagc cccaaatgcc ttataaatat acaagaaaaa atcttttaag atattagcca
 8341 acttaatcaa caaaaaatgt atcaaaagtc caagtaacat tcaccccagg aatgcaagt
 8401 tggttcagcc taagacaatc agtcatgagt ataccacgga aacaaattaa agagaaaaga
 8461 cattaatct cacaatgggt gcagaaaaag atttggcaat atcgaacatc tttcatgac
 8521 caaaggaaaa aaaagaaaca aaacaccaga aaattctgtg tagaaagaat atatctcaac
 8581 ccaatgaagg gcatttatga aaaaccaca gcatacatca cactccatga gaaagactga
 8641 aagctttccc cactgccatt gaactctgtc ctggaaattc tagtcacagc gacagaacaa
 8701 gagaaagaaa taacggccgt ctaaactggg aggaagaaat caaagcgtct ctattctctg
 8761 ggcgcataat acaatataga caaatttcta aagtcacaaa aaattcttag agtcataat
 8821 gaatccagaa atgcgtcagg gctcaagatt cagatgcaaa aatcgtctgg gttttagtc
 8881 accaacaac aattccatta acaataatac caaggaatta atttaacta gaagagaaaa
 8941 gacctgttta cagagagtta taaaacattt ggtgatgaaa ttaataaga gtaaatcata
 9001 tagaaacacc gtctgtgtt tggagaccta atgtcataaa cgtggcaaca cagagacgcc
 9061 tcacggggaa cctgagcct cttctccaa acaggcctgc tcatcttcc acaggttaacc
 9121 tgagacccta aagcttgact ctgaggcact ttgaggcct gaagagagca gtagctctc
 9181 ccatgggacc gacagtcaag gccaggga tgaccacctg gacagatgac ttcccggcct
 9241 catcagcagt cgggtcagag tggccaccag ggggcagcag agagtcgctc aacactgcac
 9301 ctggagatga ggcaacctgg gcatcagtg ccatgcagg ggctggatac ccacacctca
 9361 cacctgagga caggggccgg cttctgtgg tctcgcctc tcaggatgca cagactccac
 9421 cctctcgtc tgcattgaca gcctctgtc ttctggagg acaagctcca cttccccat
 9481 ctctcccag ggggctgggg ccaacagtgt tctctctgt cactccagg aacacagagc
 9541 caagagattt attgtctta attagaaaa ctattgtat tctgcattt cccagtaac
 9601 tgaaggcaac ttaaaaaat gtatttctg gacttccctg gtgggccagt ggctagactc
 9661 tgagctcca gtgcatgggg cctgggttca atccctgctc aggaactac atcccacagg
 9721 ctgcaataa gatcctgcat gccaccgat gcaggcaaag aaacaagtgt tcggtatgca
 9781 tgtattcac gtgaggtgt tctataatt acagccagta ttctgtctta cacttagtca
 9841 ttctttgag cacatgatc gtcgatggc cagaccacac acaggaatac tgaggcccag
 9901 caccaccgg ctgccagaa cctcatggc aagggtggac acttacagga cctcagggga
 9961 ccttaagaa cgccccgtgc tcttggcagc ggagcagtgt taagcatggc tctgtcctc
 10021 gggagctgtg tctgggctgc gtgcatcacc tgtgtgtgtg gcctgtgtg ggtcacctc
 10081 caggggccct cgagggtcag aagaacctc ctttaaaagt tctagaggtg gagctagaac
 10141 cagaccaca tgtgaactgc acccaaaaac agtgaaggat gagacactc aaagtctgg
 10201 gtgaaattaa gggccttccc ctgaaccagg atggagcaga ggaaggactt ggcttcagg
 10261 aaacctgac gtctccaccg tgactctggc cggggctatg gcagggccca ggtccttg
 10321 gtgcaaagga ctgagggtc ctggaaaata cagtctccac ctctgagccc tcaagtgaag
 10381 gggcttctc cccaggagt gggcaaggac ccagattggg gtggagctgt cccccagac
 10441 cctgagacca gcaggtgcag gagcagccc gggctgaggg gagtgtgagg gacgttccc
 10501 ccgtctcaa ccgtgtagc cctgggctga gcctctcca ccacggctgc aggcagccc
 10561 caccaccac cccgacctg gctcggactg attgtatcc ccagcagca ggggataaga
 10621 caggcctggg aggagccctg cccagcctgg gtttggcag cagactcagg gcgcctccac
 10681 catggcctgg accccctct cctcggcctc ctggctact gcacaggtga gcccagggt
 10741 ccaccacc cagcccagaa ctggggaca ggcctggccc tgactctgag ctcaagtggga
 10801 tctgcccgtg agggcaggag gctcctgggg ctgctgcagg gtgggcagct ggaggggctg
 10861 aaatccccct ctgtctcac tctaggtca gccctgaggg ctgtcctgc cagggaagg
 10921 ggggtctct tctcagag actccatcca ccaggacat gagccggggg tctgagact
 10981 gacggggagg gtgtccctgg gggccagaga atctttggca ctaattctc atcaggcagg
 11041 gggcttctg tctaggtt ttacgtcca gctacctc cttctctc ctgcaggcgc
 11101 tgtgtctcc tacgagctga ctcaatcacc cccggcatc atgtcccag gacagcggc
 11161 caggatcac gtgtggggc ccagcgttg aggtganaat gttagtggc accagcagaa
 11221 gccaggccag gcctgtgcgc tggctccta tggtagcat aaccgacca cgggggtccc
 11281 tgaccagtc tctggcga actcaggga catggccacc ctgcccata gcggggccc
 11341 ggccaaggat gaggccgact attactgtca gctgtgggac agcagcagta acaatctca
 11401 cagtacaca ggcagacggg aaggagatg caaacccct gcctggccc gcgggcccag

11461 cctcctcgga gcagctgcag gtcccgtga ggcccgggtgc cctctgtgct cagggcctct
 11521 gttcatcttg ctgagcagcg gcaagtgggc attgggtcca agtcctgggg gcatacagc
 11581 acccttgagc cagaggggta ggggttaggg ttagggtag gctgtcctga gtcctaggac
 11641 agccgtgtcc cctgtccatg ctacagcttct ctcaggactg gtgggaagat tccagaacca
 11701 ggcaggaaac cgtcagtcgc ttgtggccgc tgagtcaggc agccattctg gtcagcctac
 11761 cggatcgtcc agcactgaga cccggggcct ccttgagggg caggaggtgg gactgcagcc
 11821 cggcccccac accgtcacc ccaaccctcg gagaaccgcg ctcccagga cgcctgcccc
 11881 ttgcaacct gacatccgaa catttctc agaacttctg caaaatattc acaccgctcc
 11941 ttatgcaca ttctcagaa gctaaaagt atcatggctt gtaaccact ctcttaaat
 12001 attcttctt aacgtccatc ttccctgctc cttagacgcg ttctatcc acatgtctta
 12061 ctgcctttgg tctgtcgtg tattttctt tttttttt ttatigga atatattgc
 12121 gttacaatgt tgaattgaa ttgtttctg ttgtacaaca atgtgaatta gttatacatg
 12181 tctgaggag gggcggtgc gtgggtgcag gagggccgag aggagctact ccacgtcaa
 12241 ggtcaggagg ggcggccgtg aggagatacc cctcgtcaa ggtaagagaa acccaagtaa
 12301 gacggtaggt gttgcgagag ggcacagag ggcagacaca ctgaaacct aatcacagaa
 12361 actagccaat gtgatcacac ggaccacagc ctggtctaac tcagtgaac taagccatgc
 12421 ccatggggcc aaccaagatg ggcgggtcat gtgccatgg ggccaacca gatggcggg
 12481 tcatggtgaa gaggtctgat ggaatgtgt cactggaga agggaaaggc aaaccacttc
 12541 agtattcttg ccttgagagc cccatgaaca gtatgaaaag gcaaatgat aggatactga
 12601 aagaggaact cccaggtca gtaggtgccc aatgtctac tggagatcag tggagaaata
 12661 actccagaaa gaatgaagg atggagcaa agcaaaaaca ataccagtt gtgatgtga
 12721 ctggtgatag aagcaaggc caatgatga aagagcaata ttgatagga acctggaatg
 12781 ttaagtc aa gannnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn
 12841 nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnagaattt
 12901 gagcattact ttactagcgt gtgagacgag tgcaattgt cggtagtgt agcattctt
 12961 ggcattgcct ttcttggga ttggaatga aactgacctg ttccaggcct gtggccactg
 13021 ctgagtctc caaatttgt ggcgtattga gtgcactact ttaacagcat catcttttag
 13081 gattgaaat agctcaactg gaattctatc actttagcta attccattca ttactttgt
 13141 ttgtagtat gtttctaag gccccctgg ctctatctc ctggtatgt ggctctggtg
 13201 agtgatcaca ccgtgtgat tatctgggtc atgaaggct tttgtatag ttctcttag
 13261 gaacagatat tatgatctc atcttgcat ctggtatat ctgagaagc actgactccc
 13321 ttcatggtga cgtcagatcc tcatgactaa caaatggcct ttgtaagat gactgctca
 13381 tggattgag ctcctccgtc accaagacct tatgactgac ctcctccact gcccaggtg
 13441 cctctgaag cgtctgagat gccgctccc aggtgcact cctcatttg ccccaataa
 13501 aacttaact gcagctctc agctgtgcat ctgtgttag ttgacagtac aaatataatg
 13561 gaaaattaa attaatata atctatggg agaatccaa acatcttatg agggagagag
 13621 agggagagaa aggaagaag aagaagcagg agggaggaga gagtagagaa acagggggag
 13681 ggcggcaggg agacagagg gaggacacc aggggaaagg gaggaaggcg agtcagtg
 13741 gagagaggcc agagttcatc agagctgga ctgcagccc aatcccacgg gtgtgtccc
 13801 aagcagggga gagcctgagc caggcgga cagagctgtg tctccagtc tctggccgt
 13861 gacctggagc tgtgtgtga gccccctga cccagcctg gccctgctgg tggcggagg
 13921 cagtatcct ggacacagt tctgagcgt tgtctgaaat cctgtggag gcgacctca
 13981 ggacggacct cgcctggccc cactggatc tgcaggcca ggcccagtg gggcttctg
 14041 cctggaactg agcagctgga gggcgctct cccccagca gtggagcggc cccagggcg
 14101 ctgagagctg ccggggggac acagagcttg tctgagacc agggctcgtc tccaggggt
 14161 cccctaaggt gtctctggc cagggtcaga gccgggatga gcacaggtct gactcagact
 14221 ttgagagctg gtggctgat cctggggac agagggtgtg gtctaacct ggggtcaga
 14281 gggcaggacg ggagcccagc tgaccttg ggactggcct cctctgtgt ctccttggg
 14341 cagtcacagc ttcccggac gtggactct agggagacag ctggggcctg gctgtcagga
 14401 gggggtcga gaggccacac tcagaggagg agacctggc ctgcttgggt tgtgactgag
 14461 ttttgggt cctctaggag actctggccc tgcaggccct gcaaggtcat ctctagtga
 14521 gcaggactcc acaagattga tgaactgaat cctctaggag aggtgtgtt gtgagggggc
 14581 agcattctag aaccaacagc gtgtgcaggt agctggcacc ggtctagt ggcggggca
 14641 gggcactcag ggccgactag ggtctgggg gattcaatg tcccacagc actgggtctt
 14701 ccatcagaat cccagactc acaaggcagt tccgggatt aggtcaggac gtgagggcca
 14761 cagagaggtg gtgatggcct agacaagtc ttacagaga gagctccagg ggccatgata

14821 agatggatgg gtctgtattg tcagtttccc cacatcaaca ccgtgggtccc gccagcccat
 14881 aatgctctgt ggatgcccct gtgcagagcc tacctggagg cccggggaggc ggggcccgcct
 14941 gggggctcag ctccggggta accggggcag gcctgtccct gctgtgtcca cagtctccc
 15001 ggggttgag gagagtgtga gcaggacagg agggtttggt tctacttcc ctggctgtct
 15061 gtgtcactgg gaacattgta actgccactg gccacgaca gacagtaata gtcgggtca
 15121 tctcggcac ggacccact gatggtaag atggctgtt tgcggagct ggagccagag
 15181 aactggcag ggatccctga gcgccccta ctgtcttat aaatgaccag cttaggggcc
 15241 tggcccggct tctgtgta cactgagta tattgttcat ccagcagct ccccgagcag
 15301 gtgacttgg ccgtctgtcc caaggccact gacactgaag tcaactgtgt cagtctatag
 15361 gagaccacgg agcctggaag agaggaggga gaggggatga gaaggaaaga ctcttcccc
 15421 aagtgagaag ggccctccc ctgaggttgt gtctgggctg agctctgggt ttgaggcagg
 15481 ctgactctg agtgcctggg gaccagggcc ggggtgcagt gctgggggc cgcacctgtg
 15541 cagagagtga ggaggggcag caggagagg gtccaggcca tggtagcgt gcccagcgt
 15601 ctgccttga gcccagca gtgtgggt ctctgagacc ctattccc tctcagcgt
 15661 ttgaggggc cagtgggt ttgggttat gcaattcac ccccggggg ccctcactc
 15721 agaggcggg tcaccacacc atcagccctg tctgtccca gcttctct cgtcttca
 15781 cgtctgaca tcagactgt cctcaggac tgaggtaact gtcacttcc ctgtctga
 15841 ccacatgacc actgtccaa gcccctgc ctgtgtct gggctccca gtggggcgt
 15901 cagcttgca gcgtctggc cgtggactgc ggcatggtgt cctgggggtc actgtgtatg
 15961 tgacctcag agtggtcac tagttctgag gggatggcct gtccagctc gacttctgc
 16021 caagcgtgc tcctggaca cctgtggacg cacagggtg gtccctga agcccgtt
 16081 gggcagcca gccttgacc tctgtctct ggccgctc tctgcccc tctggctac
 16141 cccatgtgt gccttagca gagctgtgt tctcagcat aactgattac tgtctcagt
 16201 acttcatgt cctgtgacg ggctgagta gcatttca cactagaga ccacagctc
 16261 cctgtgtaa gtatcacac tctctctgt gggactttg taaaagatt tgcagccagg
 16321 agtcatgggt ggtcttagt gagaaatgt ggtcagaga gacctgataa ccgatgtga
 16381 gaggggaacc tggaagatct tcagttcagt tcatcagt cattcagt tgtccagt
 16441 ttgggatcc catggactgc cacacgccag tctccctgt ccatcaca ctctgaagc
 16501 ttgtcaaac tcattccat caagttggag atgcttca accatctcat cctctgtat
 16561 ccccttcc tccgcttc aatttccct agcattaggg tcttccgt gactcagtc
 16621 ttcgcatcag gtggcaagt ttggagtt cagttcagc atcagctt tcaatgaata
 16681 gtaaggactg atttctta ggatggactg gttgataac ctgcagtc aaggactct
 16741 caagagtct ctcaacact gcagttaaaa gccatcaatt ctccgtgt cagcttctt
 16801 ttgtgtaaa ctctcacat catacatgac taccgaaat acattagtc ttagaacca
 16861 gttggggct tccacgtg ctctagtgt aaagaatat cctgccaact cagaagatg
 16921 aagagatgc gttcaatct tgggtcggga agatccctg gagaaggga tgacaacca
 16981 ctccagtatt ttgcctga gaatccatg gacagagaag cctgggtgac tgcagtcct
 17041 ggagtctac agagtcagc acgactgaag caacttagct acttggaata gagcatgcac
 17101 gaagctgtc aaaaaacag tcaagaagtc ttgttttg aaggtttact gagaaagtg
 17161 atgactgct ccaacactc ctctcagtg aaaagatcag aagcgtaga tcaatgggt
 17221 gtcaatact tggatgcgt ccaacaggt atctctgag atggaaatga aggcagttt
 17281 tgggtaact ggaggacaag atgagatcat acattggaa cactgtctg catcaaaggc
 17341 gtgtacagta aacattagct gttattgca aaataaatt agctgaatc acccaaatc
 17401 gatggcatt taaagccac tgagtgtaa aatcaggggt gtgcagcaa aacgtccat
 17461 ttgactcatt atgattcca tgcacaaga ctgaaagtc acttctct cagcagaaga
 17521 gaagtagaa catttaacc ttttttga gtgtcaagg aattttgtt acactgtaa
 17581 gtcagtgaat atattgaagc tttcattg tggaaatat taaatatga aaattgaaat
 17641 tttaaattt attcctgggt agttttgt tccagtagt catgcatgga tgtgagagt
 17701 ggactataaa gaaagctgag cgctgaagaa ttaattgtt tgaactgtg cactggaga
 17761 gactctgag agtccctgg tctgcaagga gatcaacca gtccatcta aaggaaatc
 17821 gtctgaata ttcactggaa ggactgatgc tgaagctga actccaatac ttggccacc
 17881 tgatgtgaag aactgactca tatgaaaaga ctcatgct gggaaagatt gaaggtgga
 17941 ggagaagggg acgacagagg atgagatgc tgaatggcat caccgactc atggacatga
 18001 gtctgaataa gctctggag ttgtgatg acaggaggc cctggagtgc tgcagtcct
 18061 gggattgaa agagtggac atgactgag gactgaact aactgagtt ggtaacagat
 18121 atgagaatta tataattaa atctaaact ttgtattc ttttttggc ggttcaaaa

18181 gagctgtccc tctgttaac tatataaat cttttgaga attactaat tgataatgt
 18241 cacaagtat ccaatttctc attacttta gttgtcagta taagaaatcc catttgattt
 18301 atcatgttat agtatctgca actctaagtag ttcagttctg acaaatttt attttattt
 18361 aaaatattgg catacagtaa aatttcaaac aatatacaat tctcccttc agtttaaaaa
 18421 aaaaaacaaa aaaaaagtaa tattagttaa aaaaatccgg gaagaatcca agcatttaaa
 18481 attgcatcac atttctatgc tagacaagct gatataaagt tataattaat aaaggattgg
 18541 actattaaac tctttacata tgaggtaca tggtctcta gcaaaacatt taaaaatag
 18601 ttgtgggtaa attattgtt tcttaaaga aataaaaaga cataagcgt agcaattggn
 18661 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 18721 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnna aaatggataa ggggggagga
 18781 catgggtagg ggagcgcgat ggaggaagta aggtgtcga gggagttggg gggggaataa
 18841 gtgggtaaaa gggaagcggg cggaaggagg gggaagcagg agagagggt gggcgtcaga
 18901 tcggggggag gggatgagg gagagggaat ggtagacggg ggggtggaag cataaaggaa
 18961 aagatagggg ggggaaaagt tagaagaaga atgaggggat aggcggaaag ggaagagaaa
 19021 tgggagaaga acagaaaaat agggggagggg ggggcgtaaa gagggggggg gagggcagg
 19081 gtggagatga cagatacggg gaatgcccc gtataaaga gtatatggcg tgggcgaga
 19141 aggtgtcat cctgtggag gggggacgcg gagaacctt cgggtatag ggaggattcg
 19201 gggggatcgt tcgggaaggc agtcagcaca gcaccacca aggtgtcagg gatggtcgt
 19261 ggggtccaaa gaagaggccc aatccgcgt ctggcagca aggagccctg gagactggga
 19321 agtgtccagg aactgacct aggggtcga ggaaccaga agtgtgtctg tgaagatgtg
 19381 tttgtgggg ggacaggctc agagcttga gcagaaaagc ggcatggcc tgtggagggc
 19441 caaccacgt gatctttt aaaaggttt tttttgatg tggaccatt taaagtctt
 19501 cattgaatt gctacaatat ttttctgtt ttatgtctg gtttctcgg ctgcaagggt
 19561 tgtgtatcg tatctctca accaggactg aaccacagc cctgcactg gaaggcgaag
 19621 tcttaacca gatgccagg aactgctc cctcactga tctaatcaa gacctcatt
 19681 aaggaaaaac cgagattcaa agtccccca ggaggactg gtggggagga gagagccaag
 19741 cactcagcac tcagtcagc acggcgccct cctgtccag ggcgagggt cggccgaagg
 19801 accaccggag acctgtcgg attcaccagt aggttgtga ggaattcaa cttactttt
 19861 aaatctgtc tcaaggctg ttacaagcg accttaccag taacttaaa gtgaaaggg
 19921 acttccagg cggcattgc ggtgaagaac ccgcccgtg gtttaggag acataagaga
 19981 tgtgggttag atccctggt caggaggatt ccctggaga aggaatggc aacctctc
 20041 agtatcttg cctggaaagc ctcacggaca gaggaggtg gcgggtaca gtccagggg
 20101 tcgcacacga ctgaatgac ttacttcaa gttgagacag gaagaggcag tgactgttg
 20161 caaacaccg caccatgct ccaggggac ctgcagcgt ctggtcatg agctgtgta
 20221 aaaaaatca acccaacgag agggccagac agagggaagc tgagtcatc aaacacgggc
 20281 atgatgtga ggagataat caggaaagga cctgccaagc ccatgacaga ccggtgtct
 20341 gtctgagggc cgtctggca gagcagtgca gggccctcg agaccgccc agctccagac
 20401 ccggctggg gctacagggt ggggctgagc tgcaaggact ctgtgtgag cccacgtca
 20461 gggaggatca cctgttgt tttctagtt tctttaaaa tagctttat ggtctctgt
 20521 ctttggttt aaaataaaa ctgttctcg taaacaacgt gaaaaaac aaacaggagg
 20581 aaaacaacgc agccgggca ttaccgag aagagccgc tctaacatt tgacgggtg
 20641 ctttctatt taacctgtt ttactgtaa actgtaaaa ccacatcata aataattaa
 20701 aggtctctg gaagttaaa aagtaagcat ggccgtggcg atggtgtc cacaccgtga
 20761 acgtcgttt caaacggta aatttaggg accccctgt ggtccagtgg gtgagattt
 20821 gcttcattg caggagccgt gggttgtat cctggtggg gaactaatg cccacatgt
 20881 gtatggagt gccaaaaa atttttga aatggtgagt tttaggtgac gtgaatttc
 20941 cattgatga cttcacaggc tcagatgac ccaggccctc aggaagccc agtcaccgg
 21001 tctttactt tcttagag tttatggt tctgttctg ccttaaac caccatgtt
 21061 caacctcctc tgatttga cttataata aagttaggt gtgttcagg aaacttgc
 21121 cagtattctg taataatca aatggaaaga attgaaaa agagcagaca ctgtacatg
 21181 cataactga tcaattgt gtacacctga aactgagt cagccgtca gtcgtgtccg
 21241 acctgcgac ccaaggact gcagcagcg ggcttccct ccatcacca actccggag
 21301 ttactcaa cacatgtcc tcgactcgt gatccgtc aacctctc tctctgtg
 21361 tccccctc ctcgcctt caatcttc cagcatcagg gtctttcaa atgagtcagt
 21421 tctcacacc aggtggccag agtatggag ttacagctc agcatcagc ctccaacga
 21481 cccccatac ctgaagctaa cacagtgt atccactgt ctgcaacatg aaagaaaaac

21541 acatttttta agtttaggct gtgtgtgtct tcttctctc aacactgcgt ctgacccac
 21601 ccacactgcc cagcactgca tccccgtgg acaggaggcc cctgccccca cagctgcgtg
 21661 ccggccggtc actgccgagc agacctgcc gccagagtg gggcccttg cactggggac
 21721 aaggcagggg cctctccagg gccggctact gtccactgtt cctactggtt ttgtttcaa
 21781 aagtggaggc agcgtaatat tccctgatt ataaaaagaa gtacacaggt tctccacaaa
 21841 taaaacaggg gaaaagtata aagaatggaa gttcccagca cagcctggag atcacgccgg
 21901 gtgcacctgg ggtgtccttc caggctggac ctacatttc acgcagacat cagaaggctg
 21961 cgagatctac ccagaaggct ggtagatgg ggtagatgg agtgacaaac agtagacaga
 22021 gagatataca gacagatgat ggtagacag acgctaagac accgagcag gggacagacg
 22081 gatggaagac accatcctt gtactgacc acacaccac atgggtgtgg tgagccggct
 22141 gtcatacttg tgaacctgct gctctcaca caccagctgg gtccctccag cccagcgtc
 22201 ccacacagca gactcccgcc tccatcccca ggcaggaatc ccaccacaa ctgggggtgga
 22261 cctccccgc aggaaggctg tctgtctaa ggcttgaga gcaagtaca gacctctc
 22321 tgggaagaca gcgacaacc gcctaccccg cagagcccag gaggaccct gagtctagg
 22381 gaagggacca cggggcctgg acggggagcg gcccaggac gctgccccca acctgtcca
 22441 cctcactcct gctctgctc gaggcggggc gcagagagg gcctgaggc ctctccag
 22501 ttctgggag caccactgg gcctgaacca ggccagaagc cccctctca aggtgtccc
 22561 agaccactcc cctccacctc cgggtgctc gtctctggc agcagggagc cccagtgaga
 22621 agagacagct ccaggctgtg atcttgccc ctggctgctc tggcagtgtg ggggggtggg
 22681 gtcgctggga ggcatgagt gctgggggct ggggctgta aagcacctc aggtcagtgg
 22741 gctgttggtc gggctctcg aggtccgac ggtagagct gtccaggac acaggaggcc
 22801 tggcagtgg tccaagagt cagggccaaa ggaagggtt cgggcccctc tggctctca
 22861 gctctgagg ccggggacc cagtctggc ttgtagggg ggcgattga ggtacaacg
 22921 atccaaaaga aaacacacat ctacaggga agagtctga ggaggagaga gctacacaga
 22981 gggctctcac actgcggaca ctgctggag tctgagagct cgagtgcgg gcacagtga
 23041 cgaaggagg acggaacct caaggacac ggacgccat ggccagagac acacgcagt
 23101 cccatgagg ccggctgctc agacgaggg gagctctca ttaaggctc tgcgtgaata
 23161 gtgaggagaa ctggcccgt gtgtgggaa acttagcca gaagaaacgc tgcctggcc
 23221 ccaaggatca nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 23281 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn tgccttgc
 23341 ctccaggag ggaggaagc tggatctgg gttgcctg ggttaaagg atccaccac
 23401 tccctttta gccactcct gtctggcaa ttcttaaga ctggaggtc caaagattg
 23461 gacacactga gcgagtgaac tgcactgagc ctaagaaaag tcttgaatt cctccaaaca
 23521 aaacacactt gtctgggta ctctcttg tttgttaca aatgtctgt cctctgttc
 23581 tctggccag ctctgggtg tcatctgac ctgacgaag caaaggagc ctggaccctc
 23641 aaaatctga ggaccagca cccctccatt acactctgt tccccgca acgggcagt
 23701 gttcgccgt ctggcgtat gttaagcga cgggtgata ctgggagtc ttactctgt
 23761 tcttttct ctggggtag accaccatc gcacactct gtctgaatg gaacatttg
 23821 gtgattgat gtggccaga ccccccaac gaatgtacct tcaggttgt ttctcttt
 23881 tatatttgc ttgtgaat agacacagga tccatcagt tctatagt gagaaagtaa
 23941 aaaccactc agccttagt ggatggagat ctatagtaa gatagcagt tagccgaaa
 24001 tggaaattc agccagaatc tgaagcgt gtctggaag gagaagagg actcaggccc
 24061 gagcacactg ctccacgtg gacctcagg ctctgacagc tctacctcc ggggtctca
 24121 tgggacaggc catgcaggc acgatccgt tgagaagtt ctgcttcc catcacattg
 24181 gcaattgac gcttgcct tctctaca tggagtta ctttatccc agacagttg
 24241 gttctctc tgaattgc caattgata gatcgtaca gtattctta accacataga
 24301 attcggcagg ggggtggg ggacaggga ggggtgggtg agagtgggg gagggggctg
 24361 caccgagcag catctgggt ctgactccc tgcggggat agacctctg cccctcagt
 24421 gacagcacag agtctctc tctgaactgc caggagcgt cctgcaattg acttaagaa
 24481 aggcactaa ttaggaatt tgggtgaca ttacattt aagtgtgta gcagtatta
 24541 tagttcatat cttttatag ttctgtatt ttactagct aaagggttt tgggtttct
 24601 tttgttta aaagctaaa tctgtttt aattccatg aatacaaaa aaaaaagtct
 24661 gtagaatatt ttaaagtg aaggcttgt tgggaatgt agcgtttgc tccactgaac
 24721 cgaacgtaa taacattgt agaagagac cagagtga ggtacctt ttattgagt
 24781 gacatgacag caccatcgc gtgagtatt ggctggagt tagagacagg ccatgtggg
 24841 ctaactcct tattgtgt ctacgctt gagtaaat cagaagctt ctctgaagag

24901 agtgggggtca gctgtcagac tcttaggtgt ctacctgacag cagggtggg attaatgca
 24961 gcagccagta gatacgggat ggggcaagag gtcacctgt cctttgtg ctgctgggag
 25021 agaggctgt cctgggtcca gtggggccaa agctgtgact ttgtaccac aggatgtctc
 25081 tgacctgcc ttgggttccc tgagggtgga gggacagcag ggttccccg gttcctggc
 25141 cggagaagga cccccaccc ctgtctctc gacatcccc caggactgc cccggagtag
 25201 gttctcagg atgggcatcc gggccccacc ctgactcctg gagctggccg gctagagctt
 25261 gctgcagaat gaggccttg ccattgcgc cctgaaggag ctgcccgtca agctctccc
 25321 gaggtgttt acggcgccct ttgccaggag gcacacccat gccgtgaagg cgatgggtgca
 25381 ggctggccc tccccacc tccgatggg ggccctgatg aaggactacc agcctcatc
 25441 ggagacctc caggctgtac ttatggcct ggacctcctg ctgctgagg aggtccgccc
 25501 taggtaaggt cgacctggca gactgggtgg gccctgggtg tgagcaagat gcagccaggc
 25561 caggagatg aggggtcacc tgggaacagg cgttgggtgt acaggactgg ttgaggctca
 25621 gaggggacaa aaggcacgtg ggctcccc ccagtgtccc taaagtggg aaccaagggg
 25681 gccccggaag ccggaggagc tgtgtgtgt ggagtgcaga gccctgcgg ggtcctgatg
 25741 cccgtcggac tctgcacagc tcagcgtgt ccccgcgcc cggtaggcgg tggaagctgc
 25801 aggtgtgga ctgcgcgg aacgcccacc agggactct ggacctgtg gtccggcatc
 25861 aaggccagc tgtgtcact gctggagccc gagtcagccc agccatgca gaagaggagc
 25921 agggtagagg gtccagggg tgggggtga agcctgtcc gggcccttg gaggtgtgg
 25981 tcacctgtg cctcaaggag gacacgtgg acgagacct ctgtacctg ctgaagaagg
 26041 ccaagcagag gaggagcctg ctgcacctg gctgccagaa gctgaggatc ttgccatgc
 26101 ccatgcagag catcaggagg atcctgaggc tggtcagct ggactccac caggacctg
 26161 aggtgaactg cacctggaag ctgctgggc cggatgggca acctgcggg ctgctgtgt
 26221 cgtgcatgc cctgttgcg cgcaccgcc ccgaccgga ggagcactgc gtggccagc
 26281 taccgcca gtctgagc ctgccccacc tgcaggagct ctacctggac tccatctct
 26341 tctcaaggg cccgtgcac cagggtctca ggtgaggcgt ggcgccagct ccaagacca
 26401 gagcaggcct ctctgttc gtcccgtg gggacattg cagggtgcc ggccactcg
 26461 aagtctcac gatgccacc ctctgacct gggcatctg tcaggtact tccctgtta
 26521 gggtcagagg cgtggcctag gtaaatgct gtaaaaggg actccttct gggagtccg
 26581 atagtgggg ctggtgtga tgccttggg aattcttcc gagagagtga tgtcttagct
 26641 gagataatga cagataacta agcgagaagg acggtccac aggtgtgagg ttgaagtc
 26701 aaagctctg ctctccctc cactgcctt tctgtctg agctgttta ggtccagg
 26761 gagctgtgg aagtgggtga tctggagat gacaagaagg gatcaggagg ggaattgt
 26821 ggctcctaag cagtcagag aagagaaaa gtaataag cattattgt aaagtggc
 26881 cagtctctt aagtccaaat tataattata atttctct aagactctg aatacatag
 26941 aatcctcag taacaggta ttgtctgcc tgaacacag tgataaaagc tgggaggatg
 27001 cagcctaate tgtctgtg aatgagttg attgattcc ttttggcag ctgcaactc
 27061 caagcattag gaataaatat gtactgag aacccgaag aaagaaaga agaaaaaaa
 27121 aaagaattg aggtgtgat ggacggttg tggccctga atcttggg gatgttacc
 27181 cagggatcac gtgtaactg tgggacccc agcccatgt cactgcac cagcctgctg
 27241 tgaattccg cggatcnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 27301 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnncaat
 27361 tcagctcgg taccacaaag gtccgtctag tcaaggctat ggttttcca gtgtcatgt
 27421 atgatgtga gattggact gtgaagaaag ctgagtcca agaattatt ctttgtact
 27481 ggggtgtgga gaagactct gagagtcct tgaactcaa ggagatcaa ccagtcggt
 27541 ctaaaggaga tcagctctga atgtcattg gaaggactga tctgaagct gaaactcaa
 27601 tactttggc acctgacgtg aagagttgac tcattggaaa agaccatgat gctgagagga
 27661 attgggggca ggaggagaag gggacgacag aggatgagat ggctggatg catcaccaac
 27721 tcgatngac atgagtttg taaactcca ggagtgtg atggacttg aggcctgtg
 27781 tctgtggatt catgggtcg cagagtcgga catgactgag cgactgaact gaactgaact
 27841 gagctgaaga gtcacctgt accagagctc ctgagtcct cctgcaggcc tggctgta
 27901 ggccccagg tcaccgtct gcctcttca tccatctt tcacgacagg ctgggagtgg
 27961 ggtgaggtga gttgtctgt atctagaatt tctgcatgc accctcagag tgcaattag
 28021 ctccagagaa ctgagctca agagttcatt tttctttt ctctttatg atactacct
 28081 ctctgagca gagacctcat gtcagggaga aggggactct gcctctca gcctttgt
 28141 cctcaagac ccacagggg aggtgcct gcttactga gccggaagt tcaattgct
 28201 atgtctcca gaaacaccc cccccaga gacccaga aataagtga acagacctt

28261 gtttcccaga caagtgggac acacgttatg aaccacctca gtgattaaaa tagtaacctc
 28321 tgtgtatgtg tatttactgg agaaggaaac ggcaacctac tccactattc ctgcctagaa
 28381 aattccatgg gagagaagcc aggcaggcta cagtcacgg ggtcacagag actgaacata
 28441 cacaagcaca tgggaagtga ttttcagta ttttaaatt tgttcagttc aacatggagt
 28501 acaagaattc aaatcgtgaa gtcaattgac caagaaacca gaagaaatca ctgtgtgtg
 28561 atctctgtgg aggtaacatg ggtacctgtg ctctgacctt cacagcctct ggctctctt
 28621 ctacatgtac atacacatat attccatgt atgtatgtat tcggaagatt tcacatacgt
 28681 ctaccagtc cacagccccc gcgttccctg atgccagaa catctgtgat agctgtgagt
 28741 attgtacca gataagatct tccaggttcc tgcactcaca ttggttatca ggtctctctg
 28801 atccagcatt tctcagctaa gattcctgt gactcctggc tgcagaatct tctgcaaaag
 28861 tcccacagag aggagtgtga tcaactgtaca caggaggggc gtggttctct agtgtgagaa
 28921 aagctaactc agcccgctac agggacgtga atgtacctga gacagtaatc agttatctg
 28981 agaaatcaca gctctgctag aggcagcaca tggggtagcc agcagggggc agcagagcac
 29041 ggccaggagc cgcagggtcag aggctgggct gcccagcgg ggttcaggg gaaccagccc
 29101 tgcgggtcca caggtgtcca gggagcagcg ctggcagga agtcaggacc ggacaggcca
 29161 tcccctcagg actagtacc acctctgagg gtcacatcca cagtgaacct cagagacca
 29221 tgcctcagtc caggccagg acgtgccag gctgaccgcc ccactgggga gtccagggga
 29281 gaccacagc cggggggctt gggacagtga tcatgtgtc agacacagag aaggtgacag
 29341 tgacctcagt cctgaggac aagtctgatg tgcagacgtg agaagccgag gaggaagctg
 29401 gggacagaca gggctgatgg tgtgtgacc ccgctctca gtgagggggc cccgggggtg
 29461 aattgcata aaccaagcc ctactgcc ccacaaagt ctgagaggga ataaagggc
 29521 tcggagagcc cagcactgt cggggtcag aggcagagct cggggcgcgt ccacatggc
 29581 ctgggcccct ctgtactgc cctctcac tctctgca ggtgcggcc cccagcctg
 29641 gtcccaagt gaccaggcct caggctggcc tgcagctca gcacaggggc tctgcaggg
 29701 aatcggggc gctgggagga gacgtcttc ccacactccc ctctctcc tcttctag
 29761 gtcactggtc ttctctcag ctactcag cgcctcggg tccgtgtcc ttggacaga
 29821 cggccagcat cacctgccag ggagacgact tagaaagcta ttatgtcac tggtagcagc
 29881 agaagccaag ccaggccccc tgtctgttc attatgagt ctagttagag acctcaggg
 29941 atccctgacc ggttctctg ctccagctca gggaacacgg ccaccctgac catcagcggg
 30001 gccagactg aggcagagc cactattac tgcagtcac atgacagcag cggtagcct
 30061 cacagtaca cagacagac gggaagttag acacaaacct tccagctct ctacgtct
 30121 cctccagccc cgggaggact gtgggcacag caggacagg cctggcccgg tccccgga
 30181 gctgagccc caggcggccc cgcctccgg cctccagc aggtctgca caggggcgtt
 30241 agcagtggac gatgggtg caggccctgc tgtgtcggg tctgggtgt ggagtacat
 30301 ggagaacgga ggctggatg aggactaaca gaggacaga gactcagtgc taatggccc
 30361 tgggtgtcca tgtgatctg gctggacct cagcagcaa aatctctg attgaccca
 30421 gaacttcca gatccagac cactggct tagaaaggct taggaggtga acaagtggg
 30481 tgagggtac catggtgacc tggaccagaa ctctgagac catggcacc cactccagt
 30541 actcttccct ggaaaatccc atggacggag gagcctggaa ggttcagcc catgggtcg
 30601 ctaagagta gacacgact agcgacgtca ctctccctt tcaattcat gcattggaga
 30661 aggaatggc aaccagtc agtgttctg cctggaaaat cccagggaca ggggagcctg
 30721 gtgggtgccc atccatggg ccacacagag tcagacaga ctgaagcaac ttagcagcag
 30781 cagcagcag ccaataaaac tcagcttaag taatggcatc taaatggacc ctattgcaa
 30841 ataaggcca ctgcgtgca ctctgttag gactcagtt cctgattgt gaggggtccc
 30901 acaagacgtg tgtgtatatt ggtgtgccc gaaaacagt tcaatgtgag catccagac
 30961 tcatccct cctactcca ctattccatt gtctgtcag gtattaagca taaaggtaa
 31021 gggcttatt agatgaaga ggagtgaata ctgtctgtg cttaacacat accaagtacc
 31081 atcaaggtec ttctattta ttaactgtg ttttaacag aaatagcta ttagaagca
 31141 tccggacgat agccatgtt acagacggg aagctgaggc atgaagtct cagcacctg
 31201 ttccagctc gacctgaaac gggcagagc cggcagcaa caaggctct ctccaagc
 31261 gccgctct caccgctc ctatggctt tcaactgtct tctaaacta agctctccc
 31321 aacctgtg agacaggatt agagacttta ggagaaaaga ccaggacat cccacaccg
 31381 accgagtga gccactaaga caaggcttg taaggacaga accagcaggt gctctcagc
 31441 agccaggag agacctgca ccaaaaaca tattgtagca tctgacct ggacttga
 31501 cctccagaaa tgtgaaaag aaacgtgtg ggttaatca actaccggt gttattgtg
 31561 tatgactgcc tgagttaaga aggagtggg aacactgag ttaggtgt tatggaacat

31621 aagtcttgtt tctctgaaat aaattcccaa gggataatt cctaggttgt agggtaactg
 31681 ccacaaatct aggcagctta ttaaaaaaca aagatatcac ttgccagca aagggtcata
 31741 tagtcaaatt atggtttta tagtagtcat gtatggatgt aaaagttgga tcataaagaa
 31801 ggctgagcac cagagaattg atcccttcaa atcgtgtgtc tggagaagac tcttgagagt
 31861 cccttgaca gcaaggagat ccaaccagtc aatcctaaag gaaatgaact gtgaatattc
 31921 actggaagga ctgatgctga agctgaagat ccaatacttt ggccacctga tgcgaagagt
 31981 tgactcattg gaaaagaccc tgatgctgga aagcttgagg gcaggaggag aagagggcgg
 32041 cagaggatga gacggttgga tggcatcact gactcaatgg acatgagttt gagccaactc
 32101 tgggagacag tgaaggatag ggaaggctgg cgtggtacag tgcattcggt cacaagagt
 32161 ctgacacatc ttagtactc aacaacgaca gcaacacagg catcacacgc ttagtgtgat
 32221 aagcggcaga actgtttcc aggggtccgn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 32281 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 32341 nnnnnnnnnn tacgattcga gctcggacc tgacattgt agtcacgtca tgacagctg
 32401 tttccggtc ttagggtatt gtggacgatt tctgttggg ttgctcatg ataatttagt
 32461 tacagcttag gtcttctt tccaggccac gagcgacatg tttcaggtg agatgacgtg
 32521 gtgggggatg ggcggccaag cccccactgg ggggggaggg attctgtgt gggcaggagt
 32581 tggcagcatc cctgaactga tgacctgca tccaggtgac aagaaccggg ggatattatt
 32641 cctctgctt ctcattgcat gtctcgggt cttcatgatg aaaacatatg acaatacagg
 32701 ggagttgat tgggcgggc acaactctgg gtgggggacc cggtggcatt gtgccagca
 32761 gggccatcaa gatgaggcg acctgggtgg tcccttctc cctggggtc ttggtttcc
 32821 cctcatgga atgggatcag gcagcagcca tggaacaccg cgacctggc ttcttcacc
 32881 tctcgtctg tgatttggg tgggatacc aggcattgag acctggggc gggggacatc
 32941 actcctctg agcaggagg cgcagagtc ctcgtccat gaggactcg tccctgggt
 33001 gacctgagg actgctggag gctgaagctg gaggcacagg cggctgca ggcagggtc
 33061 ctgaggacga cagagccagt ggggctgcag ctctgagcag atggccctc gcccgggccc
 33121 ctgagctgt gtgtccagt gcaggtctc tcaggtgagc cactacgtta tgggggaggc
 33181 gccctggga gggatcggg gtgtgactc ctccagatt cgcacctct gggagcactc
 33241 tggccacct ctaagcctg caagagctg gtcatcagt ctaactctc tctgaagtc
 33301 caatggactc tctcatgag gcagtcact gatggcctt ttatccccga tgggtgctt
 33361 ttcgctgac ctggtctcc tgaccacctc ccagccccc accatacagg aagatggcac
 33421 ctggtccctg cagagctaag tccaccctg gcctggctc agatgcctac agtctctg
 33481 cgggagggc cgtccccac taggccccaa gcctgccgtg tgagtctcag tctacctgg
 33541 aacctctc atttctccc agtctcagc tcccaacccc agaggtatcc cctgcccctt
 33601 tcaagccct tgccttcc tgggggatg ggggtgatg gagggcaagc ctgatcccc
 33661 gagcctgtc cgtgacaat gtccgtctt ggatcatgc tccctggct ctgagctc
 33721 cctggtccct ggggatggg tgcgtgatg acaagtggat ggactctcag gtcacacctg
 33781 tccctccct aaggaactga ccttaaccc cgacactcg ccagaccag aaagcactc
 33841 agacatgctg gctgataat gagaaggctt ttattcagga gaaacaggaa cagggaggga
 33901 ggagaggccc ctggtgtgag gcgacctgg taggggctca ggggtccatg gagaggtggg
 33961 ggagggggtg tgggccagag gggcccgag ggtgggggtc cagggcccta agaacacgct
 34021 gaggtctca ctgtctctg caggtgtc cctcgtgag tgacctgca gctgtaactg
 34081 ccttctgatt tccagtcgt gccgtcagg ctcatgagt gctggcccg tatttctgt
 34141 tctctgtt ggaggcccgt gtggtctca cgtgcgggt gatggtctg cgtctgct
 34201 tccaggccac ggtcacgta cccgggtaga agtcgtgat gagacacacc aggttgccct
 34261 tgttgccgt gagctctc gtggggggcg ggaacagggt gaccgagggt gcggacttg
 34321 gctgacctg gtggacagag gagagggtg aagacggcg ggaggtctg acctgtccc
 34381 cacgtagcc ctgttgcct tctctgtcc ctccgacct tgcctcagc cctgggccc
 34441 cagacagccc ctcagaagcc attgcaatc actctccaag tgaccagcca aacgtggct
 34501 cagagteccc ggtgcgacc agggctgctc tctcctgct tctggcccc gggagtctg
 34561 gctgctctt ggcactgacc cctgagccc tcagccctg ccagaccct cgtgacctt
 34621 ccgtcatgc agcccagggt cctcctcgt gaaccgggt ccccccggc acctgccagg
 34681 acgtctctga tggagatgt ggggacaagc gtgctagggt catgtgcgga gccggggccc
 34741 ggctctctc tctgctcc gcccagctc agctctctg gccaaagccc ggggtctctc
 34801 tgaggtctg cctgtctacc gtccgctg cctgagtga gggccctc cctcacctg
 34861 ctcaggga cgtgcccc acacagcacc tccaaagacc ccgattctg gggagtcaga
 34921 gccctgtca tatctctaa gtccaatgt cgttcgagg ccagcggagg ccgacctcg

34981 gacaggtgtg acccctgggt cccaggggat caggtctccc agactgacga gttctgccc
 35041 catgggaccc gctcctttct gaccgctgtc ctgagatcct ctggtagct tgcctcgtct
 35101 cagctgtgtc caccggccc ctacggccag agcgggcgag acccctctct ctctgccctc
 35161 cagggccttc cctcaggtg cctctgtgt tctggggcc tggtagatgc ccccgccgag
 35221 cccccaagct cctgtctggc ctcccggtg gggcatggag ctacagcac agagccggg
 35281 gcttgagat gcccctagtc agcaccagcc tctggccgc acccagcgt ctgccctga
 35341 agaggggaac aagtcctgc attcctggac caaacaccag ccccggcgc ccgactggcc
 35401 ccattggacg gtcggccact ggatgtcct gctggttacc ccaagacaa cccgcctccc
 35461 ctcccgccc caccgagaaa ggtggggatc ggcccttaag gccgggggga cagagaggaa
 35521 gctgccccca gagcaagaga agtgacttc ccgagagagc agagggtag agaggctggg
 35581 gtaggtgag agccactac ccaggacgt gaccaggtc ccggcgcta agacaaaata
 35641 cagagactaa gtctggacc aaaaccgccc gggacagcg ctggggcctg tccccggg
 35701 gggctgggccc gagcgggaac ctgctgggcg tgacgggcgc agggctgag ccggtggggc
 35761 tgtgtctcc gctgaggggt gttgtggagc cagcctcca gaggccagg gacctgtgt
 35821 cctggaggtg cctgtgccc agccccctgg ccgaggcagc agccacacac gccctgggg
 35881 tcaccagtg cccctcact cggaggtgt cctggccacc actgacgct tagcgtgag
 35941 ggagacgtg agcggcgct ctgtcgggg cggcagagga gtaccggcct ggctggacc
 36001 tgcagccg ctctggcct cactgtaagg cctctgggtg ttcttccc acagtctca
 36061 cagtccagcc aggcagctc ctctgggg ctgtggacac cgggtatc ctacggccc
 36121 aagtgggaa cctgccctt ttctccacc caggagatg cagttcagt tgtctctc
 36181 aatgaacatt ctctgtgtc agatcactgt cttctgtac atctgttgt ccatcatc
 36241 atccaacat catcatcca tccatccc agccatcat ctgtcatca acatcatc
 36301 ttcatcat tgtcatca tcttccatc ttgcatctgt ctgtcaaca gtggccatca
 36361 agcacccg tcaccaagccc tgtgtcacac gctgggactt ggtgggggga gccctgccc
 36421 tcccaccct ccatctctc tgaactct ggggtcaagt ctaacaagt ccatccgt
 36481 ctactctg gtcctccgc agcctctct tccactct ctgctctga cccactgt
 36541 gcactggac gaccaccag ggcccttga tctctgtc ctctgacc tctttttt
 36601 ggctctggat ttatacat tctgctct ggagcgct cagctgagt gtccacaga
 36661 cgcctcagac tcagcatct ccatgaaac tgcctcagg tcttcaga cctgtccc
 36721 cacatgtc tcaattcgt agatttctc acaagccaga ggctggact catccataa
 36781 tgctgccc tcatgagtc agcctctgt tctaccata accaaacat ccttaaaaa
 36841 tctcagaaga aaaaaaaag caccagatg gcactgtc agtttatgat gacaagaatc
 36901 ctacttcag tcatcact cagctgtc cgactcttg cgacccatg aatgcagca
 36961 cggcggcct cctgtccat caccactcc cggagttac tcagactac gtccattgag
 37021 tcagtatgc catccagcca tctatctc tctgtccc ttctctct gccccaatc
 37081 cctccagca tcagagttt ttcaatgag tcaacttc gcgtgagtg accaaagtac
 37141 tggagttca gttcagcat cattcttc aaagaaatc cagggtgat ctcttcaga
 37201 atggactgt tggatctct tacagtcaa gggacttca agagtctct ccaacaccac
 37261 agtcaaaag cctcaattt ttggcgctc gccttctca cagtccaact ctacatcca
 37321 tacatgacca caggaaaaac cataacctg actagatga ctttgttg caaagtaatg
 37381 tctgtctt ttaatatgt atctaggtg ctataact tcttccaag aagtaagtgt
 37441 ctttaatt catgctgca atcaacat gcagtatt tggagccca aaaaataag
 37501 tctgacctg ttccactgt tcccactct attccatg aagtatgg accagatgcc
 37561 atgactttg ttctgaat gttgactt aagccaact ttactctc acttactt
 37621 tcatcaagag gcttttagt tctcttcac ttctgcat aagggtgtg tcatctcat
 37681 atctaggtt attgatatt ctctggcaa tctgattc agttgtgt tcttcagtc
 37741 cagtgttct catgatgtc tctcatata agttaataa gcagggtgat aatatacagc
 37801 ctgacgtac tctttctt atttgaacc agtctgtgt tcatgtca gtttaactg
 37861 ttgttctg acctgcatac agatttctc agaggcaggt caggtggtt ggtattcca
 37921 tcttctcag aatttccac agttgatt gatccacac gtaaggtt ttggcatagt
 37981 caataagca gaaatagatg tttctgaa actctctgc ttttcatg atccagcaga
 38041 tgttgcaat ttgatctg gttctctgc ctttctaaa accagctga acatcaggaa
 38101 gttcacggt catgtattg tgaagcctg ctggagaa ttgagcatt ccttcttag
 38161 cgtgtgagat gagtcaatt gtccggcagt ttgagcatt ttggcatt ctttcttg
 38221 ggattgaa gaaactgac ctgtccagg cctgtggca ctgtgagt tttcaatt
 38281 gctggcatat tgagtgcag acttcacag catcatctt caggattga aatcgtcca

38341 ctggaattcc atcacctcca ctactttgt ttgtagtgt gctctctaag gcccaattga
 38401 cttcacattc caggatgtct ggctctagat gactgacac accatcgtga ttatctgggt
 38461 cgtgaagatc tttttgtac agttctctg tttattctg ccacctctc ttaatatct
 38521 ctgcttctgt taggcccata ccgtttctgt cctcgcctat cgagccctcg cctccctacg
 38581 tagagactct aagcaggaag gtgacccgtg ctgcactggg tccagcatgc tttaattca
 38641 gcagtggaaac ttctgggtca tgattgtgt taaggatgc gcatacgatt ttgaagcaa
 38701 aatttaacag gacagcagtg taaagtcagt acttattct gattaaagaa agcaaatac
 38761 cagcctgtta ctaagttaat taactaaaga aacatcttca acttaataaa cagtatctc
 38821 tgaaacttac agcatgcttc acatttaaag gcaaaacat ttagaggcc agggttccca
 38881 cgcttacgtt tattatttaa tatatgtac agattcaagc ccatgacaca aaatggggg
 38941 aagagtgtga gtgttaggaa aaatgagata aaattggtt ttgcagggtga tgggctagt
 39001 tactttaaaa aaaaaaaca aacaagctca agatgaactg aaggactatt agaactgga
 39061 caagagttaa cctgtgatc aatacaagca ggctgggcaa aactcagcag gttttctct
 39121 atacaggcag taatgattga gaatacgaac cggcgggaag gcttacaacc tcgataacag
 39181 ttctattaaa agccctagga atgaactaa cagcgnnnnn nnnnnnnnnn nnnnnnnnnn
 39241 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 39301 nnnnnnnnnn nnnnnnnnnn cccaccctc cctcctccc cccaccac cagtgcacca
 39361 ggtctcgtgc ccagagagct gaagatgcca gcaggccgc tgcctgcctc gctcgcgtg
 39421 cccgggctcg ctgcccgtc gcctgccag cacacagatg cagcccccag tctcgtgctc
 39481 acccgcctcc cccaggcagg actctccac aacaccaagg gcgtctctg gttcaggatg
 39541 gccctcgtg aggtgtaaag tgcttcccg ggctgagacg aatgggcccg agatccaaac
 39601 gaggccaagg ccgccagcg ccttgccga gggcacccat ggtgcagagc ggcccagctc
 39661 cctcctccc tccctccc cctgcttct tatgtccc gctatgteta ttttactct
 39721 gcaattaga aatgataccg aaggacaaac accgttccc ctgtgtgtc gctctaaac
 39781 cttatctac ttatctatta gctgtccaa gtttgcgtc taagtgaatg aaggaacact
 39841 acccacaagc agcaacgtc ccagaccct cgctgttca actgggaatg taaatgtgt
 39901 tcaaaggac ctaagttct atgtcaaaa ccgtgtgtg ttcttttg gagtgaacct
 39961 aggcactcg ttgtctgcc ttcaaagca ttcttaaca cctccagaa cccagggtt
 40021 ggcttacgtt tccagaaat ccaagacag aacttgga accgtatgaa gaaggcctgt
 40081 gagcacagca ggggccggg tacttgaggt aggtggggg ctcggtgctg atggacagg
 40141 cctgtactt ctcactgtt ccgtccagga tctctccac ctcggaggct ttcagcagg
 40201 tcacgtgtg gcccagggtc gtgtatccat gatctgaac cagagacggg gctgcgtgca
 40261 gccgcgggc gggcagcagg caggagcagc caggagacgc agcacaccga ggtcctcaca
 40321 tgcaggaggt gggggaagcg gctgtggacc tcacactgc ccgatgtgg cctcttccaa
 40381 agggccggcc tggaccctg ctttccag aggcctgtc gggccgtcc cagaggctc
 40441 agccacagg cctctggga caggagggt ccagagtga ccggccggc ggaagaggc
 40501 tgacaccgtc gcagtcaca acacgaagc aggtggagat gggatgagg atgagaaaca
 40561 ctttctttt aaaacaagc cccagagat tggaagagc tctgcacac gcaacatgaa
 40621 ctcctggccc cgtgcccag ggcgtggga gccgagttc tcggcaatcc gaccacagt
 40681 tgcctaggga gccgggtgga gacggagggt taggggaagg cggtcccca gggagcgcga
 40741 ggccggggt cgccaaggct cggcagggc aagcgcagct aggggcgag ggttagtgac
 40801 cggcactgca cccggcgag gaggccagg gaggggctga aaggtcacag cagtgtgtg
 40861 acaagaggct ccggtcctg cgttaaaaga acggtgga cagaccacga cagcggcag
 40921 gacacactca taccggacg actgcggagt gcacgcgcgc gcacacacac acacacaca
 40981 cacacacaca cacacggcc gggacacact cataccggac ggactgcgga gtgcacgcgc
 41041 acacacacac ccaccacaca cacaccac acacacacac ccaccacaca cacacacaca
 41101 cacacacac cccacacaca cccacacaca cccacacaca cccacacaca cacaccaca
 41161 cacacacaca cacacacaca cacacacag gcccggtggc cccaggcgca cacagcacg
 41221 agcaaacatg cacagagcac agagcgagcg ctacgggacc ggctgccaga ccaggcgcca
 41281 cgcgatgat tggggcggg gacggggagg ggcgggagca aacggnnnnn nnnnnnnnnn
 41341 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 41401 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn aaagaagccg ggagcgagaa tatgacggca
 41461 agaggatgta ggtggggcg gggcaagagt aaagagagcg gacggtagag gggatgcgat
 41521 tgtgatcgg aagcgagac agagtgatg ccgtattaga ttgatgcaa gaggaacagt
 41581 agggggggg ggggagagga ggggaggtg gggggtgtg ggtgggaagg gaactttaa
 41641 aaaaagagg gagagttga ggggggaata aacggcggt aaaaagaac aattgaaat

41701 taccaggggtg gggcggccag ggggggtgatt cattcttga gggggcaaca tatggggggt
 41761 ggctgtcgcg gattaggaga aaataaatat caggggtgat taagtgttg gcgttgggga
 41821 ataatgaagt aagaatcaaa tatgaatcgc gttggcatcg ttagccatcg ggggaacat
 41881 ttcccatgca aggaacaagg atgtgagaat gcgtccgtct gaaccaccgt cccgggggtcc
 41941 cagtaggact cgcgagctg atagttgccg gagcaacagt taaggagca gaagctgcta
 42001 caaaaccacc acctgcaaa gtagggctc caattacgga gtgcgctcc tgggtgtcgg
 42061 tccaaacctt tggaaggac ctggaataa gtgtacca ccagatatta atataaacc
 42121 acctggccag gagaggcagg cgctgtgc acaggaagtg tcccagact cagtcataa
 42181 ggtaataat atttgggac ctccctggaa atccagtgt taggactct cggttcaatc
 42241 cctggtcggg gaactaat cccacaagtc acaagacatg gccaaattt aaaaagaaaa
 42301 aaagagagag aatatattag tgcaataggt ttagaattg aaattaagct cctgcccacc
 42361 cccaccccc aatctggatg aataaagcat tgaaatagta agtgaagtca ggcttgaca
 42421 tgcactgatg tgactacct taagcaaccc ccaccctagg actggtcggg gtccaggag
 42481 ttccaggggt gccaggaaga tggagtccag cccctgccct ctccccac cagtcctcc
 42541 actggagcgc cctacccac ctccacccc tccgaccct gctaccccc acccctgcc
 42601 ccaggtctcc cctgtctgt gtctagctc cacactttt gggcagtgt tcccttaca
 42661 gctggttct gctgccgct accggggccg tccctctgt tcagttcagt tcagtcctc
 42721 agtcattgt gactcttct gacccatgg actgcagcac accaggcct cctggccatc
 42781 accaaccacc agaactact caaactcatg tccatcagc cagtgtgccc atccaacct
 42841 ctcatctct gtcgaccct tctctggcc tcaatctt ccagcatcag ggtctttcc
 42901 aatgagtcag ttcttgcag caggtagcca aagtattga gttcagctt cagcatcatt
 42961 tcttcaatg aatattcagg actcatttcc ttgggatga actggttga tctcttga
 43021 gtccaaggga ctctcaagag tctctccaa caccacagt caaaagcacc aattctcag
 43081 tgctcagctc tctttagt ccaactctca catccatag tgaccactgg aaaaaccata
 43141 gcctcgacta gatggaactt tggggcaaa gtaatgtct tgctttgaa tatgtgtct
 43201 aggttggta taactttct tccaaggagc aagcgtctt taattcatg gctgcagta
 43261 ccatctcag tgattttg agcccaagaa aataaagtct gtcactgtt cactgttcc
 43321 cccgtctatt taacggaggg aaatttcca gagccccag gtccaggct gggccccacc
 43381 cactcccat gtcccagaga gcctgtcct cccaggctcc cggtggcgc tggtaagtc
 43441 caggatatag tcttaccat aagtgtctgt gtgtctagg aaagaaactc tccctctctg
 43501 tgcctctgt cctcatccg cagaagtac tgccaggctg gggagtctgt gacgtctca
 43561 gaagccggag gattttctcc ccattgtct aaagagagct cgggggtggg gaagctctg
 43621 caccctagg atcaccagag gagccagggt ctccagggt cccggggacc cctcagtggt
 43681 ggctcaggaa ccacagagcc agacctgat tcaaaaaacc tggtcacacc tccagatgac
 43741 ccttgtccc ttggctcgc ctcaaatgt ccaagccca acagtgaagc gcttaagaga
 43801 aggatccacc aggttgagt ttggggagga ggaagtggg gagctgggg agggcctggg
 43861 cctgggagac aggaatccac catggttca ggcagggtct ctggggcctg cgggtggag
 43921 agcgggcagg agcagacaga ggtgactgga cagacacac cctccactc caaggaggt
 43981 gggcaggggc ggggcacaga ggaacaagag acctgagaa ggggtccacc gagcagactg
 44041 ctggaccag acatctctga gccagctgga atccagctct aagccatgt cagccaggc
 44101 aggttatagg gcaggactga gtggagtggc cagagctgca gctcatggg ctgggaaggc
 44161 cctgcccgtc cctgagggt ccccagggt ctgcccagac tccaatttc gaccgagca
 44221 cacacaggag gaagtgtcg ggggtgagt ggcccagagg tctgggcagg tgcagggtg
 44281 gggaaggggg gcagctggag taccgctg aattcaggga cagtccttt ttctcctga
 44341 aacctggggc tgtccgggg gccaccgag cctccaggca gcggggggac ccagcccca
 44401 atatgtgaga agagcaggct ccaggctgga gagagcgaag caccatggtg gggagaagt
 44461 agactggatc ggggccccta ggggtcccc cgacctgca cggcagcct cagggcacc
 44521 gcacccatt gctgtcagt gctggccagt gtccaaggcc agggatgtgt gtgtgtgt
 44581 gtgcgtcgt gcgtgcgtgt gtgtgtcgt gtgtgcgtgt gcgtgcgtgt gtgtgtgt
 44641 gcgtgcgtgt gcgtgcgtgt acgtgtcgt gcgtgcgtgt gtgcgtcgt gtgtgtcgt
 44701 acgcgcgag cccagcctca gactggacc aggcagcctg ggattcctcc aaaactgcct
 44761 tgtgatttg gtcaaacgt gaggtctga taccgcat ccatcgccc cctctgccc
 44821 cctcatcac cgtgtgtgt gtcattatc agagctgtg aggtctggg aggtcatcc
 44881 acctgccagc taaacctga ggctgccga atgcactga tgcgggcaga cccgagacgc
 44941 tgtccggag acgaaggcca gctgtcacc ccgcccagag gcagtcggg ccacaagcat
 45001 catcaagca gtggttctt gagccgacg ggtgatgca aaggagccag gagacacgt

45061 cgcgtccaag ctgggggacc ccaggctgt tatgccggac agtaaacacg ttcagctccg
 45121 gagggagagg gtccccctac cttccagggt ttctattcc acaaacatcc aaagacaatc
 45181 cataccgaag gcgatccgtg ctttgctcc tgagacgtgc ggaagcacag agatccacag
 45241 aactgtctc ccaggatcct atgtatgtaa aggaaccgaa gtcccaggct gtgtgtctgg
 45301 taccacatcc caggaacag gctggactga tttcaccaa atgtagcaga aacgttaagg
 45361 agtatcagct tcaaaatatg agggccagac atgtctgaga agtcccttcc agaaaagtcc
 45421 cttggggtc cttcccaga gttgctgaaa cagagaaccg gaagggtgc agagctgaac
 45481 ttaacaact ggatcgcaaa ggtccgtctc atcagagcga tggttttcc agtggctatg
 45541 tatggatgag agagttggac cataaagaaa gctgagcgcc gaagaatcga tgctttgaa
 45601 ctctgggtgt ggagaagact cttgagagtc cttggactg caaggagatc caaccagtca
 45661 atcctaaagg aaatcaatcc tgaatattca tgggaaggac tgatgtgaa gctgaaactc
 45721 caatacttg gccacttgat gcaagaact gactactgg aaaaacctg atgtgggaa
 45781 aggtgaagg caggaggaga aggggtcgac agaggatgag atgttgggt ggcatcacc
 45841 acctgggac tcaatggaca tgggttgag taaactctgg gagggtgga tggacagaga
 45901 atcctggcat gctgcggtcc atggggtcat agagagtcag acacaactga gcgactgaca
 45961 gaactgaagc aactggcaag ccggagggtg ggtgccggt gcgatgagcg ggaacgtgca
 46021 acctgccacg tggagctct cctacacca gactcctgac ggcactggga ccctagccct
 46081 ccacggcctc tccagggcca cgagacaccc tcacagagca gagaagcgga acagagctgg
 46141 tgtgcagaac caggccccgg ggggtgggag gggctggtgg gcaggctta gtgagaagcc
 46201 cttgagccct ggaaccagag cagagcagaa cagttggcag agggccccct gggagaggcc
 46261 ccccgcccag agtaccggcc ctgggcccct ggggagaggg cgggtctggg ggcagggaca
 46321 gaaggcccag gcagaggatg ggccccgtgg gacggggcgc accaaaacag cccctgccag
 46381 caagggaag ctggggcact ttcaccccc tcaaggagg agcccacacc agcgcactg
 46441 cccaagggtc cttggccct gggggcact gaggcccagg ccaggccagg gggcccatga
 46501 ggccccagg ggtcagtga gtgtcccag gcagcccctg cctctcatcc tctgtggcct
 46561 ggctcttat cccgtggcg cccacggcct gctgccccg acagcggcgc ctcagagcac
 46621 agccccgc atggaagccc cgtcaggaaa gagcccttg agcctgcagg acaggtaagg
 46681 gccgaggag tcatggtga gggaagtgg gcttccctc gatgggacc aggggtgaat
 46741 gaccgaggg gcggggaac agaaggaaa ccagctggag agaaggagcc tgggcagacg
 46801 tggctgcag cacagcgtg acctgggccc cagtgtgcct ttgttggg tttatttt
 46861 aattttgtat tgagatgcta ttatctct ggagctttg ccgccctgag attttgacc
 46921 cgtggctgt gtccctctg cctcacccc gcctctgtag cagggcagac acggcgcaac
 46981 ggggcagggc gtgccagga ggcactgtca tttgggggc agcgccccca caaggcagg
 47041 ctgccttct cccctctac aggcagcgac agaggccag agaggtagg caagctgcc
 47101 aatgtcacac agcacaggg cgcagtcca ggcactgaga aatccggga ctagacaggc
 47161 accagagtgt cctgtgttt taaaaaacg gcccaagaga agaggcaagt ctgcaaggcg
 47221 tcccgggaag gcagcagggg cttggctcgg tctcccccga ggaggccagc tctcagcga
 47281 ggttctaag tgttaacgg agccaagcct gaaccaagg ggtcacgtc agctatggga
 47341 cactgacctg gcatggggga gctccaggca aaggagtag ggaggccaag gaggagagag
 47401 gggtgacag gcctgcagg agcttcaga gctgggaaa acgggttca gaccagggg
 47461 tcatgtccac cctcctta tctgggatc cggggcagg attgaggat ttatgtcgg
 47521 ggctgtcagg gtccagttc tctgttgaa aaattgttc agatcagaga ccagcgtgag
 47581 gtcaggtag aggatggaga agaagctgtg aaaaggtgat ggagagcgg gggacggtc
 47641 tcggtgatca ggcaccgaga tgcctatgg aatccgcagg cgaattaca gtgacgtcgt
 47701 cagagggtg tcggggagga acaggcactg tcatgaactg gctacaaaaa tctaaatgt
 47761 gcacccttt cggcaatat cagcaagtca taaaagaaa cgcatttct taaaattgcg
 47821 taattccgt ttaggaatt catctgggg cgggggaaca atcaaaaaga tgtgacaaa
 47881 ggttacaag ccaggagtc aactcgttaa tgatgggaga aaaccggaaa taacctgaat
 47941 atccaacaga aagggtgtga tgaagcgag catggcacat ccaccgcaag gaatcctaac
 48001 acaacttcc aaaacaatat ttctgactt gggttttta agcatgcgt cactttcaa
 48061 agctgtcag aaacataga aatatgcaa taatgtgtc ctagccaaat ttttaatt
 48121 ttgtttata attttataa gttataatt tatgaaatat aatgataaaa ttataaacta
 48181 taaaaaagt atgaaatgt tcacaagaag atatacatgt aattttatc tctacaatac
 48241 ttttaatac cagaataacg tcttttaa aaagattgag cacagaagcg tataagtaa
 48301 aaattgagag ttctgtca ccaaccacac gtcttactt aaaaccatt ctccagcgag
 48361 agacagtgc atgtgggtc gtacactct ggccttctc ctaggcatgt atgtccctga

48421 aaactcacac acacggctaa tggctgctggg atttagttt tcaaacgga ctcatctct
 48481 gcctatgagc ctgcaactat ttattcagtc tgtgagatt ttctatatca gccacatgg
 48541 atcccgcagc ttctctgaat ggctctgtat gaattcaaag ttggaagaa gcagcgtgtc
 48601 ttaatcatt cgcclattaa tggacgtttg ggggtgttcc actacaaaan nnnnnnnnnn
 48661 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 48721 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn atacaattcg agctcgtac cctggcttga
 48781 actatatgaa cagagaacga tgagaacagt ttctcaaact tggaacagt aacatttgg
 48841 gctaaatgat tctttttgt gtggagttgg cctatgaata gaggatatta gcagcatcat
 48901 ttaaccttta ctactacat acctgtagca actacatcct ctccatttgt gtcaatcaaa
 48961 actgtctccg gacatggaca agtgtgcccc tgggatgggt ggaatgacct ttgttaaga
 49021 accactgggt cagagattca tagattttg tctgttgac ttttaaaaa tacatctgg
 49081 ttttatttt attggttct gctctatct ttatgattac ctctcttta ctgggggtt
 49141 ccctgataga tttcccttc tggctcagct ggtaaagaat ctgctgcaa tgcaggagac
 49201 ctgggttcag tcctgggtt gggaggatcc cctggagagg agaagggcta cccacccag
 49261 tattctggcc tggaggattc catggagtgt atagtccatg gggctgcaga gtcggacatg
 49321 actgagtac ttacacac acatatgtcc ctgtagctc agctagtaa gaatcccacc
 49381 cgcaatgcag gagaccccg tccaattcct gggctcggaa gaticcttt tgttactcc
 49441 ataagatctt atctggggac aaaactaaca gctatgccag acctctgga catcaggga
 49501 cgtgaggggt gtggactgga cagatgtgtg tttctccca aacacaaaca tacatctga
 49561 tacatgtaca tggagagagg gggaggagg ctgtgagtct ccaggggacc gtgcaacct
 49621 gtgacattca tggaggcgt tgcgggtgat cactacacag ttctcttc tggttctg
 49681 gtcaattgac ttacaatt caattcctat acttattt agactgagg aattttacac
 49741 tatttaaga catatgtata catgagtat gttcagcgc atgagggtc atttgtgtg
 49801 tccacttgc ctggaacaa agttggactg atttactct aggggtgctt ggggtgttt
 49861 ctggaggaca ggagcattg aaccaagg gctcgtgaag catgagctc tctcaggtg
 49921 gaccaggag gaacgaagg ccgaggagg cagactctc tctccctaa cccaggtct
 49981 ctgctcaga aagggacaat ataagtact gaagaaaaga aagaacatca gctgtggag
 50041 gttgttctc tggagcagat tcacacgtg aggtcatgt gcaggaatt taggtgaaac
 50101 agagcagtea ccatgtgtg ttgaaaatt ttaattaca ttgcagta cgacttgtt
 50161 taagccagac agggtagcac agcaaagta ccatgtgtc acctgtgtt tgaaggag
 50221 agagaactg ctggcacatt caggaaagg cgtgtctcag cttggaggc acactgagag
 50281 gccacaagca gatgtgagg accagggtc cgggcagagg gatcaattca ctgcttca
 50341 ctttgcac atctgtgtc tgcctactt ggccagagta gttcagctt cagatgtg
 50401 agttccatt gtagaaatc caatctgggt cattttaaa cctctcttg ttctactaa
 50461 tggttttaa atctcttg ctcaagaaa aaaataaaca taattttaa ggggtgttg
 50521 gggcctgac tataaagta attatctgg ccatcaga gcatgttga attaatat
 50581 tctgtctta ctatgctc tatttctg attcttaca gtaatttt gtaggaatc
 50641 gggactgtg aatatttct tgtgaatac gggatcttg tattttcc taattttt
 50701 ttttttca ttttggtt tacctcagg aaagtaata ggactcagga aagtccttg
 50761 tccgctgtt attcagct ctacttgg gccagggcag cgttctctt gggctaagt
 50821 tcccacaac cggggcagc tctctact ctaccctg aggcctaat gaggagctc
 50881 cctgcgtctg agcagccgc cctctgtga cgtgcgtgtg tctctggca tcggcgtcg
 50941 gtgtcttg aggtccgc ctccctgc tactgtgc ccgactcga gctcaggc
 51001 tccaagcag gtccgagtg tgcagacct ctgttagct ctctctct caggactct
 51061 cctctagat gtgtgttt ttiggtcc ttggacctt gctctgaac caggcctgt
 51121 gctgagtgt atctctgg ggaagcctg gaggtgtgac gggccgccc tgcgtgtg
 51181 tgacaggtg gggctcggg cggggcctgc acgtctctt gaccgagcc gggactggc
 51241 tccgggctc aggcactact gactgaatc cctcacaga ggggtcagg cctgggcgg
 51301 ggaaccgtct ctgcaatgac agccctccc agggaggga cagcggggag ctccgaggc
 51361 tccagcccta gtggagggtc gggagacca gggagcggc ctgacggccc cacaccggc
 51421 cagggtgtg tctgtgtt tctcagctc aacagaagct ccgaggagct gggcagttt
 51481 ctgaattct cccggagtt ttgctgtga gtgtctgtc agcaccgtat ggacatccag
 51541 agtcattag cagtgtgtc tgtccctg tctgtctt atcaggctt ttgtcaggt
 51601 caccacacgg ccaacaccag gacagtctg tccgcccag ccatgtccc tgcggagcc
 51661 cctgtgcag ctccgaagg gccggaggc cgggggaacc gggcaggcc tgtccctgt
 51721 gtgtccacag tctcccggg gctggaggag agcgtgagca ggacgggagg gttgtgtt

51781 cacttccccg tctgtctgtg tcaactgtgag gattatcact gctgtcagct gactgacagt
 51841 aatagtcggc ctctctctcg gtctgggccc cgtgatggt cagcgtggct gttttgcctg
 51901 agctggagcc agagaaccgg tcagagatcc ctgagggccg ctactatct ttataaatga
 51961 ccctcacagg gccctggccc ggcttctgt ggtaccactg agtatattgt tcatccagca
 52021 ggtccccga gcaggtgatc ttggccgtct gtcccaaggc cactgacact gaagtcggct
 52081 gggtcagttc ataggagacc acggagccgg aagagaggag ggagagggga tgagaaagaa
 52141 ggaccccttc cccgggcac ccacctgag gcggtgcctg gagtgcactc tgggttcggg
 52201 gcaggcccca gccagggtc ctgtgtggcc ggagcctgcg ggcagggccg gggggccgca
 52261 cctgtgcaga gagtggagg gggcagcagg agaggggtcc aggccatggt ggatgcgcc
 52321 cgagctctgc ctctgagccc gcagcagcac tgggtctct gagaccctt attcccttc
 52381 agagctttgc aggggccagt gaggtttgg gtttatgcaa attaccccc gggggccct
 52441 cactgagagg cggggtcacc acaccatcag cctgtctgt cccagcttc ctctcggt
 52501 tctacgtct gcacatcaga ctgtctca gggactgagg tcaactgac ctccccgtc
 52561 tctgaccaca tgaccactgt cccaagcccc ccggcctgtg gtctccctg gactccccag
 52621 tggggcggtc agcctggcag cactctggcc gtggactgag gcatggtgt ctggggtca
 52681 ctgtgatgt gacctcaga ggtgtcact agtctgagg ggatggcctg tccagtctg
 52741 acttctgcc aagcgtgt ccttgacag ctgtggacc gcagggtgc tccccgaa
 52801 gctcccttg ggcagcccag cctctgacct gctgtctct gccacgtct gctgccccct
 52861 gctggtggag gacgatcagg gcagcggct cctccccga ggtaccccc agggccctgt
 52921 cagcagagag ggtgtggacc tgggagcca gcctgcctg gccagcact agaggccgccc
 52981 tgcaccggga agttgtgt ctgtgacct gtctcagggc ggagatgacc gcgccgtccc
 53041 ttgggtgt tagtgagtg gaggtccgg gatgactta gccgtaact gccaggctcc
 53101 gtagcaacct gtgcgatcc cccggggacc cagggtctct tgtgtgtg taccaagggt
 53161 ggcactagtc ccacccagg agggcactc gctgatgtg ttcctggcag ttgagtcat
 53221 ttgagaact acatcttt catcatcaca tctcatcag cagtatcacc accaccatca
 53281 ccattccatc atctcttc tcttttct ttatgtcacc tcacaatctc acaccctca
 53341 agagtttga ttgtagcat atttactta gcacagtgt cctctttta ggaaactggg
 53401 ggtctctgc tgataccct gggaacccat ccagaaattg tactgatggc tgaaccctg
 53461 cgttggatt ctgcccagg agaccctagg gcctcaaagt tctctgaatc actcccatag
 53521 ttaacaacac tcatgggcc ttttatact ttaatttga aaaatactct tgaagttagt
 53581 acctacctcc acatttaca gcaggtaaag ctgcttcgca ttgagagca agtccccaga
 53641 tcaataaaga gaatgggatg aaccaggat ggggccagg ggtctggat tcagactcca
 53701 gccgttagg acagaactg actaggtacg aagtgagcgg ggtggggggg caatctgggg
 53761 ggaactgtg cacccccagg gctcggggcc atccccacca cactctggt tcatcagta
 53821 gccccctcag cctgcgtgtg gaggaggcca gggaagctat ggtccaggtc atgtggaga
 53881 atatgtggg ctgggtgt gctgggtct aggggtctg ccaggtctg ctgctctgc
 53941 tgggcagtga taattgttc tcatctct gagaagtcac gagtgcagg tctctatgg
 54001 ccaagctatt ggaggaggca gtgagcact ccacctgc agacatctct ggaggcatca
 54061 gtgtctctg aggtgtct ggggcttgg cggggggacc tgagattcag ccattgactc
 54121 tcagaggggc cagctgtggg tgcagcgga gggctggcg gtggaggata cctcaccaga
 54181 gccaaaataa gagatcccc aacggataga aattgactca cacccttgg tctggcacat
 54241 tctgtctga aattctgt ggacaggaca cagtccctgg ataaagggt tctatcttg
 54301 cgtgtgcaat agagctgtc acacgttgg ctgggacatg taatccttg aacatggtat
 54361 taaattctgt tcaatacat ctgaaaggat ttgtcatca ataaacctaa ggtatattgc
 54421 cctgtcatt cctgtcttg tagtgtct gagtaggtg gaagggtga ccagcttcac
 54481 aaatcgagt aggaattcc ctattctc cactgtctaa tagacttca taagattagt
 54541 gtaattct cttaaatc ctgtataat cactactgt gccaccgga ctgaatttt
 54601 ttttaggat attttaaac aagcattta atgatttt ctttattt cggtgtgt
 54661 ggtctctgt gctgtgtcc ggcgttct cgtgtggcc agtggggcg ctgtctcgc
 54721 gttgcgaagc tgggcttct gactgcagt gcttctctg ttgcagagc cgggtccag
 54781 ggcgtcagg ctgcgtggc tgcggcact gggctcagta gtctggggc acaggtgcag
 54841 cagctctca ggacgtttg tcccagatg gtgggtcgt cgaaccggtg tccctgcgt
 54901 tgcaagggtg attctcacc gctggaccac cagcgactt cctggaggt ttttaattat
 54961 ggatttaagc tctattaga tctctctca catttctat tcttttga gtcagttga
 55021 tactttgt gtgtgtgaa gttgtccat ttatccaag tcatctaat tttgataga
 55081 caattattg ttatctct aattgtgt ttacaattt gagagcatt tctgtcaatt

55141 ccttctatct gcaagattgg taataatc tcccaagagg agtcacaaac tgaaatgaga
 55201 ttanatacag gctttttt taaaagaatg aacttatgtt gttgccttc tcatagatct
 55261 tacttcttag catgactgta ctactgact ggggcgttt catgtctgtg tggagagcta
 55321 ccattagtae ttctatcgc ccaagacat cgggctcctg ggcacagtga aaacactcct
 55381 ttctgtggct atttgcaaa atatggccta gcctagcgtc ataagggtc acagctgaca
 55441 actgctggaa cagagggaca tgcgaagcaa cgtgagggct ggaacctgga gggctcctc
 55501 tggggacagt ttaaccagct ataattggaca ttccagcctc tgggacatgg agctgtgaac
 55561 tggaccaatg actgtcatt ttggaagaga aatcccagga gagaagggtc caggggaatc
 55621 tgaggccgca tgcagtgcct caggacaggg gacaccttct ccagcagagc aggggggccc
 55681 gcccaggccg cctgcagtga ttccaccagg aggagatgca tccctgcaga cctctgacag
 55741 cagggccctc tctgagaca cagggtcaca cccggggccc tggaacctt tgagacccta
 55801 aaccttctc ttctgacca cctgacagc agtctagctc agaacagaca tcttctttt
 55861 cagcaggaaa atcctttcc tegttagg gagcactgg caccggagga gctgagtctt
 55921 ttaaacacag gctgcctgaa cctcaggat gacctgcagc tgcagagg aggctggagt
 55981 gtgatagctc acttaattg tactaaaagg aacatattgg acacccctc tctgaaaaat
 56041 ttccctctg cctctatct cttagtcac ttatcgccg ttctactgt tttctatita
 56101 ctactctaa cgccaacctc tcttattcc cctccagtt taacacggtt tccctccac
 56161 ccgtctctt taatctcaga agattctgcc tttctctc ttatcacag cccctactt
 56221 ttattttt tcttaccgc ctttattcc cctccctc cactctct ttaattacat
 56281 cttactaca ccgctgcgc tatctcgaa tctatccaa tttttccc ttatataca
 56341 ctccaggccg agcggctaac ttattataat ttcttatag cgcctacctc atttccctt
 56401 atttctaat atctatat acccatgcaa ttccnnnnn nnnnnnnnn nnnnnnnnn
 56461 nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn
 56521 nnnnnnnnn nnnntgggt gtacgtata gagtaaacgc gcatgaagaa tgggtcaat
 56581 ctatggctgt gagaggcaga aaataatatt atcatatata atttatgta taacacactg
 56641 aggtgggtgg ctcgtagaat agtgcggagc gggagaaagg tgggaaggag aagacacaag
 56701 agagagatgt tgcctcgcg ggtatggatg gcggaggat agaagaataa aaagaggaga
 56761 ggtatagagg ggggcggggg gcataacgtg tgggtgggta aatagtaggc ggtattatg
 56821 aaaaaagaa agacgggggg ggcggtaaca tagaatcgc aaaaagtca tatactgaac
 56881 ggggattagg gagaagaggt ggggggcgtg ggtgcgggg gaaagaggtg tgtgtataat
 56941 tggatggag tgtattga atatatata atgtaatagg gagtgaatt agtgaaattg
 57001 tgggagtatt atattgggt gtgggggaca tggcaaagt atgatcgga taaaaaagt
 57061 aaagcaagag gggaggggaa aataaggggg gggagaagg cgaagaaat aagaggaaga
 57121 agaaagaac ggggtggcgg gcgggggggg cggcgtctt gtatctggt tttgtgtg
 57181 gtcggtggt gttcgcgtc tgtgggtcc ggggcgggtg tgcggaaaa aaaaaggcg
 57241 ggaggcccgg gggcggta cgcggcacc ccgcggtcc ctggtctc cttcggcagc
 57301 tccgggggtc ggtgagcctg cgcctccgg gccgcggcc cgagctgtgt gcgcctgga
 57361 gaatcggagc cgtgtggca gcacgcggag ggcgcgcga agggccacgg gacggactt
 57421 caaaggccgc ggcggagcgc ggcaagcca accgagggcg gtcggcgat cggccgagcc
 57481 ctgctcccc cctccgcgtg gcccagggt cgcgggtgga ctggggcggg taaaagcac
 57541 tcacccctg cccgcccc gaaagcctc caggactctc acagagcacc cggcaggagg
 57601 catcgggtc cccctcggc tcaattcagt tctcagtc tgtccaactc ttgcgaccc
 57661 catggactgc agcaccaca gcttccctgt ccatcaca cctccggagt ttactcaac
 57721 tcatctattg agtcagtgt gccatcaac cgtctatcc tctgtgtcc ccttctctc
 57781 ccacttcaa tcttcccag catcagggtc tttctatg agccagttct tcatcagg
 57841 tggtagagt attggagtt cagctcagc atcagctct ccaatgaaca cttaggactg
 57901 atttcttta ggtggactg gctggatgca gcgcagaca ccgaccgct ttacccctg
 57961 tgcctttcc aatggctgt cctgcgggc ctaggggcat tggtcgggt ttgaactctg
 58021 tggcctttaa ttatgcct tagttccagg tccagggcag ggccatccgg attcaggatg
 58081 cttccagcc cttcaggaat ggcaggttt catgtctt tctgagtga ttctgagtg
 58141 tcatattgt gcccttgca gggagggtc ctgacttcc tatcttaca tcaactgtcc
 58201 caaccccaa gagaggctc ttggcccagg gactgcagg aggatgaagt caggagcaga
 58261 agcatgggt agggggctc ggtggcaga ggaggccct ctgtaggag gaacggcaag
 58321 cgaggaggga acaggggac cgcagtgcc tggcaagct ggtgatgca cgactacgc
 58381 ccgaccacac agtctctca gccagcccga gaagcagggc cctccctga ccccatctg
 58441 ggctgggtc tcaatttct cctccctgca atgggtgac tgttgcctc caggagagg

58501 gagcatgtaa aggtggccac tctctctgg cagacatgcc aggcctgggc cagcctccac
 58561 cccttgctc ctgcagcccc tctgacctg ctctgttg ccacaccggc ccctctggg
 58621 ctgatcagg ccccccctc gcaggaagcc ctctgggaca agcccagctt gctgtaactg
 58681 tggcttcca ctgtacctg caacgtggga ggctgttact taaaactccc atgactgtg
 58741 gattgccggg cccagaaca aggcacgca tccctggagg cctcagagac catttaagg
 58801 agttaacat ttctactta tgcatttca tgtgtatcag aaagaaaaa aatgtatcat
 58861 cagttcatca aatccatgat ttctgacca atattgctaa gatgaggctg aaataggcat
 58921 ttccatttt aaaaaactga atcactctga agaaacagat ggcaggcttc cctgggtggtc
 58981 cgggtggtta cagtcctgc ttccagtgtt gggggcatgg gttcgtccc tgaaaattt
 59041 aaaaaggaag aaaaagatgg ctccccctc cctgggattc tccaggcaag aacactggag
 59101 tgggttgcca ttctctctc cagtgcata aggggaaaag ggaaagtga gtcgtcagt
 59161 cgtgtgcgac tcttagcaac ccatggact gcagcctacc agactctcc gtccatggga
 59221 ttccaggc aagagtactg gattgggtg ccattgcctt ctccaggcaa acggcctgt
 59281 actgtactg ctgctaaatc gcttcagtcg tgtccaactc tgtgcgacc catagacggc
 59341 agcccaccag gctccccct cctgggatt ctccaggcaa gaacactgga gtgggtggtc
 59401 attgcctca gctgtgtgt gctgtgcta agtcgttca gtcgtgtccg actctgtgtg
 59461 accgcataga cggcagccca ccaggctccc cgtccctgg gattctccag gcaagaacac
 59521 tggagtgggt tgcatttcc ttccaatg catgaaagt aaaagttaa gtgaaattgc
 59581 tcagtcgtgt ccgacttta gtgaccaat ggactgcagc ctaccagggt cctccatcca
 59641 tgggatttc caggcaagag tactggagt ggggtccatt cggcctaggg agtgagaaat
 59701 caggtgtgt ttccctctc tgcctctca ggggtctctg tggagcctcc ctggagaggc
 59761 cgcggcggct cgggggactg gagggggagg ggggtgtgag tcagccgggt gccctccct
 59821 cgtgcccgt ctctccctt tttaggcaca agctgggcgc ctttttagg cgcagcctca
 59881 cctgcccgc cactgccgt gttcggctc cccggagata aaacagattg cctgcacccc
 59941 gggcatcac aaggattga tgaccgttc ccagtgtgt caccacctc cctctgattc
 60001 tcagagacgc gccctgcct caggaggctg ctatcccag gccaggggc ggcgtgggtg
 60061 cccagcgcc ccgacagac actgcctct gaccacctc tccaacagc ttacctgcca
 60121 agaaggctc ctgacccctc atctgcccgt gtggttga gaaagctca tctggccct
 60181 cctctcggg gctcagtt cccctctgt gaactggcg attctgcaa gctgacgtc
 60241 tggccagccg cctccccgt gccagtgtc cccgggacac agctgaatgt cctgtctgg
 60301 gatgcacct cccaagtgg cctgtcagga ggcggggcg agcagggaaa cccgactct
 60361 ctgagacggc ccatgcatt ggggacgtg agggccggag cagcggcacc ctctggcca
 60421 gggctattt cccgccccg cccgtccct cgggcctcc agaccgcagc ccggcccgcc
 60481 ccgggaagga ccgatccgc gggccgggc acccccttc cctggccgc ggcgcggggc
 60541 gattgcagaa caaaagcgg ggcggggcc gggcggggg cggggcgag gatataagg
 60601 gcggcgccg gcggcacc agcagccct gcaccccg ggggatggc tcggccgccc
 60661 ggctccgc gggcgccct gcgcgctt ttgttttg tgggggtg gggggcggtc
 60721 gcgggtact attttcat ttataattg gtattagta gcgagtga ccacacctt
 60781 attcactat agccaattt tgcgggggca tttaacta cagactgcc cgcctttat
 60841 ttggtagc catatcagat cgtctctta ctgacact agtgattatt gtctatagta
 60901 caaaaaaga acggttgtt cggcgtaat gtgcatctt cctctctgt ttctctgac
 60961 cactcaatt acaccaaac tctactatt aatcacgta ttgtacgcca cctccgccc
 61021 gcgaactaaa agaattgca gatattctga agataaaatc gttcattgtt acgccccgc
 61081 cgttcgct atattact tagaattt tttcggcg agcagttatt cccccccg
 61141 aactagatg cgccttaata ttgttctaa ccgtttgga tttaacgat aggcgggaaa
 61201 ggtagacatt cgaccgtac gacaactaaa atcgacgag acaggctatt tatatcgga
 61261 ccacacgag cgggtataca naccgtaaaa ttatctaa tcgagagtaa ggcacagag
 61321 cgaaatacaa cggcgtgtt gggaggtgt tctgtagtga attcgacct cgcgccgccc
 61381 cctctgtgc tgnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn
 61441 nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn
 61501 tattaataa cagcgatag atgtgtgta gggaggagt gcataagaga taaagagag
 61561 gcggcgag agaaatag tagaggagga tgagagaaa aagaaagca gcgtaggtac
 61621 aacggcggg ggtagtatg ataaagtga ttatatatt tgagtaaagg aaggtagat
 61681 ggagtataa gaagtaagga gaggagagg cggcgagag agagagtga aagaaataa
 61741 tggggcaaag gcgggtggg tgagaagcag tagaagaga gatagaga ggggaaaag
 61801 aggaaatga ggattaga acagtaggaga ggatagatg gaaatag atcaggtcaa

61861 ggtggagaaa aagtagaaac tggggcgtga ttgtaaaaaa gggaggccgc gatggggcag
 61921 caccataagc gaagagatga attaatgaaa gcaaggcagg gagaatcaaa tgagtgggt
 61981 ggaggaagga ggctgtgact tccttcgctg ccggaagag aactagaata gcctcgggt
 62041 gtggggggag gtaagataa agtgacttct gggccctggg ggaggcccag gagtttctac
 62101 cgagctgagc tgggtgcctc tcccaaatgc ccaacccct gagagtcgac gggagagcac
 62161 agcctggcca aacctgggca gggcacacgt gtccttcacc ccacagtgt cagagccca
 62221 gcgtgggtccc tgcgtctggc gggaacaca gacctcaca cccacacaa gggtcgggc
 62281 gctttcaaat aacagcagcc gtgcccctg ggccggtgac ccggacacag agagatgaag
 62341 tccgcatctc tcagagtgcg ctgtctccg cccggtcagg cccgggtccc ctgttctct
 62401 gagtcacca ggagggttg catgtgggtc tcaggacac aggttcagt atgtacaga
 62461 gggtagtggg tcccagcagg gccggtctt ggaccggtt ttctgaaaag ccagttggcg
 62521 acctgggtc acagcaaagc tgatctgtt tggccaggag tctccagt acggcctccc
 62581 ccagaacatc gggcccagt ggggtccag ggggtagact tgcctccag ctacgcccg
 62641 tgtctgaca agtccatgat ttgtaaaat taattgtgt tggatggagt tgattagt
 62701 gtgtgtgagt ttctgtggc cagcaaagc aatcagttac gcatacacat gtatccagt
 62761 cttctacga ttctgtccc atataggta ttatgggtg tcaggtagag cttctgtgc
 62821 tacgcagta ggccttatt agtcagctc agtcgtgtc gactcctgt gacccatgg
 62881 actgcagcacc gccaggctcc cctgtccat accaactcct ggagcttatt caaactcatg
 62941 tccatcgagc cggatgacc atccaacct ctcactctc gtgttccc ctcctctgc
 63001 cttcagctt tcccagcacc cctagagaa gggaatggca aaccactcg gtattctgc
 63061 cctgagaacc ccatgaacag tacggaaagt ccttattagt ttctattt atatagca
 63121 gtgcacacgt gtcagccca atctcgcaat ttacacccc cctcgcgcg cgattgtag
 63181 tcatgttgt ttctacatc tgcgactta ttctgttt gtaacaagt tcattacac
 63241 cacttttta gattctgcac atacgtggca agcccacagc aaacatgctc aatggtgaaa
 63301 gactgaaagc atttctcta agatcaaaaa caagacgagg atgtccactc actccgttt
 63361 tactcaacac agccctgaac gtcctagcca tggcaatcag agaagagaaa gaaattaagg
 63421 aatccaaatt ggaaaagaag aagtaaaact cactcttgc aaatgacatg acactatac
 63481 ccagaaaatc ctagagatgc taccagataa ctattagagc tcacagtga attgttga
 63541 ggatacaaaa ttaatacaca gaaatcctc gattcctat agactgaca caaaagatc
 63601 gagagagaaa ttaaggaaac catcccacgg catgaaaaag agtaaaatac ctaggataa
 63661 agctacctaa agaggcaaaa gacctgtact cagaaaacta taaaatactg acaaggaaa
 63721 tcagacgaca cagagagaga gagataccac gctcttgat gagaagaatc gatagtgtga
 63781 caatgactat actaccaga gaaacataca gattcagta aaccctatc aaattccaa
 63841 tggcatttt cacagaatca gaattagaac aaaaagttt acaagttca gggaaacaag
 63901 aaagatcta aagagccaga gcaatctga gaaagaaaaa tggagctgga agagtcaggc
 63961 tccctgagtt ctgactgtgt atacaagct ggcagattt ttaacagcag ggggtgaaat
 64021 gaactgttc acaaaacaga tgggtgggtg ggcttccctg gtggctcagc tggtaaagaa
 64081 tcctctgca acgcaggaga cctgggttcg atccctaggc tgggaagatc cctggagaa
 64141 gggaaaggct acccactcca gtattctggc ctggaaaatt ccaaggacca tatagtccat
 64201 gggtttcaa agagtcggac acgactgagc gattccaat cctgaaaacg tccattgtg
 64261 gacggtgaac tgggtgtgc caagtcagg gtaaccgtt gctgagtac tgacactc
 64321 tctcatgggt taaatgtgg ggccaaggc caggaccaga cccgcagtc agccaggcag
 64381 accctgtgca gcccagcga gtgtgtggc gccgtggagt tctggccc catggcctc
 64441 gactggagcc cctggagtga gcccattccc tcccagccc tgagaggctg ggtgcagccc
 64501 taaccattc ccaccagt acagatcgc ctgtgtgga acctgtctt gtcccagg
 64561 aacctggcag gactcaggga gaatgttca gggcgccac agatcagggg ctggggggg
 64621 agggctgggt ccagcagagg cctgtgccc actccccga aagagcagct gatgtcagc
 64681 atgaccacc agggcaccga cgcgtgttg cacacaggcc gcccctcat ggtgacactc
 64741 tttctgtg gccacatc gcccctcag gtcctctg ctcccagct cctggcctg
 64801 gaacctctc cccgcccgg ggacgtcagg gctgtgtcc actgagcatc ccatgcccgg
 64861 gactgtgtg atcaccagca cctgcacccc ctctcgggtc tcaccaggat gggcaactcc
 64921 tgccatcca gcaccagcc tctgggtac acatcggggg aggagggaga agcctgggccc
 64981 agaccccag tgggtcctt aaggaggaca gaaaggctgc cgtgggcccag ccgagagcag
 65041 ctcttgaga gacgtgggac ccagaccac ctgtagcca cccgagtg ctctgtcac
 65101 acgggccacc agcccagc tagtgtggac gaggggtagt gggtaggccc caggtgcacc
 65161 agggcaagt ggtgaggccc ggtgagcag ggtgagtggt tgaggcccag gtagaccagg

65221 gccatgtgg gtgaggcccg ggtggaccag agtgagcggg tgaggccag gtggacagg
 65281 cgagcgggtg aggccaggt ggacaggcg agcgggtgag gcccgggtg acaggcgag
 65341 cgggtgagg ccgggtggac agggcgagcg ggtgaggccc ggtggacag ggcgagtgg
 65401 tgaggcccg gtggaccag gcgagtgggt gaggcccggt tggacaggcg gagtgggtga
 65461 ggcccgggtg gaccaggcg agtgggtgag gccaggtg acagggtgag tgggtgaggc
 65521 ccaggtagac cagggccag agcaaagccc cggtcagca gtgattct gagcggccac
 65581 tgctgcagg gacctcagc atggaaggc agccctgtt ggggtccc actggggaca
 65641 gcatgcagag agcgagtgt cccctggaga aacagccagg gcatggccg gcgccctgc
 65701 aggtgcccc aggggccaca gctgagcccc gaggcgcca ggggcccga cagccctgat
 65761 tctgggttg gggctgggg ccagagtgc ctctgtcag ctgggcccgt gacagtggc
 65821 cctcgtccc tggggcccg ggaggacgg tcaggtgaa aatggacgt tgcgggtctc
 65881 tgggttgac agtgtgcct attggcact ggctgttgg gccagcagc ctaggccag
 65941 cccccggg gctcccaag gggccgac ctcaccca cgagctggc ctggcgaac
 66001 caagaggccc tgacggcca aatagccagg aaacccgac cgaccgcca gccctggcag
 66061 caggtgcct cctccccg ggtgggggg aggggtgtt ccagtctg agcttccac
 66121 cagccagct ggagaaagg ccacatcca gacccaggc cgccaggcc cctgtgtca
 66181 ggcctggcg cctgagacca cgtccgtcag aagcgcatc tctatcca cgtcctgt
 66241 tctgggatc tggaggtcat ggccctctc ggggcccag gagccatct aagtccagg
 66301 ctgagctg aggtgccc gggacacaga ggagctggg ctggcctagg gcaccggt
 66361 cacactccc ctccgccc tcactgga ctcttgcgg ggaggactg agccaagt
 66421 ggggatggg agaaaaatg ggacctcac gatcactgc ctgggagccc tgggtcgt
 66481 ggagtaaca tgggtgact cgaagcacag ctgtccca cgaggctca cagggtctt
 66541 ctccaggga cgggacctc gatggcagt cactcatca tccccacg ggcctcac
 66601 ggtcctctc caggggacg gacctcagat ggccagtcac tcatcatt ccatgagg
 66661 ctacagggt cctctccag gggacgggac ctcatgagc cagtactca tccattccc
 66721 acagggcct acagggtct tctccaggg acgggacccc agatgggcca gtcactac
 66781 catcgtctg tgacccatc cgtccaaca taccctcc ctcacatc ctgaaagct
 66841 cctgaggcc tcccgggga ccagcctgc atgggccc cagctgtca tccaggcca
 66901 gtcaggccc gcacagtaa ggcaaaagc agacctgaa ggtgctgt taccacggg
 66961 agggggggg ctgtggacac aggggcccc atgcccgtc cagcctgccc cctgtctg
 67021 gccgagatg tgagggaac gggggggcag gaggtggac agacaggcca gcgtggggg
 67081 ccagctgccc cctggtgag ggtgagcaga ctgccccct cccccaggt acaggtctc
 67141 ctgatgtccc ctgcccctc tgcctcctg tccggctca atcagagagg tcccgcat
 67201 ccagggtcc gtggtctca tgggaataa aggtgggga caagtaccg gcacgtctc
 67261 ctgagccac cccaaacac acacaaaaa atccctcc cgtgggact taccagctc
 67321 gtctcaggg gagctgccag ggggtcccc agcccagga agccagggc caggctgca
 67381 agtccacag cataacaca tctcagctga cacagagaga cagtgtctg tggacagg
 67441 cccccactg cgagcctga gagtgtggc ctgcctgcc ccagccggg tcatcggt
 67501 cagcaaccg tgcactcc cagcgccctg gctccccct tgggcccag tcaagtctg
 67561 ggggtgaagc taagtacgg agcctcatc atgcccagc cggagccac agcgccatc
 67621 agaatgct cttccctca tcaggaaaca ttatgggaa agacaagagc tggggggtc
 67681 tgggtctc ggggatcaga tgaagggtc tgggagcagc agcagctca ggcacccaa
 67741 aacaaggccc aggagctga ccccagggc tgaggggcag aggggaaggaa ggcctctg
 67801 ggggttgca tgagcaagg caccagggt ggggtgagc accctcggc tggcacac
 67861 agggccccc tgcatcct tccccctg agacctggg ctccgtctc ccgctggc
 67921 tgcatcctg ctaccacc agaaatcct gagtgggtg ccatgtact gggcctg
 67981 ctggggagg aggagattc gacagacagg atgcccaggc agagaggggc gagcagagg
 68041 tgctggagg gggccgggg aggcctggg ggcaggggg caggagtct ccagggtga
 68101 cggcgtgtg ctatgctcg tgagcacaga gggccgggt gtcccaggc tgggaacca
 68161 gcagagggc agggacggg ctcaaaggc ccaaaggcc agcctgacc agactgtg
 68221 gtccagaagg cagctgcgc ctgaggccac tgagtggccc cgtgtccga accaccgtg
 68281 aaacatgga cacagttc caggcggag cactcctgc tccgggagg ctccagcgg
 68341 gctcatgct ccatccaca gggaggaaa ccaggccca gatgacga atcccggcga
 68401 gcaggtcaa gccagcccct ggggtcccct ctccggcct gggcctccc ctctcagg
 68461 tgggaaacc agggcacaca ggggtccat ggggtgccc tctccaggc cctggacac
 68521 ccgcggtga ccccgctc tatcatcca gccctgccc gccctggca cccgtggat

68581 gacccccgcc tctatcatcc cagccctggg ggacagatgg gaggcccaag cgtggacccc
 68641 ctggccaccc cctacccac agccgggagg agccgggagc tggggccaa gggcctagag
 68701 gagccagann nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 68761 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnca atatagagg
 68821 ggtgggataa agggtaatat gatgttagg tagttagagt taaattagaa gggtttggat
 68881 aaagattaat aaaattacaa gcgtacatat cgtgtgagtg tgggtgataa ttttgtgta
 68941 tgtggggaat agaagtgagt gtgagtagta ttcaagatgt aagtgtgcca atacaggtct
 69001 gagcgatttg aatggaagtg aaaaaagcg tgtgtgtgga ggaggcggga gaggaagata
 69061 gtgtggggga agaaaagaag gctagtgggt aaagaaatat cagtaggcgg ttgacgaaag
 69121 aagaactagg aagaattaat ataaaaataa agggaggatt aaaaaataa gagggaggag
 69181 gtaacggaaa tagttagta agaaaagaat ggagagtga ggtaagataa ataaggaggt
 69241 aatgggagtg aggaggaata aataaaaaaa tggtaggga aaatagagta gaatgagaac
 69301 aagaatgaaa aagggagtga agggggtgaa aaaaagtga gttgaaaaa gaggaaaaaa
 69361 aaggagaaga taaaaaata aaataaaaaa aggaaaaaa agaaaaaag aaagaagggt
 69421 taaaggacga aaagaaggga agagaaaaa aatagttta gtggggagg gtaaaaaaga
 69481 attaataaag taaatatggt tgggtcgaa aaaaaaaa aaatgtgtg gttgatgaga
 69541 agaaaagaaa aaagaagaaa gggaagaa gaaagaaagg agagaaaag acaacccac
 69601 cgccgggag catggagggt gaggtggcg cagcccgcg gatggcacag catcacagca
 69661 atcctaaac gtttcagac cgtgcatct tcaccgcgcg cgcgccccg ccggccctcc
 69721 tcccgcctg accgcggacc cccaccgca ccggggagcc taccaccac ccggggagcg
 69781 tccgccacgc taaggtagg actgacctga agacgcgcg gggtagaac gttttatct
 69841 catgacataa gcgagtgggt ttgaaacagg ttacaaacc ctctgaaga cgcacctta
 69901 gcgttaggt ttgttttt accatgtgac gatgcaacta ttttctct ctctccaca
 69961 gtggctagtc gcctccagag cgagggtat ctctgtaca gagacctcg gaacatccgg
 70021 aggtagttc ccacctagg gtaagcgag aaggtcatt acgaggcg gggtctctg
 70081 gggaaggga gggccctggc gcagaggctc tgccacctca gtgacacgca gaccacgcg
 70141 ggcctgcagg cgccgggctc tgaagcagg caaagccga tctgtgaca tcagggttc
 70201 cgcagcagc aaggtctgc cgcacctgg cccactggca ggggtaagc tctgctccc
 70261 gacgacgca ccaagttcag gaaggccac gcagacactg gtgagacacg gccccccg
 70321 agctgcccga gaagctctga cttgcacta aagatctct gcgcgtcca aaaaatgaag
 70381 gcctctctc ctttatctt aagacttga tttttacg atgtaataa taccaagaag
 70441 ggctttaat ttacagaca ttaggataa tttccctgt agccctgct gctttgtta
 70501 gtaacgaaac tcaaccaga aataccaaag gaatttcca aagagttca aaagcgtta
 70561 tcagcaatca ctgactgct gcatacatca tctgcccc aaacaatagc ctgctgtgc
 70621 cagttactca aagtactact tactgacga aaacaaatct agtctaagc ttttataaa
 70681 gaaactccac tctccgca actttcaga acaaccact cgtacacgtg gcaggggacc
 70741 gtgctggac tgggtgctg ctcctctgt gaccaggca cactgcccc ttctggcct
 70801 ccctacgct ctgacaaat gttcatcagc tgaaggtc accccacgag ggaccactt
 70861 ctgctattc ccacgtact accccattat aggagtttc tttgtgacg ttctgcatt
 70921 ttcatggat ttgagggtt acataatcag ggctgctgaa cagcatgaga gacgtggca
 70981 caaggctcct cctgcactt gccgcagggg caggcgagt tatctggctt gagcgtggt
 71041 accatcagg ggtaaacaca gtttcagga cgttttgac aagacactga cccgatgcc
 71101 cccactacca ccgtgcagg cctgcaggcc tccagcctc ccaggccct cccgaggtcc
 71161 cttcggaact taggggactc ggtctgccc cctgggttt cctgcacca gctttgccc
 71221 cctctggacc caggtttccc aaatggaaaa cgaagggtg ggtatggaag ctccctgggc
 71281 tctctcagc tgtgctctg catggtgatg acggctgccc atcgggggg gcaggactg
 71341 ggcagctgc gacacctcc caaggctgt accccgagt ggtgtggggc gctgtgggca
 71401 cgctctgct agcgacctc ctggaacca gcgcctgccc tctgcccgg gcaaccggcc
 71461 cgggagccaa gcaccactgc cgtcagagga gctgctggt gtgagtggac gccagtctag
 71521 ctctgaacce tggcaggcc tctgaggct tgaacattgt aaaatcagg cccggacggc
 71581 aactgcctc cctcctgccc gctgtgtc cataaactgc atctcaggac aaatctctc
 71641 actcaccagg gctgaaacag aagactgcag ctatcttct caaatctaag gtgtgctaca
 71701 gggcaagtc cagaaactgt ctggcctaag catctcatca gatgcctgag acaagagctg
 71761 tggagccaa gctggagcca gagctcctc gttctgccc acctggcacc gcgtccacc
 71821 cagtaaacgc aggttgatt tcaaaagta ccaccgact agagccaatg ctaaaccgac
 71881 cactttct gccattaga ttgggtgaag gttctttaa tcaatctgc agtcaccaca

90

75301 ggatctcagt ccagctcggc cgttggccg tgacgtcca ggtcacagg cctgccggca
 75361 cagaggagca ggccttcag tgccgtcgag cactcggagc tgctgcctcc gctgagttca
 75421 ctcaagtgt acgcacagag cgcctcgt gtaccaggcc ctattccag tccccagtc
 75481 accgagcccc cagggtggt ggggacctgc cctcgggtac actgtgtccc gtcagtggc
 75541 ttacgtgtg tcttgaggg aggtggcat tgcggtccac ctctcagcac aaacatctgt
 75601 cccctgggaa ggggggtcca ttctgggtg cgagcagccc cctgggggtcc gtgtctctc
 75661 cttacctggc tcaaggcccc ggtcctggg tcttgagag caggagagccc accctcggg
 75721 gctgtggagg gggacctgc ttctggaggc cagccgagg gccagggcg cgctccggc
 75781 cgtgccctg agggagcagg ccgacgcca gcgcggctc tcttgaggc ccgggaaacc
 75841 ctgctgagg gtgcgggtg gcaggtgccc ctgccccag gctctctgt gtgagtaca
 75901 ctcaccagc agctctgat gccaccatc cgggttctc aggaggcact catagcgggt
 75961 ggggtccct cctcccccc tcttgaggg gagggagtct gatcactgg aggtgtgtg
 76021 tccgtaccg ccccccgac tctggactg ttactacc ccgctgggc tcaggacagg
 76081 gcattgatg ggaaggacag ggctgggtc tggccaggct gggggctct caggcatgg
 76141 gtgcccctg ctctcttat attcaactg cactgcagg gggcgcaat ctggacccc
 76201 acttactgat gatctgcatc aggacatagg tccccctc tgcagcggg ggctggccac
 76261 ggagggcgct ggggaaggc cctctccag cccctcggc aggtcacca ggtgcccac
 76321 ctcagccagc agggcgagc tgcctggag ggcggagagg gaggcagggc agggctggt
 76381 cgacccccg tggggcggg gggccctcag ccggtctcc agcaccctg ctgcccccc
 76441 tcaccgtcag ggggcacct gccgtctgc ctcagtggt cggtaggggt cccaaggcca
 76501 caccaggtg tcaccagtc ccagcagctg gcttgaggg aggggcagag gtggcgcat
 76561 ggcacccgc tccccccag accagatgc tctgcttcc tccgcccac cteccagac
 76621 atctgaagga ctctgcctc caccatgcag cccgcctcc accagaagct caggttccc
 76681 gccccctc ccgaagctg caggacctc gaccagcga gagatggac agttggaaca
 76741 cagctccc cagcagcgc acagcagct tgtggcccag aagagccgc ctgttccct
 76801 caagcaact ccatggatg tcatccatg gacacccct tccccacac gcctctcgt
 76861 tctccccct caaggcagag ggaacgacc ccacctgc tctaggaca ggggaccca
 76921 cttacctcg aacatcacct tgataaacat ggccgtggt gggacagatc cctccgacc
 76981 ccaacttcg acctgggga ggagctggg tgagctcga ctgcagggt gggccctgt
 77041 ggaggtgtac ggggtgagag ggtgatggg ggtggtgctc aagcggagct cctgtcag
 77101 tccagcggt cctgcagct agtcaggat cctcagcct cccccctc ctggatcag
 77161 gaagactgag gttccctcc ctgcccccc accagcttc caagctgtc tctgtggcag
 77221 tgggagctgc caagaggtc gagcgccag tatccgggt acggggttg tggagggtc
 77281 gggcattccc ggtcagggc tctagtggg gctggagct cgggccaga gctgtccaga
 77341 gaccagtgc ctcaccgc cgcccccgc aaggagagac agagctcca ggcggggagt
 77401 cggaggtcc tggagggga gcactctca ctctcagc ccccttcca ggcgactcc
 77461 cggcctccc gtctctgc cctgtctt gttgaagtat gattggcata cagttcacg
 77521 ccactctcg gagtgtct cactaagg atacagaaca tgcctcgt cccccaac
 77581 tcccagccag gctgtcaga agaggaggc ggccgacgg gcagggcct gactctgc
 77641 gtgtgggtc cacagggtc tccccgtg cgggtggccc tctctcag gccaggagg
 77701 tccccctg tggaggtgc gtggtcgc tgcgtcctg ctcttgggt tgttcccgc
 77761 atgggtgat gatgaagagg ccagtacaga cactcggcag caggtctct ggtgaacagg
 77821 cattattc tcttctga gggcagatc tggagtgg gtccggacc gtccgggag
 77881 agtatgtc tgttctaag aagctgccg gttctcag gtgtgcac atgtcaggc
 77941 cctctgtg gcttgactc aggagacct ctctcagc gccctcccc ccaggtgtc
 78001 aggcactct tgcctctg ggggcagagc tcagcggcg aggcgggagg aggccagat
 78061 cccagcag cccaccagc ttgtctgt tccctcggc tcatagct gagaaaggc
 78121 aaggagcacc ggtgaagc ccactggg gacgcactc gatggcaga ggtgtcaga
 78181 ggtggcccc ggcagcttc ccagacgca caggcagt ctcttccc aggcaccac
 78241 tgtgtctgg gaccgagc ctgcagcag gtcgggagc gctgtccca gattccggc
 78301 tgcacctg gctccagca cccacctg ttgtcaagg gttttgt tctgagccg
 78361 cgaggaggga gctttgtc tgcagtgtc agaagtgc ataaaggagg gccacagt
 78421 ggagcttat aacattgtg cggagggtg taacaggtc gggaggcact tgaggagcc
 78481 tctagggc atggagtgt tctaaattt ggtctgggt caggctacag agatgtgtg
 78541 gtgtgtgt gtgtgtgt aaaacctc agccacag gtgaggtc tcatgtgac
 78601 cgtacacagg agacctcgt gaaagcag cactgtct gactcacct gtgattcc

78661 agctcctgcc ctcaggcggc cctgcggggc ccactggctg acggggagac ggcaccgccc
 78721 tcccccgctg tcagggtggg ggggctgacg attgcatgt cgtgtcaggg tccagcggcc
 78781 tcccttgcgt ggaggtcccg aagcacctgg agcgccgccc gcagaacagc ggactcctgc
 78841 ctgccicctt gctcttgcc atggcctgcc cgcctctggc cctctttctg ctcggggccc
 78901 tcctggcagg tgagccctcc caaggcctgg ctcacctagg ggtgtgtaag acagcacggg
 78961 gctctagaag taaatcgcg ggaagtaaat cgtagtggg aggggggatg gttccgaag
 79021 gggccctgag ggggacagga gacctggcct cagtttccc actggtgagt gaccagatag
 79081 ccagggtacc ttggactct gactctgggg ggctctcaga gactggtctc ctactcagt
 79141 tttcagaggg gaagctggtg tggcctgtc actgccctgc agggcctcag ggacaagcta
 79201 tccctgagga ggtctccagc agtcagtggc cggaggctga gccgatgat atagtaacag
 79261 cccaggcggc ctcttggggg tggtcagcct gtagccaggt ttggacgag ccgaagtac
 79321 ctaagtgat ggggtctga gagcaaggga tgagggtggg cagcaggagg acccagagcc
 79381 caccagccca cctctgaat tctggacct tagctgcatg tggctcctg ggaagacggg
 79441 gcttaagggt tggcctctt gtggccaca cagtctgat tccacagcac tggctgtgag
 79501 cttttgggag cagattctc cggggagtct gaccaggct ttgtggggca ggggctggag
 79561 ggaaggggccc caggccagac ctgagtgtgt gtctctcagc ctccagcca gccctgacca
 79621 agccagaagc actgctggtc tcccaggac aagtggccca actgtcctgc acgatcagcc
 79681 cccattacgc catcgctggg gacctcggcg tgcctggta tcagcagcga gcaggcagcg
 79741 cccccgcct gctcctctac taccgtcag aggagacca acaccgggccc cccggcattc
 79801 cggaccgctt cctgctcagct gcggatgcag cccacaacac ctgcatcctg accatcagcc
 79861 ccgtgcagcc cgaagatgac gccgattatt actgcttgt gggtgactta ttctagggt
 79921 gtgggatgag tctctcctg ctgcctgcca ctctactcc tgacctggg accctctctc
 79981 tgagcctcag tttcctcct ctgtgaaatg ggtaataac actcaccatg tcaacaataa
 80041 ctgctctgag ggttatgaga tccctgtggc tgggggtgtg ggggtaggga tggctctggg
 80101 gattactgca gaagaggaag cacctgagac ccttggcgtg gggcccagcc tccccaccag
 80161 cccccagggg cccagactgg tggctctgc ctctctgta cgggaggagc tggagtgaga
 80221 gaaaaaggaa ccagccttg ctggtcccg ctctgcatgg ctggttgggt tccaacactc
 80281 aacgagggga ctggaccggg tcttcgggag cccctgccta ctctgggtg gggcaagggg
 80341 gcaggtgtga gtgtgtgtg ggggtgcaga cactcagagg cacctgaagg caggtgggca
 80401 gagggcaggg gaggcattgg cagcagccct cctgggtag agaggcaggc ttgccaccag
 80461 aagcagaact tagccctggg aggggggtgg gggggtgaa gaacacagct ctctctctc
 80521 ccggttctc taagaggcgc cacatgaaca gggggactac ccatcagatg nnnnnnnnnn
 80581 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 80641 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn agagggtggg tgggtggaat ttaatatagt
 80701 ggtgcgcgtg gagcgtgggc ggcgcattha aggcggtcat ctaaaatagt ggatagggg
 80761 tgggtgaca ataacgggtg gtggatgtg ttacggggg gtgcaatagt tctgagttg
 80821 ttagtgtctt ctgtatggg tgcggcgtg tggacctacg cctgagtat gtggggggg
 80881 aaaagcagtg agggtagtag ggatgggaaa tattgtgga ggttcttgt tgggtattt
 80941 ttgttatta ttttgggtg tggagtgtg ggttgggtg aatttcgctt gcgttatgt
 81001 ttttttct tttctgtc gtgggtggg ttggtggtg cttgtggtg gtggtgggt
 81061 gtgtataaa aaaaaatgt tgggtgtgt cagcttagcc ctataacggt cggctttgt
 81121 tctgtttgt tctgtggcg tgagcggatg gtcgggccc cgtgctccg cggcgcgcc
 81181 tcgcgcgcc tctgtctcc gctgtctg ctgctgtgc tcccgccgc gccgtctg
 81241 ctggcccgcc ccccgccgc gccggtgagt gcccgccgc ctccagcccc cccgccccgc
 81301 cccgcccctc acgcgaggg gcgcggctc gcagagctgg atccaagggg gtgcccggga
 81361 gtggcccgcc gcggccggt accccgaaac gctgtctggg tggccgggg gtgtggtgga
 81421 tagtgagctt cccgtccctg gaagtatgca agtgaagccg gcgcgggat cgtcgggt
 81481 ggctggtgag cggcggggac tggctcggg gctagacgca cggcgccagc cccccagctc
 81541 ccagacctgc ccactccgc cccgcccgc cgcgatccc ggtgtgtgtg tgtgtgag
 81601 gggagggaca gcgggagtgg ctacagggt cccgactcac cgcagggaca aagaccgcg
 81661 ggtccccagc tggcgtcagc cgccaggtgt gtggcctcgg tgagcacacc tccaggcggg
 81721 aggggtgagg gaagcgtgt ggggagggca tgcgggtct gagcctggaa gagacggatg
 81781 ctaccgctg ggacctgtga gtggcgggat tgggaggcta tggatcagg aggcagcta
 81841 agcgtgagag ctccggtgt gcctggcggg ggtggtagg gggggagcc cctgtgtgtg
 81901 ccagcctgcg tgtgccctaa aggtctgcgc ctccccact gctggggtt cgggggacca
 81961 gtcacagcct aggtactgc aggcgcacag ctccccggga gcccgccca cgcgggtgtg

82021 ccgctgagcc tccagcctgt cggggcaggg gtggggggca gggatggggg cttagcggg
 82081 gttgggggca gacggccagg cagactctct gggcacagct ccggtgacaa gggaggctctg
 82141 gcaagcctgg gcccctctg tccagccacg ccagctctgc cctggccagt ctgccccct
 82201 ggcatgtctg gggatggaag ggggagcggg tacctcagtc tgggggccct gcctcctccc
 82261 cagccccgcc cggcccccta ggcctagggg cagagtctag gggtcaccct ggggagctgc
 82321 tgaatccgcg ggfttaggaa ccggaggggac ctgggctttt gaaccacgtg gccctagggtg
 82381 agccctcggg cgctcggta gccctaccc ccagcctgt ccagggtggg ggggtggagg
 82441 cgacagtgc cactgtggg ctgaacagcg tctgcaggga ggccaggaga gctgggcaca
 82501 cggacacgtt ccatcacctg gagctgccac tgtgccactt gtgcggggc aggcggggc
 82561 tgagccgggc tgcctctgt cagccacag atatgcagg ggcactcggg gtcgctcgg
 82621 acatgcttat ccttgagcg ctgttgagc ggccgggaag gctctgtaa tatttatcca
 82681 tcccagctca cagcttcag ggtgatgaa agccccggg cccggccact gtgggggacc
 82741 ccgcttccc tctggagcc agcggggtga gggggtgggg gagatggacc tgcctgcca
 82801 ggagcaggcg gtgtactct ggcaggtcac ttgacctc tgagcctcag ggaggggccg
 82861 ggatgggtg cgatgctct ctgcttcc cccagcctga ccagtgtcct cccctcggg
 82921 tgcctcctg cccaccgag agggggtggc tatggggacc tggccgatg gcaggcaggc
 82981 cggagaggcg atccccgct cagccgtgc cagcactcc cagtcaggg gccccgcca
 83041 ccccagccg ctggctgct cccatttcc cgattgcagg ttggccccga ggctgaccg
 83101 agcctctggc tcagctggga gactgaatt cccaagcaat tctcaagga tgtgtaggc
 83161 tgtgtgtgg tgcctatcc ggagaggtg ggtgagcga ctgggcacct cggccaggg
 83221 caggccagg gagacgtgg ctgacgagca ggcaggcctg caaggaggac gagcagccat
 83281 ctcaggaatg tgggtttgg agacaagca cagctggggg ggtggggggg ccatgggtg
 83341 ggaggcctga tcccaggtc taggtccagc tctgggctcc ctgcctgtg gacctgggc
 83401 caagacctg acctctctg gcccgtctc tccccggg aggtggggcg atgctgctc
 83461 cccaatccc cagggtgtg gatgaggcag acgaggtgtg tgcctatcc cactcactg
 83521 cctccagca gcccggggc ggggggggtg tgggactgg cgcaccagg tgaggatcag
 83581 gccttgagc tagggaggc ccccagccc caggccagaa aggacacggg gagacagaat
 83641 gcaggaggcg ggcagagcag gggccagcg tgggaaact gaggccaaga gcctgtggac
 83701 gatgtctcc aggaaggac ctgctgctt ggggctgga tctagagcc tccaggagcg
 83761 gtgacctga cgtggcagg gaaccggagg ccccgcttg caggtggacc cggcgcgagt
 83821 cacttctct ctctggcct gagagcttc tccagctgc cgtcctgtg ttctaatgc
 83881 aagtctggg gcctggggg caggtgggg ctgactgcca ggtgggggag ggcaggaatt
 83941 tggcagagca gcgtccaga gtgggagaag ccagccatg gaggggactc tctccatgc
 84001 tctgcccc aaggcggtta tagagagagg tgggttacc ctccgcatg gcccgttcc
 84061 cattgaacag atgggaaagt ggaggctgag agaaggctgt gacttcccc ggtctcctt
 84121 ggcatggaac tgggctgct gactctcagg ccggggtat cgtctgca ctgagcacg
 84181 caggatgcag gggcttggc ctggacctg cgcctcgtg gggcaagaga ggaaggcacg
 84241 ctgggctgc ctgtaccct ccacccacc gtggttgtt gctcaggcct tctgggggc
 84301 agaggagagg ggagattca ctgctggca ggctaggccc tgggtctct ggggtcctg
 84361 gggaacaatg cagccctgtt ctctgagg aggttcttg gacctcacc aggttgagg
 84421 aaaggattc tcttctctt ggaggtcac gagccgacat ggggaggagc aggggcaggc
 84481 ccggggccca cactcaggt gtgagacctg gacgtgtgt cttccacctg acgtggggg
 84541 tggggggtg gggccgggg ggatccagt aacctgccc ccaattgtc tggaagacag
 84601 cgggtactg gtcatttcc ctctctctc ttcttggc ctggtgggga cagtccctc
 84661 cctggggaag ggggaccca gcctgaagaa cagagcagag ctggggtcag ggtgtgctg
 84721 ggagcgaga gacctcctg ctctgctgc tggctattc tgggtgctt ggagtcggca
 84781 gctgtgggg agcggctgg gtgctgtct gagctctgg gtgcccagg cctgggagag
 84841 ttccagagg ctgaggccga ggtggggcc ctggcgccc ggctcctgcc ccaaatatg
 84901 ctgggaagg ccacagcgg actgagcaga caggccggg cagacggcg ctgaggctc
 84961 cggcctctc cccagctcg ctgtgacct cactgcggc cgggggtgc agggccccg
 85021 ctgtgtctg ccgtgtctt gcaggctgat cccacgggt ctccctgct ctctgagctt
 85081 ccgctttc caggcagggg aaccgcgacc tccaggctg gacgcgggga ggtgtatgc
 85141 gccaggtcag aatcacccct ccaccgggag agcgtgttc aggggacctg gcagggtgg
 85201 gaccagcat ctgggaact ccagccacc ccacccatg agaggggaca tacagaccac
 85261 acggaggctg tgcctcgtg gcagcaactg gagaacacc agcccgggc aaacataat
 85321 aactaaataa taaaagttt aaagatcgt acttaaaaa acaagtgtc cccagtgtc

85381 ggaccccagt tcccgggtgcc ctgagtgggt cgggccctgt gctgagcatg gcctgggtgg
 85441 ttcaccccca gatccacact aaaggggtggg atcaccccta ctatgcaggt gagcagatgc
 85501 aggggggggag ggcggcagcc cctccatgct ggtgggtggc cgtgggtgggt gtcctgggca
 85561 ggagccagct cacggagctg gagaggacag acctgggggg ttggggggcg ccaggaagaa
 85621 acgcaggggg agaggtgtct gccgggggtg ggggtccctt cgaggctgtg cgtgaagagg
 85681 gcaggcgggc ctgcagcccc acctaccgt cccgggcccc aacggcggga gtaagtgacc
 85741 ctgggcacct ggggccctcc aggagggggc gggaggcctt gggatcagca tctggacgcc
 85801 agtcagcccg cgccagagcg ccatgctccc cgacggcctc cgtggagtg aggtgcgt
 85861 gacaccaca ccgtgaccc gggcctctct cccgtcagg atgcccccg ccgccacccc
 85921 gtgagcagag ggccacagcc ctggcccgac gcccctccc acagtacgc cccgcctg
 85981 gccaccagg aggcctccc gcttctggc cgcccagac ctcccctg cggcgtgcct
 86041 gacctccc atggccgag tcccgcaac cgacagaagc ggttcgtgt gtcggcggg
 86101 cgtgggaga agacggacct cacctacagg tagggccagt ggccacgagc tggccttga
 86161 tctccactg ctgtctgaga cacgtggag ctggggggag ggcagatccc tatggccaac
 86221 aggtggagt gtccccaac tccgtgccc actgtcaac acccaaac cacactaga
 86281 tgcactcca tgcctcct tgggagcag gtctccacac ccactggcc acccacaca
 86341 cccgtgggc acggcgtta gtcaccacg caacctctg gggcaccgtg ctgcgggcca
 86401 ggccctggga ctctcagtga gggaggcaga caggccct cctccgggg agcgaggtgc
 86461 tcccacgcc cgttcagct ctgaccgc actcgggacc ctacaggga gggaccact
 86521 ggggcaggcc aggtgacgc tgggtgacc tggccctg gcgtgagac tacacttct
 86581 gcagtggcg gcgaagatg gtgtgtgt ccacgtgt gcagcggga ctctggggc
 86641 ctcggaagt tctggcg ggagcctgg gagcaggaag ggcaggtct ggggtccaag
 86701 gcctcccaac ggtcaggtt gggagggggc ctggggctc ttgggtctt tccgccagt
 86761 gcagaccctc gcggccacct aaggcacac agaccacaca aagctgtgc catgcagtgt
 86821 ggggagtgt gcgcaccctc agagcacact gggccacat cagcacgcc tggccctca
 86881 ctgtcatcc ggggaaactc ctggcccg cagccagcg ggtgacgt acccgtgag
 86941 ccagaccag gcccctca cggccctgt cctcccagg atctccgt tccatggca
 87001 gctgctcgg gaacaggtg ggcagacgt ggcggaggcc ctccaggtgt ggagcaggt
 87061 cacaccctc acctcaccg aggtgcaga gggccgccc gacatctga tgcattcac
 87121 caggtgagcg ggggcctgag ggcacccca cctgggaag gaaaccatc tgcggcagc
 87181 cactgactt gccctaccc acccccgac aggtactgg acggggaca tctgcccctt
 87241 gatggacct ggggcatcct ggccacgcc ttctccca agaccaccg agaaggggat
 87301 gtccacttc actatgatga gacctggacc atcggggaca accagggtag gggctggggc
 87361 cccacttcc ggaggggccc tgcgagggc ccggagccgg gcccgggctc tgcgtccgt
 87421 ggggagctc cgattgccg ggtgtctc ctctccagg cacgcatc ctgcaggtg
 87481 cggcacaga gttggccac gtgctcggg tgcagcac gacagctcg aaggccctga
 87541 tgtccccct ctacacttc cgtaccac tgaacctcag ccagacgac cgcagggga
 87601 tccagcagct gtacggccg cctcagtag ctccacgtc caggcctcc gacctgggc
 87661 ctggcaccg ggcggacacc aacgagatc cgccgtgga ggtgaggccc tgcctccct
 87721 gccacggct gcctctcag ctccaatg ggtcctct aaccttgc tctaccca
 87781 gccggacgc ccaccggtg cctgccaggt ctctttgac gcagcccca ccatccgtg
 87841 cgagcttct ttctcaagg caggcttgt gtggcggtg cgcggggggc ggctgcagc
 87901 tggctacct gcgtggcct ctgccactg gcaggggctg cccagccctg tggatgcagc
 87961 ctgcaggac gccaggggc acatctggt ctccaagg gagtgggagc cgggtcacac
 88021 tcaggagact gcagggagc aggaacgtc tggccaagg tagggacaga cagacgtgat
 88081 gagcagatg acagacggag ggggtcccg agtttgggg cccaggaaga gcgtgactca
 88141 ctctctggg cacagctgg aggtctctg gaggaggcg ttctgaagc gggagtagga
 88201 taaaaggtat tgcacccat gaagcagtg tgccttgc cctagagac aaggctctg
 88261 ggctcagagg tggtaagt acccatga ggcacagct tggagaatg cgggaggat
 88321 gtgagctcag tgtccagag atgggagcct ggagcatgc aaggggcagg gcctgtgcc
 88381 tgagagctg cactggggg ggcagccaag tgcagggat gagcgggccc ccaggtggc
 88441 tcttctgc tcagaacac ctctccatg tatactccc agcggcgtg gcattgcca
 88501 gtgtctct tggggcagg agtacaagc aggcattat actggcctt tgtgtttat
 88561 ggacaacga actgaggct ggaaggtcc aggtgtgtt ggtggcgga ggtggcgt
 88621 gggcagcct gttcagcac acaccccca cccacgtt ctccaacagg agctcagta
 88681 tgggtgtat acggtgaga gccgtctg gggccgccc cctctcca gctgggctg

88741 caggggtccc cgatccatgc cgccttggtg tggggctccg agaagaacaa gatctacttc
 88801 ttccgaagtg gggactactg gcgcttcag cccagcgccc gccgcgtgga cagccctgtg
 88861 ccgcgccggg tcaccgactg gcgaggggtg cctcggaga tcgacgcggc ctccaggat
 88921 gctgaagggtg tgcagggggc aggcctctg cccagccccc tccattccg cccctctcc
 88981 tgccaaggac tgtgtaact cctgtgctc catctttgtg gctgtgggca ccaggcacgg
 89041 catggagact gagggccgtg cccaggtccc ttggatgtgg ctagtgaat cagtcgagg
 89101 ctccagcctc tgcaggctg ggtggcagct cagaccagac cctgagggca ggcagaaggg
 89161 ctgccaag gtagaaaga cctggggct tcttggtg ctacagacgt aaagcgtctg
 89221 cctgcaatgc gggagacctg gattcgatcc ctgggtcagg gagatccct ggagaaggaa
 89281 atggcaatgc cctccgtac tgttcctg aaaattccat ggacagagca gcctggaagc
 89341 tccatgggt gcggaagagt cagacacat ggagcgactt cactgttta agggccacct
 89401 gagtcctca ggttcaagg aaccagcag tggccaaggc ctgtcccat cctctgtcc
 89461 acttaccagg cctgacctt cctgtctct caggcttcgc ctacttctg cgtggccgccc
 89521 tctactggaa gttgacccc gtgaagggtg aagccctgga gggcttccc cggtcgtgg
 89581 gcccgcactt ctccagctgt actgaggtg ccaacactt ccgtgatca ccgctggt
 89641 gtctcaggc cctgacacct ccacacagga gaccgtggcc gtgctgtgg ctgtaggatc
 89701 caggcagggc acggagtcgc ggctgctatg ggggcaaggc agggcgctgc caccaggact
 89761 gcaggagggg ccacgcgggt cgtggccact gccagcgact gtctgagact gggcaggggg
 89821 gctctggcat ggaggtgag ggtggtctg ggctggctcc acgcagcctg tgcaggtac
 89881 atggaacca gctgcccag gctccatcc acacctca ggtcggggc tcagcagggc
 89941 tgggggagct ggagccctca ccgtcctgc tgtgggtcc catagggggc tggcagtggt
 90001 gtgtcagggt cctgcgctc ctgcctcca caggggttg ctctcgtag gtgtgcctt
 90061 ccagtttgt ggtctggag acctattccc caagatcctg gccaaaaggc caggtcagct
 90121 ggtgggggtg ctctgcca gagacctgc accctggggg cccagcata cctcagtcct
 90181 atcacgggtc agatctcca aagccatgta aatgtgtaca gtgtgtataa agctgtttg
 90241 ttttcattt ttaaccgac tgcattaaa caggtcgtt tctacctgc ctgtggggg
 90301 gtctctgtga gtgcaaggcc agtatagggt ggaactggac caggaggtg ggaggtgtg
 90361 ctggggaccc gtcagtcct ctggtctca gggctgggtg ttggtcagg gctcccctg
 90421 ctccatctca tctgttga atgctacag tggcttaca gtctgctcc catctccca
 90481 gcggcctctc agaccgtct ccaccaagt gtgtcacgt ttccgatcc agccactgc
 90541 aggacacaga accgaactca aggtactgt ggctgactcc tactctctg ggtctactt
 90601 gcctgccacc ctacagagc caaggatccg cctgtgatgc aggagtgtg gaagtcgctc
 90661 agccgagtc gactcttgc aaccatag gactgtagcc taccaggctc ctctgtctat
 90721 gggattttc aggaagagt gctggagtgg gttgccatt cctctccag gggatcttc
 90781 caacctggt cctccgata gcaggcagac tcttactgt ctgagccacc aggcaatgca
 90841 ggagacctag gtcagctc tgggtggga agatccctg gagaaggga tgacaacctg
 90901 ctccagtatt ctgattggg gaatccatg gacaaaggag cctggaggcc tacagccat
 90961 aggtgcaaa gagacagc tgagcaagtc acacacacag agccctacgt ggatgctcat
 91021 agcggcacct catagctgc atgtatcagg tgttgcatg ggcagccatc agcagggggc
 91081 cattctgac ccactgctt gtccaccgg atacacgggt gcctctctgt gtgtcgggc
 91141 cactcggtg tcagcgcga agggcagggc tctcggagg cacagggcac agagttaagg
 91201 aggggatggg gacgttagct cctcccagc tctcagcga tgcagcaggc aaaacaaacg
 91261 ctaggatcc tgccaaacc gtagtctct gccatgctc gcccaccc cagagccaca
 91321 agaacgggag ctgggggggt gcccggagct gggatactg tccctgggccc cgccatgtg
 91381 ctggccgca cagcgtctc cggcgggga aactgaggca cggcgccctc cggttctc
 91441 cccgcttcc gggcctcgc tcttctcc tcaccagggc agtatccag ccccggtgt
 91501 gagacggaga agggcgccgt tcgagtcagg gccgcggctg ttatttctg cggtgagcgg
 91561 ccttccctgg tacctcact tgagaggcg ccgggaaggc cgagaaacg gccagggctc
 91621 cttaagggg cctgtgggg gcgcgccgc cctttgtcc ggggtggcggc ggcggcgacg
 91681 cgcgcgtcag cgtcaacgc cgcgcctgc cactgagggc ggctgtctg tctctcgg
 91741 cggcgcggc ggcggcggc gaggaggca acccatctg gcttgcaag agactgagnn
 91801 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 91861 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn gcaggtgccc gcggtgacgc
 91921 ggacgtacac cgcggcctgc gtcctacca ccgcccggt ggtaccgcc cccgggggtt
 91981 gccaaaggta cgattgacc ctccccccc cagccctgt cccctagggt ggggtgggtc
 92041 gggggcagtt tctaagatc cctggttcc cagcagctg aactcctag tccctccag

92101 ctctacttca acccgcacct cgtgtccgg aagttccagg tgaggccgcc ccgccccttg
 92161 cacttgetgg cccaacccct ccgcccagc gctggcctga ccgccccca ccccgcccac
 92221 cccacgcagg ttggaggct catcaccac ttctcttct tcgggcccct gggattcagc
 92281 ttcttttca acatgctctt cgtgtatct gcgcctggt ggaagcggga ggaggcgagg
 92341 gcgggggacc gggcgaggagg cagcgggccc cgggaagctg agacctcca aggggcacgc
 92401 ttctatacc aaagccgag gticcgtac tgccgcatgc tggaggaggg ctcttccgc
 92461 ggccgcacgg ccgacttct cticattgtt ctcttcgggg gcgtctgat gactgtatcc
 92521 ttccgggct cggggacct tgggtccgg cctctgctgg cctgaggcc ctgctgagc
 92581 gcatgccaca gaggagagt tgcgacccg agctgagggt gttttgagc gtacatcac
 92641 tgctcagctg cagggtcccc tgcgaactc cagggtaca ccaaaatc cacagggcag
 92701 ggtgccagg ggtgagtcc tgaatgcagg tagccaggag gatctagggc tgggcccggg
 92761 ggctggggtg aagtggagag gcaggggcca tcaggggggc cctggaggcc accgttgggt
 92821 cttagagtgg gaagcgaac caacctgctt gagggttca ggggttagg aagtcagagg
 92881 ggccctgggc agggcacaag acctgactc tggccagct actggggctc ctgggtagcc
 92941 tcttctctt gggccaggcc ctacggcca tctggtgta cgtgtggagc cgccgcagcc
 93001 ctggggtgag ggtcaactt ttggcctcc tcacctcca ggcgccgtc ctgccctggg
 93061 cgctcatggg ctttcaatg ctgctggga actccatct ggtggacct ctgggtgagc
 93121 ctgctgtcca gggagcctgc ccaagctgg gtgtgctgg ccagagccct ggtcctcc
 93181 ccgccccac cctcttccc cactcctggc gccccatcc ttccagcccc tccaacaagt
 93241 cagcctatag gtttactta ttgagcctg acccattgc tgacgttgt gtggggccc
 93301 acccgtagg gatgggtggc tcagggtgac tgctcacag tccacttct ctgactcct
 93361 caggcctgac ctctcccag gtctgccta ctctgggcca agcctggccc cagctgggc
 93421 tggctggccg tgcaggcat cagacccca tctttgggg gcttcagggc tgtggagggt
 93481 ggccctggca ttggcgcct tccacaggg attgcggtg gccacgtcta ctacttctg
 93541 gaggacgtt tcccaacca gctggaggc aagaggtgc tctgacccc cagcttctg
 93601 tgagtgtga cagcctccc cccccctc ccagatggc tcttaccac atgaggggg
 93661 gggacctgc cagctgccg tcagcgtgg cctctccca caggaaactg ctactgatg
 93721 cccagagga ggacccaat tactgcccc tcccagga gcagccagga cccctgcagc
 93781 agtgaggac acctaccca gagcgggtc cccaccccc accctggcc tgcaacgcag
 93841 ctccctgtc tggaggcgg gctggggcc agggccccg cctgaataa acaagtgacc
 93901 tgcagcctg tgcacagc actggctct ctgccggc cagcctctc acgcgggca
 93961 ggtgtgtgt gccgagagc agggccacca agcctgact gctctccgac ccagaacatt
 94021 ggcacagctg gaggccaga gagggtccag aacctgcca ctgccagca gaactctgag
 94081 cacagaggc agcctgtg ggtttctat cctgccctg cctgtccgt aattcagctt
 94141 ccactgatg ggctcacat tcaggggcg ggctgggact gggatgctg gttgtgctga
 94201 gctttggccg tgggggccc cctgtccga actagcaacc ccaaggga cctctgttc
 94261 atttccagc caggccact aaggacggc caggtgcaga agagggccag gcccttctg
 94321 tgactccga gcctcaagt tcagtgttg cagagtccag tggctaggc agaggcctt
 94381 gggaagctt gccctgccg ttgcagctg aggcggcag gacctcac tggccccag
 94441 ctacgggca ttggaggacc agtcgcacg gtgtttact cctgggtcgg caccagccgc
 94501 cgccggctgt cctttcaca gaggataaaa gtactcgtc tggagtga ctttaattt
 94561 gtcataaac ctctggcca gcagcgggt ccgagtggt tggcaggtga agggccctc
 94621 ccgggcctt ccaggcagg gccgcctgc cagcaggga ggcaggcagt gtcacccc
 94681 actggctct gggctcagg tactctctg tgggcccga acatctccc cagtgtgga
 94741 gccagtgtc cgtgaggcca gctggcctg aaaccttct ctctgaagc ccgctgtcc
 94801 ctggccctg atggaggga gagggtggag cgcaagtcc taggatgtc ttgcagacc
 94861 cccagacca ggggcgagg ccatctcag ccccccca actggaaacc ctggagctc
 94921 tggccctct ggtgtgagg cctgtatg cgacctcag cctgccagc aacggaagg
 94981 gcaggcccc gggccacgg cttaacgcaa ctggcctgg gtcacctgc gggcctgtc
 95041 ccaggaggaa gaccagggt ccacctctt ggtgccag tccaggtac gtggggacc
 95101 gtccatgca cagaagatg aggtcaccc ggtgagctg cgccgggcc tgcagagca
 95161 ccagccgcg gtggagggt gccccagct tctgtcagg cacgtgtgc tgggaggtc
 95221 ggccggagca gtgccacca cgtgcagc gacaggtgg cacaggcca ccagcagtc
 95281 ccgcacggga tgggcccct caaggccag agaagccag ctctggctg ggggctggc
 95341 tgggactgac aggtggcct gccctctgc cccactact tccagccac ccgggactc
 95401 aaggactgc tgagctggc aggtgggac ccgaggggag tcaactgct cgtggggga

95461 ggaggggagg tccacagggc tgagccctga gctgaaccct ggccctgctc gtggttgg
 95521 ggggtggggg gtccagtggc gccctagccc tgctaggcc cagctgggac gtgcgcggc
 95581 gagggcgagg ggccagccca tgccatgtg tccccgttc tcagctccat gctaccact
 95641 tgaagaaaca gaacctgtg ctttttatt tagaaagtgt tcttgccct gcctggggct
 95701 tctataaaa aaacaaacac agctcaacgt ggccctcct gaccagagac gggcgggtgg
 95761 gactggggct cagcagacgg aatgtgtccc cggcggcggg agaccaggag gccctggcc
 95821 cgctctcag gacggctggg ctgtcccccac ctggctccct ccgagccaga agatggagga
 95881 gaggtgggct gatctccaga tgctccctgg gagccaagcg ccacgggggtg gtcaccaggc
 95941 cggggccgtg ttggccagac gcctcatccg cctgtgggag ggggagggca gcaacccccg
 96001 gatctctcag gcaaccgagt gaggaggcag gagccccag cccctccctc ggccgctctg
 96061 ctgctggggg cctgaagtc gtctctgtc tcgccccct cccagggag agtgagcctg
 96121 ttctgggctg tggtagacc tggcgagg ccagcctgc ccggggcct gtctgctg
 96181 gaaggggctg gggcagcacc ttgtgtccg gtctggctc cggatctct tctcatctc
 96241 tgcctcgtc aggtgtcca gcagcgggca cactggta gcgtgcctg tgtccggat
 96301 ggcaatctc accgtgggca ggggggttc actgtggagg acgagagagg tagacggctc
 96361 acagagcagc tcagggagag gccctagaa agcagtgtc acccgtgc gggcagacag
 96421 gacatggagc ctggttctg caccggctc ccgacacagg gcggcgggc acgtgcaa
 96481 catggcatc cgggtctgc atgtgggag gggccacag gacagtgtg caggtccagc
 96541 cattccagt ggactgtg ggaggaggag ggccgtccg cccgtcagt gtccaggaga
 96601 aaggagagca aaggagtcca tccaccagg agtgagtc caggggcct gccctgacca
 96661 gcctgcaggg ggccctcgg cccacatcac agggggccag aatcataag cctgactgc
 96721 tccaccccg ggccctcaa agacgcct agactcgtc cgaggccac ctgcacacc
 96781 tctgggaag tggactcagg gctgggggtc agcctcgtg agggcgaaa ggtggggac
 96841 tctggcga gctgtgct ctgccagg ccaggccag cctgccggc agcctagcc
 96901 acgcccac ccacctgcc cggcgccga cgtggcctc cgggtctct cctctggct
 96961 cctgctggc cactgtgtc cagccccagc agtcggcct ccaggagccc tgcagagta
 97021 gccccagag ggaggagggg gcccgggga acagcacagg aacaaacaga cccctggct
 97081 tagtttagc tctcatctg gaaatggg acagtgtct tctgcagg ggttcagag
 97141 gaccactgc atgcaacac cagcacac cactgctg ggggctcgg cccagccgg
 97201 tggcccgag tccaggctg gtggtggg cggccagcc accctgccga cagtgttc
 97261 ccagccggg ggtgtgcg cagtcagaa gccagactg cagacccaaa tgcactct
 97321 cacgttgcg gctccagct gccttctt ggggcagcag acacgaaagt caccaagccc
 97381 acggcgacg gagcaaac gtcttctt taaacaagt cgggtcccg aggcctgtg
 97441 ttacctcc tgggtctc ggaagattgc atccagggg gtgttctaa accaagggt
 97501 gctcgggca ggctggaag gaggggctg gagccaggag cccacctta cgggcattc
 97561 gcttctggg tctcaaggc ggctgggacc ctgattccc accaccgcc aggtgcaagc
 97621 agggaggccg tgcggagga ggcagaggc ctggagggtc gtcttcagc tgacctact
 97681 ttacaacct cacaggtgc gcaggccagc tgggagcat ggtgtgccc tctggtaga
 97741 tgagaacaag actgcaggga gtatcccc tgaactccc caaccaggag gagacaaaac
 97801 tgggtgtgc cctctgtt aagatcaact gactctggc aaggggccca gccacccga
 97861 tggggaaagg gcagcttc caacaagcg tctgggagc ggaccggca ggcatggt
 97921 tctcagctat gacaccagc gcacaagc cccgagaaa acagctaagc tgggactgt
 97981 cacacaagt aactcaaac ccaagaaac caaaaaagc ctgcggatct tcagatatg
 98041 gggaaggagc ctgtatctg aatgtataac gaactcctga aaagtgaag tgttagtcac
 98101 tcagtctgt cagctctt caaccctat gacgtagc tgcaggctc ctctgccc
 98161 gggattctt aggaagaat actggagtg gttgccatc ctctccag gggattctc
 98221 caaccaggg attgaacct tgtctctt gactggcag gcgggtctt taccagtagc
 98281 gccacctg tagaaacct ccagggtccc tgagtgtcag agcaggagg actcggccca
 98341 ggctgtgag gggacctct ccaggtccc tctgcacag cagtggagg tctgtctga
 98401 gtcagctcc agggatgag gacttggtg cagatcact cccaggacct cagatctgc
 98461 tctgggaagc gaggtcccc aggtggccc caggccgct ggctcagct cgtgagccgt
 98521 gcgtggacag gtgcatgag caggcctccc acgggactc gggcgcggc tggacccgg
 98581 ggctgccagt ggtgcgggg ggcccgtgt ggcggtgtt cctctctg ctccagtc
 98641 taggaacat gttggcgt cctctggg ttctggaga agcagctgag atgcaaacg
 98701 cccacgcgc tccctcagct gtccctgtc acgggtggc ccttggtgac ggctccatg
 98761 cagggacggt gacagctga gcagccgct aaaaccacac ggggacggt gcagctcag

98821 cagccgcgta aagcctgaca tccaatttgg aagcctccc cagtggaga ggggcccggg
 98881 gacggggctg cccggggcga gctccaccgg gtcgggggtc acgaggagcc caccgcgctc
 98941 cccgccacca gcacctggga ccagataccc tccccgtct gagggcggcc tgaacgccgc
 99001 cccctccac gggggcgccc accgcctgct cgtggactga acaagaggcg gcagtggcct
 99061 ccagacccc tcgggggagg gcagacctgt ccgagactga gcacaagtcc agggaatgag
 99121 caagggtctc agtaatgtcc ccaccgggac gggacgggag gaggcgacag aggccgctga
 99181 ggtgcggggc agccctcagt agctggcctc aaggccccag gcagtcccgg ggcctccccg
 99241 cagggggcgg gggcgaccac cggcccgagc ccaggcagtc cggggcctc cctgcagcgg
 99301 gcgggggcga ccaccggccc gagccctacc tgaaggcgta ggtcttctga tgccagctca
 99361 gctgtccccg gatgtctag gcgatgtggtg tgacgaactc cccgcccagc cccagctcgg
 99421 agcacagctt cagagcgaac ttctcgggag agttctctt ctccgacatg tccactcga
 99481 actgtccac caaggagatg tccccacgt ggatgttcag ctggcccggg agcacagaca
 99541 tgagccagag cggccccctc tggggccagg ccgcaccctc accaccctt ctccccggaa
 99601 catccccgcc tegtcttgg ccgcgccct gtgtgtctac ttgggtaag gaaaacaacc
 99661 cccatctctc tgaaggggt taactagcga ggaagatgcg ctgtaactg gaaaactccc
 99721 taaaagaaa gcttgatct gatggctca ctggtgaatt ccacaaaca ttcaagcac
 99781 taacaccaat cctatcaaa tctgccccaa aaactgaaa ggaaggaaca catcataact
 99841 ccctgccttg atacaaagc cagacaaaga tactacgaga aaggaaagg gcagaccggc
 99901 acttactgtg gacattgatg tgaacctca gcagacaga gcaaaactac attcaccagc
 99961 acgtcagaag aatcacacac cgttataaat gatgggatga tgacacaacc acattataaa
 100021 cggggggct tactctggtg atgtaaggac ggctcagtaa gaaaaccgt caatgccatg
 100081 aaccactga acagagtga ggacaaaac cacacagtc tctgataat tgaggaaaa
 100141 tcattagaca aactcaacg tgcttcacg ataaaagcac tcagtaact aagatcagat
 100201 ggaaaccaca tcaacaagat taattcagtc aaaaaattca ctgcaagtat caccacaat
 100261 ggcagaagac tggtaactt tctctaaga tcaggaacga gccaaagata cccagtcttg
 100321 ccactttgt tcaatatagc gttggaatt ctactcagtg cagtgcagtc gctcagctg
 100381 gtccgactct ttccgaccc atggatcaca gcagccagg cctccctgtc catcacaac
 100441 tcccggagt caccacaact catgtgact gagtcagtga tgccatccag ccattcctc
 100501 ctctgtctc ccttctct cctgcctcca atccctcca gcagtaggc aagaaaaata
 100561 aatcaaagg atccacctg aatggaagaa gtaaaactat ctctggtcc agatgtaca
 100621 atcttatag cagagttta gatgctaaca aaatactatt agaactaat aatgaattca
 100681 gcaaggtagc aggatacaaa gtcaactgc aaaaatcagc cgcatttcta catgtaaca
 100741 ctgcacaatc tgaagaagaa aggatgaaca aattacaata acataaaaa gaataaaatc
 100801 cttagaaatt aactgatca aagagatga caatgaacaa tataaacat actgaaagaa
 100861 attgaagata taaataaat gaaaacatc ctatgtccat ggattggaag acttaaaatt
 100921 attagctgt caaggctatg gttttccag tggatcatga tggatgtgag agttggacta
 100981 taaagaaagc tgagaccga agaagtgatg ctttgaact gtggtgttg agaagactct
 101041 tgagaggtec ttggactgca aggagatcca accagtcct cctaaaggag atcagtcctg
 101101 ggtgttcatt ggaaggactg atgtaaagc tgaactcca atacttggc cacctgatgc
 101161 gaagagctga ctcatgtga aagaccctga tgctgggtaa gattgagggc gggaggggaa
 101221 ggggacaaca gaggatgaga tggttgatg gcatcaccga ctcaatggac atgggtttg
 101281 gtggactctg gaagtgtg atggacagg aggcctggcg tgctgcggt catggggtg
 101341 tgaggagtgc gacagactg agcgactga ctgaactga catgaatacc caagcaatc
 101401 taaaagcca aatgtaatc ctatcaaat ccaatagca ttctgcaga aacaggaaaa
 101461 aaaatctaa aattcatatg gaatctaagg aaaagcaaag gatgtctgt caaaacaatg
 101521 acgaaaagaa caacaaagc ggaagactca cacttctga ttacagaact tactgcaaag
 101581 atacaataat gaaaacactg tggactaac gtaaaagcag acacgtggc caacgggaca
 101641 gccagaaat aaactctca atagcagtc aatgattt caacagagat gccagacca
 101701 ctcatgaag gaaagtgtt gcaaccaac gtttgggaa aaaagaacc acatgcgaaa
 101761 gaatgaagt ggaccttac ccagcccat ctacagaaat caactcaaaa cagacagaa
 101821 atatggtca agccataaaa cgctcagaaa aacagagcaa agctttatga tgttgattt
 101881 ggcggtgatt tctcagatat gacgtcaaag gcataggtga taagcgaata aataaactg
 101941 acttaccaca aatacaaac ttctatgat ccaaggacac taccagacg ataacaaggc
 102001 agccaggga aaggaggaaa catccgaaa tcacagatc tggaacaga ccgtgcctg
 102061 tgagatacag ggaaccgata aaaacaagaa aacagcaaaa cccgactca aaaatgggaa
 102121 ggactccagc agacacagga gacagacaag ccgcccagc gtcactaat agcaagcaag

102181 gccgcgcaaag gcccgatatcc aaggctgtgg ttttccagt ggcatgtag gaaagagagc
 102241 tggatcgtaa gaaagctgag cgctgaagaa ttgattgaac tgggtgttg gagaagactc
 102301 ttgagagtc cttggactgc aagatcaaac cagtcattc tgaaggagat cagtcgga
 102361 tagtcactga aggactgatg ctgtagctcc aatactttgg ccactgatt cgaagaactg
 102421 actcattggc aaagacctg atgctgggaa agattgaagg caggaggaga aggggacgac
 102481 agaggatgag atggttggat ggcactactg actccatgga catgagcttg ggcaagctcc
 102541 gggagagagt gaaggacagg gaagcctggc gtgctgcagc ccgtgggtcc caaatcttg
 102601 gaccaagcga ctgaacaata acaatcaac agggaaatgc aaatcaaac cacagtgaga
 102661 tactgtccac caccaggcag gcgttctca gcggggttcg gggcagggtg tgcctcttc
 102721 tctcgtaacg cccccaggac cgcgggggct gctgagacag catgggtgtg gcttggccta
 102781 gcctgcccac gacaagagtg gcagtgtgct cgcctactg cgccttccc tgcctgccc
 102841 accagctggg ccaccttg gaccaccag ctccgctcc gtggacggca agggcgacg
 102901 agcggccgga cagcccca acgtgtgccc ctctcagaa gtcggcctgt gccctctg
 102961 ggacaagccg cccaagagac agcttccag agccctgcc cacaacacgg acccagaca
 103021 ggctccttg gaggcctcca cgcacctcg cactcgca gcccagga caaggcaggc
 103081 ccgtgcggg tgaggagccg ctacctga taatgacgc ctggtctgac tggcttcca
 103141 ggatgctgc cgtgggtg gactgatct gctgtgat ggcagaggca atggtggca
 103201 cgaatgtcag tgggtcaga tccaggctg cacagagaat ctctgagaac atctcggg
 103261 tcatcagct ctctgaaac atgacggagc gggggaaccc ccagtggacc acaggccta
 103321 cggcagcgt gctcagccc gccctcccc agccttgcct cctctgccac cgcggccg
 103381 ggtgacgaca ggacccctg gcagcagca gacagagctg agtcacgcc agccaggcg
 103441 gcggacggc cattcatgt ccaggtaaag gcacccgca gcttctgcc gtaactcc
 103501 atgtccagc ggatggggac cagcacctg ggctgggacg cgttctgtg gatcacggt
 103561 gggctgtgt cgtcgaagc ggaaggggag cggccgctg ctacgaaag cgggctggc
 103621 ccctgtgccc agggcctccc tctctgacc actggtcgt gagacctgc cagagaggac
 103681 ctgtccacta cgggcccggc cggcagaaac agggctggc ggggtccac cgggcccggg
 103741 ggggagctgc cgactggca gcgggacaag ctacaggtt cctgcagga agagaggtt
 103801 aagccccaga gcaggcagga ttctccagc agctgtggg aagaagggt atgtccaga
 103861 gaagaaaccc tgaacaaag gccgagggc aggagggtt aggagctgt tggagagcag
 103921 tgaagggggg ctggcggtt ggggggtgt ggggagcctc ggtggccaag caccagggc
 103981 tcccacctg cagctggac ccgagggag cccagagga cggagagca ggcagctcg
 104041 cactcacacc tgcctttag gatggggaag agggagaga cgggggctgc ggggggcaag
 104101 gaaaccaggc acgcccgt tagaccggg ggcgagaacc acttccaag aacgcaggg
 104161 cgccaatgat gaacaatgg tagcagccc caggcgggag gcccgtgac caggccct
 104221 caccagagc ggaaggtcg ctctgtcg cggccatgc ggttctgt gatgtgtg
 104281 gagcaggga cggcgtccag gtgtgcgag ctgtgggca ggtgggac cactggctg
 104341 ttctcttg ccttctgt cctgggagac acagacgcc gtcgctcag cctatgggc
 104401 aaaagccgc cccagccgc caggtgtg ccagtggac cccgcatgc cctctggg
 104461 ccaggcccc atggggacct ctgtcgccc agtccgcg tggttatcc ccaggctca
 104521 agcggcacct gctcggggt accagttta ggggaggagg agaggcagg gggccagcc
 104581 cagtctgta gctgcacc ccaggctca agcggcacct gctcggggt accagttta
 104641 ggggaggagg agaggcagg gggccagcc cagtctgta gctgcacc ccaggctca
 104701 agcggcacct gctcggggt accagttta ggggaggagg agaggcagg gggccagcc
 104761 cagtctgta gctgcacc ccaggctca agcggcacct gctcggggt accagttta
 104821 ggggaggagg agaggcagg gggccagcc cagtctgta gctgcacc gtgtatgt
 104881 ctgggctgg cactcaggaa agagggtcag ggttcacgg ggggtggcg gcagatttc
 104941 aggagagccc caggggcagc agagaggagg ctacaggtca tggttgggca gggggccagg
 105001 gctggagaca cagagagggt cccgattcg gggggtgccc taccaggtg gctgggagc
 105061 cctgggggtt tgcacactt cagcaggct gttattcag acgttggc cagcctgaga
 105121 caggtaatgc ctctgcctc cgggccttca gggatgaaa gatacttag aaagcgggac
 105181 tcaaagtaac tcaaggaact cgcgtccac agtggggagc cttctctcc aattcatg
 105241 gggcgtttac tacaggaaa ataccgaagg ccgtttgag ctgaggctcc cgggcccggc
 105301 tgtcgttg tgagactgt cgtaccctt gggccacat cctggtggc aagggggca
 105361 tcagtgcgt gactgcaga cacactctg cagccctgc ccacagctg caccatggt
 105421 gacgtccac cctggagaa cctgaccact gcccgttcc ccgtaaaac agcggcctc
 105481 caggatggg ggcagaggga gaggccttg cctttcact cctctctgc agcgggggc

105541 cctcgcaccc cagtggccgg gccagggagc gccccttggg gtggggcagg gagggatcca
 105601 cacaccaagg ggagccagga ccccccaaa tctgtgccc tgccctgata cccgagacct
 105661 ggggaaacgg gggactgggg ctgatgcggg caggaccaag aactgaggcg gtgagacggg
 105721 gtccccacca caggccatct ggctggcagt ttctactccg ggcttgcagg ccaagaggga
 105781 aaaggtgccc cactcagatc aggcgcctcc cgtccccagg gagggcctac aaggtcagat
 105841 cctttgtaac ttccacgggc aaaactggct tgctgggcct gtgcggggcg catgggcgtg
 105901 gaccaccaca ccttcccca ctgagtctcc agccggagct gtcaccagg tccccagg
 105961 ccagccccac cccgccacct tgcatagcc tctgtatcc aggcggaggc tgcccggtcg
 106021 acccctctg cctgatggcc tcaagtggac aatgcgagtc acgttgcagc acgtgagtgg
 106081 gacgggcagc gccacgcggg gtccgggcat ccgagtccca cactcagcc tccctccgc
 106141 tgacagaggg tctgtccaag agccctgggg gccatccagc cctgtccga cctggccggt
 106201 gtggaagagg ggggtgtcca cccctctgg ggggtgtgct gggcgtggg caggccctc
 106261 ctaagagtgg agccactgg tggtttct gcagccccac ctccacacag cagttctac
 106321 tgccagtaa caggaggcta ctggcctagc tctctccctc gtgtgatga ctcaaccagg
 106381 agcgttcag gccccacaca gggttctcg ctgtgcatg aggatctca agcccatcc
 106441 acgtgcatgt aatctctcc ggtaacttct ctagggaagc ccggctatcc tgccatctc
 106501 accgcaccac cagggcgaga aaagccatct ccagcgctca catccacaat gggccaggcc
 106561 gtgagcacac cacctcttc gggaggttgt gggggcgggg nnnnnnnnnn nnnnnnnnnn
 106621 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 106681 nnnnnnnnnn nnnnnnnnnn cgcgcccccc ccccccgcg gcgccgacc ccggcgggcg
 106741 gccccggcg ctgggagcag gtgcggggcc gcggcgctc gtgagcctcc agcccgagg
 106801 acgggccccg ggggcggcc cgggtgccag gccctgggag ccccgaggc cagagtcca
 106861 gagggccgga ggaccggga aggcggaga gaggtggga gcacggggt ccagccctag
 106921 gccattcag ccccaaagc atcggtgaaa ccattgctgg cccagataa aagcgtcgc
 106981 aacttttca ccccgcgga gactttagc ggtagctgcc cctaggggg aatggaaaaa
 107041 ccaggattta ccaggtgggt ggaggtcaca actgccaga tctgagaaa gaggggtcag
 107101 tggggcgga agattagtg ggagaggagc tticagaacc caagggaatg aaacgaggct
 107161 tgaggttgt tatccagcag ccgccccctg cccgtgagt gagcgaaggc tggccctt
 107221 attgtacat ctccagctc ttcgtagaa aacctagat tttaatact gtggcagctg
 107281 agtcaaaaa taaggaaaag cccgactct tgagagccag gcacaaggcg tctgtgacag
 107341 ggtctccagg ctgcccatt gcagtcttg aaacggaggg ttctcgaga aggaggtct
 107401 ggggtgcctg ccagaattgg aggggggggc gcgggaagtg aggaccaga agagagggt
 107461 tggccgctg caaggagtc actggacact ggagctgaag cgcagccga aactggaac
 107521 tcgaaatctg tctcgtgcc agccacaagg cctatgatt tcttggcga cgttcagcat
 107581 cttaggagga gctggcgggg gaggcggga gtctgtggc ggttgcagca gggcaggaag
 107641 gtgaggaacc tgaggctgt cagagagctg gttggagtga tgccatcg tgaccgct
 107701 ggagaaggcc tgagtagaga aggtctaagc ttaacggga aggggtggc cagggtgga
 107761 atgggtggg aagttgagg agggggagca gtggagatg ggtgtgag gaatggagt
 107821 gagcttagac gtctgagga tactgcagt ctgtgcttt ttccacct ggctgaaat
 107881 tactgaaa caaaacaacc ctgctctgt gacgcctag aggggtggga gggaggctta
 107941 agaggaggg gacgtcgtg tgcctatgg cgattcatg ggtgtacgg cagaaagca
 108001 cacagtatgt aattaccct caattaaaga tcaagtaca cttaaaacc ccaaacaca
 108061 cattgtaagt cagtagact ccagtaaca tttagtaag aagattcaac tgggaatgag
 108121 tccgccgtg actatctga tgaattccc gtgtctctt gaggccatc ctcttgaac
 108181 tccgtgtt ggggaagcgt gccttgtat ggagctctga ggagtaaatg agacgggct
 108241 gtagaaggcc tagtagtgc ttgcacggc cagatgtca ataacctga gttgtacca
 108301 ttatgtacc tcaagagtct cctggagct tgcacggtt ctgaatggg tctgcggg
 108361 ctccttggg gctcccat ggggtggg gctgagtg ggtgtcccg ctcttgcct
 108421 gtccctgtg gaacacccc ttccaccga gcagctctgc tttgtctct tgtgttgt
 108481 tatatctct agattgtgt tcatgctc agtctgtcc aacttccga cccatggac
 108541 tgcagcacac caggccttct gccttacca tctccggag ctgtctaaa ctctgtcca
 108601 ttgagttgt gatgcctc aaccatctg tctctgtc tcccttct ctttgacct
 108661 cagtcttcc cagcatcagg gtctttcca atgagtcag tcttgactc aggtggcaa
 108721 gtattggagc ttacgttca ttatagtc ttcaatgaa tattcagggt tgattctt
 108781 taggattgag tgactgac tcttgcagt ccaagggact ctcaagagtc ttcaacacca
 108841 cagtcaaaa gcatcagtc ttggcactc agcctctt atgatcaac gccacatcg

108901 gtacatgact actggaaaaa ctttggctca gagataattg acttgattga atacaaagt
 108961 ctttggcaaa aaataaaagt gtggcaagca gtactgacac aaaagcaagt ggcttttct
 109021 ccgttgagtc atttattat tcagtgggtg tgtgcgtga gagacggagc ggctgtgctg
 109081 ggagctgggg ctccacttc agaggagccc cggacctgcc ctccgggagt tcacaggcag
 109141 tgcctggggg ggtcctgcca ggacgcctgc cctgcgagt cccagtgtg tgatggatgc
 109201 gtgtcccgca tctgcggcca ctggggccac gtgcccaga ttgtccgggt ctgagggtgc
 109261 agagaagagg aggcatttgg actgagtctg gaaaaatgag catgtggcca cgtgagaagc
 109321 cagtggtag ggaccagtc aggcggagga aagagcggct catacagatt gtggagctgg
 109381 aagcatgagg gtgtgtggaa gcagaggccg gggacagggc cgcagggccg gccatggagg
 109441 gcgtgggctg ctgcaggctc ctgagaaggg ggacgtgcc atcatgaccg ggtttaggtg
 109501 ttgacctg gtgtccagt agaggacaga tgtgtggggg gggagctgga gatgggcatc
 109561 catcgggagt cagcctggag agaggcagag acccgtcag tgggccctca ggacgtggat
 109621 ggggcggatg ttgggaagat ctgactctg gttccggct ggggctccg gctggagggg
 109681 tgcgcccac cgagcacagg aggcaaacag atgcccctc ccagcaagac cccagcccca
 109741 gcacctccg ggcccgact cggccctct tccagaatgg ctccctgtg gtccctgccc
 109801 atcttccg tgcctgagc ctctagatc tggacaccg cgtccgctt gcgctgtt
 109861 ctgggaagtc tctgctgt ctctactca cccaggaccg tctcgaggg caaggttgg
 109921 tcttgggtc catctgtt ggggtccgc tctcgtgc ttgacctg gatgtgacag
 109981 tgtcttgt ttctttca gaaccgaga gcagctgtg gtgtccaga cagaccagc
 110041 cgtgggatg acggccct ctgtggatg cccccggc gccaaagtgg gtgaggctt
 110101 cgtgttgc ggcgggctg acatgcaggc agacctgtc gcggaggagg acctggggg
 110161 ccccttct caggggagg ctctggagca gatggcgtc atctacaagg agatccctt
 110221 cggggagcaa ggcaggagc aggcagatta ccggggggac ttgatctgt gctccagccc
 110281 tttccgct cagagcgtc cccgggaga caggggccag gacgatgagc tttccggccc
 110341 gacctctc cagaaaccg acccgactg gtaccggatc acgggcagc gggaagccg
 110401 cgatccgct gccaggagg cgttgggag ggtgacttg gggctgcagg ggcgcccag
 110461 gaccgcag cccccaagc cctacgctg tgggagtg ggcaaggcct tcagccagag
 110521 ctgcacctg ctccggcacc tgggtattca caccgggag aagccgtatg agtgcggcga
 110581 gtgcggcaag gcctcagc agagctgca cctgctcgg caccaggcca tccacaccg
 110641 ggagaagccg tacgagtgc gcgagtgc caagccctc cggcagagct cggccctggc
 110701 gcagcacgc aagacgcaca gcgggaggc gccgtacgt tgcgcgagt gcgcaagga
 110761 ctccagcc agctccagc tgcgaagca cgagcgcac cacaccggg agaagcccta
 110821 cgcgtgccg gagtgcggc aggcctcaa ccagagctc ggcctgagc agcaccgcaa
 110881 gatccactg ctgcagaggc cgcacgctg cgagctgtg gggaaggcct tctccaccg
 110941 ctgcacctg ctgcggcacc agcgcgtca caggggcaag aagccgtac cctgcgcgga
 111001 ctgcggcaag gcctcagc agagctcaa cctcatgag caccgcaaga cgcacaggg
 111061 cgagaggccc taccgtgct acaagtgcg caaggcctc agccagagct cggcgtcat
 111121 cgagcaccg cgcaccaca cggcgagag gcctacgag tgcggccagt gcgcaaggc
 111181 ctccgccac agctcgggc tcatccagca ccagcgcag cacacgggc gcaagcccta
 111241 cgtgtgcaac gagtgcggc aggcctccg ccaccgctg gcgctcatg agcactaaa
 111301 gacgcacag cgcgagcgc cctacgagt caaccgtgc ggcaaggcct tccggggcag
 111361 ctgcacctc ctccgccac agaaggtcca cgcggcgag aagctctagg gtccgccgg
 111421 ggcgagggca cgcggccct ggcgccccg gccagcggg tggacctggg gggcagccg
 111481 gacggcgga tccggccgg ctctctctg cgtgacccc ggggggttg tttgccctc
 111541 catctgtt ttctaaagt cagacgaata cagtcagag ggacgaagt gggtaagcc
 111601 cccgggagac gtccggcag ctctaactg agacattga agaagtgaag cggactcga
 111661 gccgtacag cccggggaag atgagtcaa agtcgagggt cacctggcc actgcagggt
 111721 cgctcggcg tggggcggg cgggtgcagg agggctctc ctgggcttg ggtggcaggc
 111781 gaggacccg cgcctctag cctcggcct gggttgctg agggcgggc tggctgtagg
 111841 cctccagcg gaggtggag cgtgcccgg ctacgacag cacaggacc tgcacaggg
 111901 agtagccct cccagaccc ggcgtccag ctggggcgc tgcggggcct cgttctgtg
 111961 gctgggcag ctgcacctg tccagggat aagggttcc ggtctgaag gctgggtca
 112021 gggctcag ctggccctc ctgcttggg gctctgagg aagcccaag gctccgttc
 112081 cctctcagg agtggggac gttgggaat ccacattccc ctgggggtg tgtgtgtg
 112141 ttaaggct ccatcagac tgggactgg cactcacag ctttggaac tggcaactga
 112201 ggacggagac ccagggtgac accccacct ctgctcggc cccccggc ggggagacac

112261 aggcccgctct ggttcccaag atggcagggc cctccccct ccagcttggt cctgggtgt
 112321 ggtgectggg gctacagcga cctttccgg tccccgggc cagttcagct ggcatcctc
 112381 agggcggggc tctgagggtg ccatgttcc agagctctc ctctccac cagtagcagg
 112441 cgggcggcca gctcccaggc agccccctgg catgcctag gtgcacacct gcccgctgtg
 112501 acccagcaag gcttgaagggt ggccatccca gttaaagccc ctgccctgg cccaggaatg
 112561 ggctcgggca gggccgcac tggtgcccc agaagcgtct gtccctggcc tctgggagtt
 112621 ggcggtggtc tctgtactg tccctgcag ggccccttag cactgctgg ggaggaggtg
 112681 ggctgaactg atttgaagt ttacatgtc tgcggccga gtctacgag cccgtcaggg
 112741 tcatgttgtt tattcagca gatggggctt ggctcggcag ctaggatgtt cctgaataa
 112801 aatgggaagg ccagagctgt tctccatca gcaggcttg cagctgggga cgttgaaagg
 112861 acaggctgtc tggctgggg agaccagctc tgtgcagccc ctgctgtcc tgggggtact
 112921 aaaccagccc ctgtgtgccc ccatctgagt ggcagccgc ctggaggatc gccatcact
 112981 tgtgagaatt gagagaatgc tgacaccccc gcttggtgca gggggacagg gcccctaag
 113041 atctacctc ttgccccacc cccgggaccc cctcagcctt ggccaggact gtccttactg
 113101 ggcagggcag tcatccactt ccaacctttg ccgtctctc cgcgcgtgt gctcccagcc
 113161 aaattgttt atttttcc aagcatcact ttgcacagt caccactctc cttaaacca
 113221 ccttccgga gtctctgt cgtaatgc cgtttcagc caacctgggt cggccccc
 113281 gccagcaag cctgtgagc cccgcgcctc ccagctactt cagctcgc tcaagcttct
 113341 aaacgggac ctctcccc ccaccccat cctttctt tctgattat gtaacacggc
 113401 aggtaagact cctctctga aggggtgaca gactcacaca aaacgtgtt cagaccaggc
 113461 aagtgtttt ttccagaagt gtgagcgga cctagtctc agctcatgt cttcctgt
 113521 ttcttatgt gttctaagtc cttgacttg ggctcccaga cagcgacgtt gtaagaggcc
 113581 gtctggtag catttgaatt gtctcagat ttctgtctg gatttgtt tattgttta
 113641 gtttccctt ctttagcag acgttgtga ctgtgtaaa gctccagtc ttggtctgt
 113701 ttactaatca aattgtttg tcaagtaca tgtattctg tctttctt atcttttg
 113761 ttgctaata ttaacattt acatttcta gattaattat ttaggtaatt aataatttt
 113821 aacatttcta gtaaacgtg gtacttgggt ctgtgttgt ttctttag ttacagctt
 113881 ttctgtcta tactgtgac gctgggtt ttttgctc ttaggattt cctttgacc
 113941 ccattattat tatttaatt agtattttt aataattaaa aattagtgt ttaaatat
 114001 ccctaactc aacccagtg atgactgctt cagtcattgc tgttactat tatgtgctg
 114061 tgtcaggatt ttaagtgc catagacatt ctctgagcct gaatatatta tcagtttat
 114121 acagcatttg tgtactctca agaaacgtgt ttactctg tcagttcgtt ttgttacct
 114181 agtcttatg ttatttctt ccagtcgca ctgtctta ctgtcttc ctgcaggtg
 114241 tgaggacgcc tggcagccg tgagcatgcc ggggtccggg gtctgggccc caggcgcca
 114301 gcaaagccct gtgggtgtgt gcacggctgg gctgtccgg gaggaagcct gtggccccac
 114361 ggtagttagg agcgtgtgt tactgtgca caccacggtc tggtttgtg tctttccc
 114421 tgacgtgtt ctgtttgccc ttggttcta tctgttta ttagtgccgt ttacgtttg
 114481 ttactatgc cgttatctg atagacagg gtacgtgat caagtatta cgtatttg
 114541 agcagatgc tattaacag agatgaactg agaacctgt ccttgcag cctctttg
 114601 ctctttaat gctttagct tcaactctc ttcccaaac attataatgg aaaccttg
 114661 cttttttt ttaattgc attgcatga gagttattt agctcgcat ttttttta
 114721 aaatttgtt atatatttt gctatatc tgaacttat aaacagcaa ttattggat
 114781 ttgctttct attcttctg taattctt tacataaga gtctctat gagtaacatt
 114841 gctgtttaga gtgaggcatg attatttcc agcttagtat gtattgggtc ggtaacccc
 114901 caaaggcat gctatcccc gcccatctc tgtagttat tgtccagtg tggagcgccc
 114961 tgttaggcc gacgagagac ccacatcgg gcacacctgc cctctgtt ctgtcagtg
 115021 ccgggtctg tctgagtc actctgatg tcacaggctg gtgttcagc gacctggct
 115081 gtgacacgga ggtgtgatg gactgccc gcccatggg gcttgaggga ctaaaggatg
 115141 cacacctgc tggcagactg agggcacagg tgtttctac actgtcagc tttgaaata
 115201 ttctttgat ttctacct aactccaaa ggccgttcaa cataagctag aatgtactg
 115261 ggtgctgat tacatttag aaaagttca gcaataacca cgagatgcag caaagaacta
 115321 gacctcacag atcagggcgc ctgcataagg gagccacac agtcgtggga gacggggacc
 115381 ctctccagc tctgtctgt cccaggttg tccctcacc cggccctct ctccctgc
 115441 cctctgtgg tggggcgcc ccacatcac agctgcagag cctcaagaag ggggtcgccc
 115501 tggccactcc cgtggcagga gggacagag ggcaggagct taccgggt gcagtgtct
 115561 cggatcagct cagctggccg ctgcgggtc ggggggacag ttactggga ggcaggagcc

115621 cccactacag ctgccaggac ttctcagagg tgacaagggg gttcagtcac ctcagcccag
 115681 gtggaacca aatggcctct tgcgcggctc ctggggccac gcggagggtc gctgggatca
 115741 caggtatctg gatgtgtgcg ccatggacat gcaccacctt cggggggtaa ggggtgggga
 115801 aaggcagccc cttcttttg ggggacccc tcttcagtgt ctgataacca ggaaacaaa
 115861 tcagaagggtg gctcgggggt gctgagcagg gtgtctcta caccacagge cacacactca
 115921 cacagcctcc aggactccag tggggctgag cgctggagac tcaccacgt ttgtacccc
 115981 cccaccaag gccatcccag aacagctgcc tgcgtctca cggctggccc ctcctctg
 116041 gtctaacca gtgtgggtgg gccggcctgg ggtctccacc tgcctctgc tgtccctgg
 116101 gctgctggt gctgcagat gcggggccct ggcccggaga agcccatca gagcccagag
 116161 gacgggagtg gagcggggag gtgagcccc gagtctcag gggccagagg caaataactg
 116221 ggctgtgtcc ctggaaggca gttcccatg aaaccttcaa tataggccgc cccagacgat
 116281 cagcctcatc tgctacgtgg attcctccc gtagcgaatg gtattgggt tctacatgga
 116341 cccgggactt ctgttgaat tataatctt cccccactgc cctccaggg atctggaaaa
 116401 tggaggcctg ggctagacgg aagcttctc caagattctt tattgaaggg attcgaagag
 116461 aaacaggtgg tcagtaatc gtgggggatg gaggggtgag cgctacgtg aacggttta
 116521 ctgtgtctac gggaccagtt ttgatgtct tccccitcaa gaagcagacc caaacaccga
 116581 gatgtgagg ttagcagcac agagcgggtt catccacaag gcaaccagge agggagacca
 116641 gagacgtct ggaatctgcc tccctatggg cacgggctgg gtgtcacgg atgaagacca
 116701 agcagcaggt ggcgtggggc gtggggagcc tgcggaaagc gatggacaag gtgcgggacc
 116761 gcggtccgcg cgggtgaccc aagctccgc tctgcgtgc agcgcagct gggggcggag
 116821 cttccaggga cccgcgaccg cggcagtg gagggtccgc ggtccacca gtcctaacag
 116881 ctcagctcca gctagacgc gctgagtcg gctttctaga gagcaacccc ggcgggtatt
 116941 ttatggttct ggcttctga ttgaggaca cgcgagctt agaacaccct tgattagtgc
 117001 gggcaggcgg aatggattg actgatcac atctgcagt tcaccatctc aggggccgcc
 117061 ctcaccccca ctatctgc caaaggggg gctcgggtc tgagatcgg gccacacgtg
 117121 cactagacgg tcggtcagcg ctgctgtga gcggaccgg gccatctc acaccgccac
 117181 tggccctgt gctcaataaa aggaaggaaa gcgggaaaag cgcttctgg ccgcggtggc
 117241 ctgcgcgtt cctccatgc catctgtgg cagagccgg catggcacc gctgcacaga
 117301 aacctcgtg tccgttggg tccccatcc ttgacccga gagagaccc tccgtccaa
 117361 atgaaaaaca gctgtccca agagtcatta taatcacag caattgtgt aattgtctt
 117421 cggatccact cacagttca cgaacattc tgtaacctc tgacaactcc tacataaagc
 117481 aatactgaga agaaaagaac gtggtgata aatacaagg catacaaaa taaggagcaa
 117541 agaaaaaga cagtcctgc agttctgtt tttcatctc tcatgagtag gatggcagat
 117601 aaaacacaga atgccagtg aataattta gtctaagtat gtcccaata ctgcctaac
 117661 ttcaaatca acctatttt taaaatatat atttttgc gttactcat cagttcatgc
 117721 accaaagcct ttgttctg actcctaact tttgacccc tctgggtga ggagacccc
 117781 taacctcgag agcccatcac acagtccct tgggactaga ccttctttg cccatcacag
 117841 ctgaccgaa gggccagccc atggcagcg ctgcgcccc ctggcgaca gactctgcgc
 117901 ggcagcccc ggagcccagg tgcgacccg cggctctgg cgcctctag tttgaaaga
 117961 tctctctg gtgtcccag tcattgggt gtattttt agagaagatg ctgcgtgac
 118021 gatgatgat gtccttacc gggaggcagc ttggggcgc gtcggctcag gggccagct
 118081 attagcctgc atcgcgccc caggcatcgc gtccccctga gccgggtcag ctgtgggctg
 118141 tctgacacg ggttcccc agtctctgg ccgtgtcc tccaggtca gtgtccagc
 118201 ttccctct ggttggac ttgtcagcg gtctcagc atggagggg gaccctaaag
 118261 gatgtattga ggcatctcag cactgtctc cggcagggt tctgtgtcag cagtgaagt
 118321 accgggaaa ggggtgtct tgggtcct tcagaggcct gggttagacc aaagtttct
 118381 agaagattca ccattgcagg gagtcaaga caaaactagg gtgtcagca atctgtggg
 118441 gattcggcg tgagggaatt ctgaatgta catgtaagg ttactatt gttagggaac
 118501 attttccc cctacaaaca gcaggccaaa atactgagat gtcaggttg catcaaagag
 118561 cgggttcatc cacaaggcaa ccagagaacg ctctggaatc tgcctcctg cgggcacagg
 118621 ctgggtgctc acggatgaag accaagcagc aggtggcgtg gggagtggg agcctgggga
 118681 aagcagtgga caaggtgca ggacctcgg cgcgagctg aggcggagct tccagggaca
 118741 cgcggccacg cccagtggga ggtcagcgg tccatccagt ctaacagct cagctccaac
 118801 tagacgtgc tgagtctggc ttctagaga acactccgg cgggtattt attgtttg
 118861 ctctgtgact ggaggacgt caagtctaa aacaccttg attagtgcg ggaggcgga
 118921 tggattgac tgatcagac ccgcagttc accatctag ggccgcct caccctcc

118981 taccctacca aaggtggggg catcgggtgct gagatctggg gtgacacata aaatcaggtg
 119041 aagtcttagg acagggggcc gattccaggt ctaggggtgc agaaaaaacc tacctggccc
 119101 cgggctagac agcgtggagg gcgtggcccg ggctgggtga cagaagtggc cccaactgg
 119161 tcagaaggtg tgggagccca gggctgggtc actgcagaag gggctgcctg gtggacagag
 119221 tggggcctga gtgcctgctg aactggctcg tcagggtgc tgagcagaca cgggcatca
 119281 tcaactggctc ctgtgctga tagaaggagg ggaaccagg aaagcaaagg cgtttatgg
 119341 ccgctttgt gtgcgctt cctctagcac cgtctgccg cagaacgcg cattacatcc
 119401 gctggccaaa cctcggggc cggttggtat gtcccatcc ttgtctgga gatctacct
 119461 ctacagcagtt cccctgggga caatgtcag aagatgcag cttgacccg agctcgggtg
 119521 agagggtgcc ctgggttct tccgcagtg cttggagtgg aggtgcctca tgttgggtg
 119581 ggaacgggag gaaggaaaca ggctatgatt gagatgctc agacagactg tccctgctc
 119641 tgccaaattt cagaagattg tcttaataa atattccatt ttgtatgc ccttaggtc
 119701 attccagac actttaata tattgaaaga ctttaatat ttatataaaa atattatta
 119761 tagactgtat aaaaggaaca gttagaactg gacttgaac aacagactgg ttccaaatag
 119821 gaaaaggagt acgtcaaggc tttatattg caccctgct atttaacta tatgcagagt
 119881 acatcatgag aaacgctggg ctggaagaaa cacaagctg aatcaagatt gccgggagaa
 119941 atatcaataa cctcagatat gcagatgaca ccaccctat ggcagaaagt gaagaggaa
 120001 tcaaaagcct cttgatgaag gtgaaaggag agagcgaata agttggctta aagctcaaca
 120061 tttagaaaac gaagatcatg gcactgtgc ccatcactc atggaaatag atggggaaac
 120121 agttgagaca gtgtcagact ttatttttg gggctccat gaaattaaa gacgcttact
 120181 tcttgaagg aaagtatga ccaacctaga cagcatatta aaaagcagag acactactt
 120241 gccagcaaag gtccgtctag tcaaggctat ggttttcca gtgtcatgt atggatgtga
 120301 gagttgact gtgaagaagg ctgagcaccg aagaagtgt gctttgaac tgtgtgttg
 120361 gagaagactc ttgagaggcc cttgactgc aaggagatcc aaccagtcca tctaaagga
 120421 gatcaccccc tgggtgggtc ttggaaggac tgatgttga gctgaaactc cagtacttg
 120481 gctaccta atgcgaagagct gactcattg aaaagacct gatgtggga aagattgaag
 120541 gtgggaggag aaggggacaa cagaggatga gatgttga ttgcatcact gactgatgg
 120601 acgtgagctc gagtgaagtc tgggagtgg tgatggccag ggaggccctg gcgtgctggc
 120661 ggttcaggg gtcgcaaaga gtcggccatg actgagtgac tgaactgaac tgatccagaa
 120721 atttaaat aatatataa ccaaatccat gcagacaatt ataagcatat attataaat
 120781 cataattata agcaagtata ttttatatt ataattgtt ataattgtt tataagcaag
 120841 tatatttat tataagcata attgtaagta gaagtaacti tgggtttcc tgggtgctca
 120901 gacagtaaag aatctgcctg cagtacagga gaccgggtc gatccctgt ttggggaat
 120961 tccctggaga agggatggc aaccaactcc aacatgttg cctggagaat tccatggaca
 121021 gaggagcccc gaagggtgca gtccatggg ttgcaaagag ctggatacaa cagagtact
 121081 aacacatgta tataataaa ttacctata tattgtat atattataa acataticag
 121141 atattataa taattagaaa catattatac atgtattta atactgtat aaacataaat
 121201 taaaaaata atttcagcc cttggcttg ggggtgtgt tgtggacgc ttgtgctac
 121261 tttcttgaa gtggagctc cccctccaa accagcttt gaaatgact ggaaagcaat
 121321 ggaatacata agcatcagga agatagcaac agagctgtc ttctcacag aggtgtgct
 121381 tgagtgtga gcaagtcgg cagaatgtg acagattaat atagtctat aaaaatagt
 121441 tagcaaat acgaggtgc attcaagta taaagacta ctgggtctc cagttcagt
 121501 cagtcgctg gttgtgctg actctttt accccatga ccgcagcag ccaggcctcc
 121561 ctgtccatc ccaactcctg gagtcactc aaactcatg ccatcgagc ggtgatgcca
 121621 tccaaccatc tcatcctg gcgtccctt ctctccac ctcaatct tccagcatc
 121681 aggtcttcc cagtgagtc agttcttgc atcagggtgc cagagtagt gagttcagc
 121741 ttcagcatc gtcttcaa tgaattct ggactgatt ccttaggat tgactggtg
 121801 gatctcttg cagtcaagg gactcctag agtctctc aacagcacag tctatgaata
 121861 gaatagcaaa tgaatagaga ataactta cgaggatata ttaccatt gcataaata
 121921 tatcagctg tagagaacag actgttccc aggggagagg gtgggtagg atggagtggg
 121981 agtttngat cancagaagc gagctgtat atagaagatg gataaaagg atacacaaca
 122041 atgtctact gtgtggcacc gggacctata ttcagtagt tgtgagaaac cataatcac
 122101 aagactgagg aaaagtatat atatgtat gtactgagt tgccttctg tacagaagaa
 122161 attaacaca cattgtaat cgatattca atagaatcca ccccccaaa tatataagt
 122221 tcttgagat ggagacggca acccactcca ttcttgac ccaatattc tgcctggagg
 122281 atcccatgga tagaggatc caagactc gacataacc agcgactaac acttccctt

122341 tcaaatgtgt aggtttacta gctgaatct acagagatgc ccaagacatt cgtttatgag
 122401 gaaaactcca cacgcagctt cactgagaat tattaaacct attaaaggga gagagcgcca
 122461 ggatattcat ggattgaaag attcgatgtg gtcaagtgc cagttttccc caaactgatt
 122521 ggtaaattcc ccaggagctg gctcaaggcg caaaattccc ttacctttt ttaagagac
 122581 gaagccaagg agccgattct ggttgagaga cgctcaggtc ctctgcggg agagcagccc
 122641 tcttctccc ggtegcctgg gcagtttga ggccacgacc agaaggactt ggctccctgt
 122701 gtcgcgcact cagaagtctc cctctccgtc ccaaggactc agaagctggg cgctctgccc
 122761 gcagcagagg aggcagcctg gagggggccc gcgggcacag cggtcgggt ttacgcccag
 122821 ttccccccc cgccctcta cctgggcgt gccggccggc tccggggccg gccgtgccct
 122881 ccgtggccgc aaggcgtgc tctcccccg ctggaagtgc tgaccggag gaaggggccc
 122941 agacggaggg actcggagcc tccgagtac accctgggac tccgagcgt ggagcctggc
 123001 gtcacccag gcaggggcag tggggggccc gggcggggtc aggggcctcc cccggttctc
 123061 attgacacc gcgggggtgc gctgggcaca gtgtccagg gccacgttc gagcaggggc
 123121 gcgatgcagg cccgggcgc gctgtccc ggccgcagtc cagctgctt gcagaggtg
 123181 cggcaggtc cagtgacct cacagagacg cccactctg cggtccagg tgggcctgtg
 123241 cccccagaa gtgtgacct gtgcaccgg aaggcacagg gccccagc catgtctgcg
 123301 atggaagagc cgaaccgc ccatgccgt cctcgtgac cggcaggac ccgctgtg
 123361 tccacacgt gagccatctg gctcccttg ctgacatac acccaggacc tgagtgtga
 123421 ggaagttaga aggggcaggt gtgtgacac gatgccatc agcatcact gagaacctg
 123481 acaacctca gggggccagc ctgtctgtg agggcccgag ggccggcccc tccccggacc
 123541 cctgcctga atccggcac actgcccgc ttctgtctc tgcggttgt cagacagcc
 123601 tgagcccagg gctgtgac tegetgtccc ttctgcagg actgtctc cccaggtct
 123661 tctggggct cccctctc attcgggggt ggctctct gttagtggtc ttagctgtc
 123721 ccagtcttg caacccatg gactgcagca cccaggctt cctgtctt cactagctc
 123781 tggagttgc tcaactcat gtccattgag ttagtgatc tatcaacca tctatctt
 123841 tctgcccac ttcttctc gctctaatc ttcccagca tcagggtct ttcaatgag
 123901 ttagctctc gcatcaggag gccaaagtat tggagttca gcatcagtc ttccagtga
 123961 tatgcagggt tgatttcc tagaattgac tggttgac tcttctct cagagaact
 124021 ctcaagagtc ttccagca ccacagtcg agagcatcag ttctcagtg atcaggttc
 124081 tttagccc agctctcaca tggtagatg actattgaa aacccatagc ttgattaga
 124141 tggacctca ttggcaaagt gatgggctt cattggcct gtttttaac acaccatca
 124201 ggttgtctg agcttctt ccaaagagca aacatcttt aatttctg ctgcagtaac
 124261 catccatag gatttggag cccaagaaa taaatctgc cactgttcc acttttccc
 124321 ctctattg ctatgaagt aggggactg atgcatgat cttagttta accagcagt
 124381 gtcacccga ccgttctt tctaaagag ctcacacac ctcacttg aatgcaatg
 124441 gttgctgtc cgctgttc acctctggg acttctgtc aggtctgt ctctaggcc
 124501 cctgccgtat cccagggcc cagagcagtg ctgggcttg agtcgatca gggactatg
 124561 gtgtggactg gatgtgctt gcttctctg gggaacgaga gacctggcc tggggaacga
 124621 ggggacctg tgtaccgga tctctctc cgggagagga gccaagcag tggacacagg
 124681 tcagtgtgtc ttctctctg gtggcaggt tccgtctgt gctgtcatc ttggcattc
 124741 ggtgttctg tgaaccagc cctccccct ctgataccc atccatcag cacagaggag
 124801 actgggctg gggactctt ggtctgaga ttctctcc catgtgact cccctctg
 124861 gggggagcag gcaccgtgtg tgaggagggt ggaagcttt caagacccc agctttctg
 124921 tccaggggg ctctggcagg gccttgggag ctggaatgag ctggaatct ggccagtggg
 124981 ggttccctg gtgtaaaga acccgctgc ccatgcagc ggcataagag acgcgggtc
 125041 gatcactgg tgggaagat cccctacagg agggcatggc aaccactcc agtattctt
 125101 cctgaagaat ccttggaca gaggagcctg gtgggtaca gtcttgggg tggcaaggag
 125161 tggacacga ctgaagcag ttaccatgca cgcacgcgg gtcaggggtc agggccgcgc
 125221 tcttacctg ctgttgacc ttagccaggt cacaccccc aggtgtgaa agagaacagt
 125281 ctcccagac tgggcatcc aggttttac agacgtcct gtgagcttg ttagctggc
 125341 tctgtggcg ctagaggcg ctgtccgccc ggccctatgt gcgtgcacgc atgtgagcat
 125401 gtgcatac gtgtgtcat ctgtcgggg cgcacgggtc ggggacacgg gcacgcggtc
 125461 aggaacgcag cccggacac tccagtggc ccgcgagtac cgtcaggtg gggctgtggc
 125521 tccgtgtgt ggtgacccg cctcccccc gcgaacgtg tgcatagta ccgctggct
 125581 gggctctga gctcagcat cctgcccc gggtcagtc ccgacagcc cagctctagg
 125641 cccaggcgt ggaccgagg cccaggccc cggcctgtga gatgggacct cgtctggg

	125701 ggctcattct gtccecgag gcctggcagg ccctcctct ttggcattgc ataccctgc 125761 attggggtgg gtaagcacag taccatgc ctgtggcccc gtgggagcgg cctgctcagg 125821 gagggcgag cctcagctac agggctgtca caccgggctg cagaggaaga agacgggagc 125881 gaggcctaca ggaacctagc caggccctgg cccactgagc cgacaggagc ctggccagag 125941 gcctgcacag gacggggtgg cgggggggggt ggggtggggt gctgggcccc gtggccttga 126001 ctgcagaccc cgagggtcc tcagcttaga acggccaagc ctgagtcttg ggggtgcagg 126061 tcagggggg

Primers

In another embodiment, primers are provided to generate 3' and 5' sequences of a targeting vector. The oligonucleotide primers can be capable of hybridizing to porcine immunoglobulin genomic sequence, such as Seq ID Nos. 1, 4, 29, 30, 12, 25, 15, 16, 19, 28 or 31, as described above. In a particular embodiment, the primers hybridize under stringent conditions to Seq ID Nos. 1, 4, 29, 30, 12, 25, 15, 16, 19, 28 or 31, as described above. Another embodiment provides oligonucleotide probes capable of hybridizing to porcine heavy chain, kappa light chain or lambda light chain nucleic acid sequences, such as Seq ID Nos. 1, 4, 29, 30, 12, 25, 15, 16, 19, 28 or 31, as described above. The polynucleotide primers or probes can have at least 14 bases, 20 bases, 30 bases, or 50 bases which hybridize to a polynucleotide of the present invention. The probe or primer can be at least 14 nucleotides in length, and in a particular embodiment, are at least 15, 20, 25, 28, or 30 nucleotides in length.

In one embodiment, primers are provided to amplify a fragment of porcine Ig heavy-chain that includes the functional joining region (the J6 region). In one non-limiting embodiment, the amplified fragment of heavy chain can be represented by Seq ID No 4 and the primers used to amplify this fragment can be complementary to a portion of the J-region, such as, but not limited to Seq ID No 2, to produce the 5' recombination arm and complementary to a portion of Ig heavy-chain mu constant region, such as, but not limited to Seq ID No 3, to produce the 3' recombination arm. In another embodiment, regions of the porcine Ig heavy chain (such as, but not limited to Seq ID No 4) can be subcloned and assembled into a targeting vector.

In other embodiments, primers are provided to amplify a fragment of porcine Ig kappa light-chain that includes the constant region. In another embodiment, primers are provided to amplify a fragment of porcine Ig kappa light-chain that includes the J region. In one non-limiting embodiment, the primers used to amplify this fragment can be complementary to a portion of the J-region, such as, but not limited to Seq ID No 21 or 10, to produce the 5'

recombination arm and complementary to genomic sequence 3' of the constant region, such as, but not limited to Seq ID No 14, 24 or 18, to produce the 3' recombination arm. In another embodiment, regions of the porcine Ig heavy chain (such as, but not limited to Seq ID No 20) can be subcloned and assembled into a targeting vector.

II. Genetic Targeting of the Immunoglobulin Genes

The present invention provides cells that have been genetically modified to inactivate immunoglobulin genes, for example, immunoglobulin genes described above. Animal cells that can be genetically modified can be obtained from a variety of different organs and tissues such as, but not limited to, skin, mesenchyme, lung, pancreas, heart, intestine, stomach, bladder, blood vessels, kidney, urethra, reproductive organs, and a disaggregated preparation of a whole or part of an embryo, fetus, or adult animal. In one embodiment of the invention, cells can be selected from the group consisting of, but not limited to, epithelial cells, fibroblast cells, neural cells, keratinocytes, hematopoietic cells, melanocytes, chondrocytes, lymphocytes (B and T), macrophages, monocytes, mononuclear cells, cardiac muscle cells, other muscle cells, granulosa cells, cumulus cells, epidermal cells, endothelial cells, Islets of Langerhans cells, blood cells, blood precursor cells, bone cells, bone precursor cells, neuronal stem cells, primordial stem cells, hepatocytes, keratinocytes, umbilical vein endothelial cells, aortic endothelial cells, microvascular endothelial cells, fibroblasts, liver stellate cells, aortic smooth muscle cells, cardiac myocytes, neurons, Kupffer cells, smooth muscle cells, Schwann cells, and epithelial cells, erythrocytes, platelets, neutrophils, lymphocytes, monocytes, eosinophils, basophils, adipocytes, chondrocytes, pancreatic islet cells, thyroid cells, parathyroid cells, parotid cells, tumor cells, glial cells, astrocytes, red blood cells, white blood cells, macrophages, epithelial cells, somatic cells, pituitary cells, adrenal cells, hair cells, bladder cells, kidney cells, retinal cells, rod cells, cone cells, heart cells, pacemaker cells, spleen cells, antigen presenting cells, memory cells, T cells, B cells, plasma cells, muscle cells, ovarian cells, uterine cells, prostate cells, vaginal epithelial cells, sperm cells, testicular cells, germ cells, egg cells, leydig cells, peritubular cells, sertoli cells, lutein cells, cervical cells, endometrial cells, mammary cells, follicle cells, mucous cells, ciliated cells, nonkeratinized epithelial cells, keratinized epithelial cells, lung cells, goblet cells, columnar epithelial cells, squamous epithelial cells, osteocytes,

osteoblasts, and osteoclasts. In one alternative embodiment, embryonic stem cells can be used. An embryonic stem cell line can be employed or embryonic stem cells can be obtained freshly from a host, such as a porcine animal. The cells can be grown on an appropriate fibroblast-feeder layer or grown in the presence of leukemia inhibiting factor (LIF).

In a particular embodiment, the cells can be fibroblasts; in one specific embodiment, the cells can be fetal fibroblasts. Fibroblast cells are a suitable somatic cell type because they can be obtained from developing fetuses and adult animals in large quantities. These cells can be easily propagated in vitro with a rapid doubling time and can be clonally propagated for use in gene targeting procedures.

Targeting constructs

Homologous Recombination

In one embodiment, immunoglobulin genes can be genetically targeted in cells through homologous recombination. Homologous recombination permits site-specific modifications in endogenous genes and thus novel alterations can be engineered into the genome. In homologous recombination, the incoming DNA interacts with and integrates into a site in the genome that contains a substantially homologous DNA sequence. In non-homologous ("random" or "illicit") integration, the incoming DNA is not found at a homologous sequence in the genome but integrates elsewhere, at one of a large number of potential locations. In general, studies with higher eukaryotic cells have revealed that the frequency of homologous recombination is far less than the frequency of random integration. The ratio of these frequencies has direct implications for "gene targeting" which depends on integration via homologous recombination (i.e. recombination between the exogenous "targeting DNA" and the corresponding "target DNA" in the genome).

A number of papers describe the use of homologous recombination in mammalian cells. Illustrative of these papers are Kucherlapati et al., Proc. Natl. Acad. Sci. USA 81:3153-3157, 1984; Kucherlapati et al., Mol. Cell. Bio. 5:714-720, 1985; Smithies et al, Nature 317:230-234, 1985; Wake et al., Mol. Cell. Bio. 8:2080-2089, 1985; Ayares et al., Genetics 111:375-388, 1985; Ayares et al., Mol. Cell. Bio. 7:1656-1662, 1986; Song et al., Proc. Natl. Acad. Sci. USA 84:6820-6824, 1987; Thomas et al. Cell 44:419-428, 1986; Thomas and Capecchi, Cell 51: 503-512, 1987; Nandi et al., Proc. Natl. Acad. Sci. USA 85:3845-3849, 1988; and Mansour et al.,

Nature 336:348-352, 1988. Evans and Kaufman, Nature 294:146-154, 1981; Doetschman et al., Nature 330:576-578, 1987; Thoma and Capecchi, Cell 51:503-512, 1987; Thompson et al., Cell 56:316-321, 1989.

The present invention can use homologous recombination to inactivate an immunoglobulin gene in cells, such as the cells described above. The DNA can comprise at least a portion of the gene(s) at the particular locus with introduction of an alteration into at least one, optionally both copies, of the native gene(s), so as to prevent expression of functional immunoglobulin. The alteration can be an insertion, deletion, replacement or combination thereof. When the alteration is introduced into only one copy of the gene being inactivated, the cells having a single unmutated copy of the target gene are amplified and can be subjected to a second targeting step, where the alteration can be the same or different from the first alteration, usually different, and where a deletion, or replacement is involved, can be overlapping at least a portion of the alteration originally introduced. In this second targeting step, a targeting vector with the same arms of homology, but containing a different mammalian selectable marker can be used. The resulting transformants are screened for the absence of a functional target antigen and the DNA of the cell can be further screened to ensure the absence of a wild-type target gene. Alternatively, homozygosity as to a phenotype can be achieved by breeding hosts heterozygous for the mutation.

Targeting Vectors

In another embodiment, nucleic acid targeting vector constructs are also provided. The targeting vectors can be designed to accomplish homologous recombination in cells. These targeting vectors can be transformed into mammalian cells to target the ungulate heavy chain, kappa light chain or lambda light chain genes via homologous recombination. In one embodiment, the targeting vectors can contain a 3' recombination arm and a 5' recombination arm (i.e. flanking sequence) that is homologous to the genomic sequence of ungulate heavy chain, kappa light chain or lambda light chain genomic sequence, for example, sequence represented by Seq ID Nos. 1, 4, 29, 30, 12, 25, 15, 16, 19, 28 or 31, as described above. The homologous DNA sequence can include at least 15 bp, 20 bp, 25 bp, 50 bp, 100 bp, 500 bp, 1 kbp, 2 kbp, 4 kbp, 5 kbp, 10 kbp, 15 kbp, 20 kbp, or 50 kbp of sequence, particularly contiguous sequence, homologous to the genomic sequence. The 3' and 5' recombination arms

can be designed such that they flank the 3' and 5' ends of at least one functional variable, joining, diversity, and/or constant region of the genomic sequence. The targeting of a functional region can render it inactive, which results in the inability of the cell to produce functional immunoglobulin molecules. In another embodiment, the homologous DNA sequence can include one or more intron and/or exon sequences. In addition to the nucleic acid sequences, the expression vector can contain selectable marker sequences, such as, for example, enhanced Green Fluorescent Protein (eGFP) gene sequences, initiation and/or enhancer sequences, poly A-tail sequences, and/or nucleic acid sequences that provide for the expression of the construct in prokaryotic and/or eukaryotic host cells. The selectable marker can be located between the 5' and 3' recombination arm sequence.

Modification of a targeted locus of a cell can be produced by introducing DNA into the cells, where the DNA has homology to the target locus and includes a marker gene, allowing for selection of cells comprising the integrated construct. The homologous DNA in the target vector will recombine with the chromosomal DNA at the target locus. The marker gene can be flanked on both sides by homologous DNA sequences, a 3' recombination arm and a 5' recombination arm. Methods for the construction of targeting vectors have been described in the art, see, for example, Dai et al., *Nature Biotechnology* 20: 251-255, 2002; WO 00/51424.

Various constructs can be prepared for homologous recombination at a target locus. The construct can include at least 50 bp, 100 bp, 500 bp, 1kbp, 2 kbp, 4 kbp, 5 kbp, 10 kbp, 15 kbp, 20 kbp, or 50 kbp of sequence homologous with the target locus. The sequence can include any contiguous sequence of an immunoglobulin gene.

Various considerations can be involved in determining the extent of homology of target DNA sequences, such as, for example, the size of the target locus, availability of sequences, relative efficiency of double cross-over events at the target locus and the similarity of the target sequence with other sequences.

The targeting DNA can include a sequence in which DNA substantially isogenic flanks the desired sequence modifications with a corresponding target sequence in the genome to be modified. The substantially isogenic sequence can be at least about 95%, 97-98%, 99.0-99.5%, 99.6-99.9%, or 100% identical to the corresponding target sequence (except for the desired sequence modifications). In a particular embodiment, the targeting DNA and the target DNA can share stretches of DNA at least about 75, 150 or 500 base pairs that are 100% identical.

Accordingly, targeting DNA can be derived from cells closely related to the cell line being targeted; or the targeting DNA can be derived from cells of the same cell line or animal as the cells being targeted.

Porcine Heavy Chain Targeting

In particular embodiments of the present invention, targeting vectors are provided to target the porcine heavy chain locus. In one particular embodiment, the targeting vector can contain 5' and 3' recombination arms that contain homologous sequence to the 3' and 5' flanking sequence of the J6 region of the porcine immunoglobulin heavy chain locus. Since the J6 region is the only functional joining region of the porcine immunoglobulin heavy chain locus, this will prevent the expression of a functional porcine heavy chain immunoglobulin. In a specific embodiment, the targeting vector can contain a 5' recombination arm that contains sequence homologous to genomic sequence 5' of the J6 region, optionally including J1-4 and a 3' recombination arm that contains sequence homologous to genomic sequence 3' of the J6 region, including the mu constant region (a "J6 targeting construct"), see for example, Figure 1. Further, this J6 targeting construct can also contain a selectable marker gene that is located between the 5' and 3' recombination arms, see for example, Seq ID No 5 and Figure 1. In other particular embodiments, the 5' targeting arm can contain sequence 5' of J1, such as depicted in Seq ID No. 1 and/ or Seq ID No 4. In another embodiment, the 5' targeting arm can contain sequence 5' of J1, J2 and/ or J3, for example, as depicted in approximately residues 1-300, 1-500, 1-750, 1-1000 and/ or 1-1500 Seq ID No 4. In a further embodiment, the 5' targeting arm can contain sequence 5' of the constant region, for example, as depicted in approximately residues 1-300, 1-500, 1-750, 1-1000, 1-1500 and/ or 1-2000 or any fragment thereof of Seq ID No 4 and/ or any contiguous sequence of Seq ID No. 4 or fragment thereof. In another embodiment, the 3' targeting arm can contain sequence 3' of the constant region and/ or including the constant region, for example, such as residues 7000-8000 and/ or 8000-9000 or fragment thereof of Seq ID No 4. In other embodiments, targeting vector can contain any contiguous sequence or fragment thereof of Seq ID No 4. In other embodiments, the targeting vector can contain a 5' recombination arm that contains sequence homologous to genomic sequence 5' of the diversity region, and a 3' recombination arm that contains sequence homologous to genomic sequence 3' of the diversity region of the porcine heavy chain locus. In a further embodiment, the targeting

vector can contain a 5' recombination arm that contains sequence homologous to genomic sequence 5' of the mu constant region and a 3' recombination arm that contains sequence homologous to genomic sequence 3' of the mu constant region of the porcine heavy chain locus.

In further embodiments, the targeting vector can include, but is not limited to any of the following sequences: the Diversity region of heavy chain is represented, for example, by residues 1089-1099 of Seq ID No 29 (D(pseudo)), the Joining region of heavy chain is represented, for example, by residues 1887-3352 of Seq ID No 29 (for example: J(pseudo): 1887-1931 of Seq ID No 29, J(pseudo): 2364-2411 of Seq ID No 29, J(pseudo): 2756-2804 of Seq ID No 29, J (functional J): 3296-3352 of Seq ID No 29), the recombination signals are represented, for example, by residues 3001-3261 of Seq ID No 29 (Nonamer), 3292-3298 of Seq ID No 29 (Heptamer), the Constant Region is represented by the following residues: 3353-9070 of Seq ID No 29 (J to C mu intron), 5522-8700 of Seq ID No 29 (Switch region), 9071-9388 of Seq ID No 29 (Mu Exon 1), 9389-9469 of Seq ID No 29 (Mu Intron A), 9470-9802 of Seq ID No 29 (Mu Exon 2), 9830-10069 of Seq ID No 29 (Mu Intron B), 10070-10387 of Seq ID No 29 (Mu Exon 3), 10388-10517 of Seq ID No 29 (Mu Intron C), 10815-11052 of Seq ID No 29 (Mu Exon 4), 11034-11039 of Seq ID No 29 (Poly(A) signal) or any fragment or combination thereof. Still further, any contiguous sequence at least about 17, 20, 30, 40, 50, 100, 150, 200 or 300 nucleotides of Seq ID No 29 or fragment and/ or combination thereof can be used as targeting sequence for the heavy chain targeting vector. It is understood that in general when designing a targeting construct one targeting arm will be 5' of the other targeting arm.

In other embodiments, targeting vectors designed to disrupt the expression of porcine heavy chain genes can contain recombination arms, for example, the 3' or 5' recombination arm, that target the constant region of heavy chain. In one embodiment, the recombination arm can target the mu constant region, for example, the C mu sequences described above or as disclosed in Sun & Butler Immunogenetics (1997) 46: 452-460. In another embodiment, the recombination arm can target the delta constant region, such as the sequence disclosed in Zhao et al. (2003) J Immunol 171: 1312-1318, or the alpha constant region, such as the sequence disclosed in Brown & Butler (1994) Molec Immunol 31: 633-642.

Seq ID No.5	GGCCAGACTTCCTCGGAACAGCTCAAAGAGCTCTGTCAAAGCCAGATCCC ATCACACGTGGGCACCAATAGGCCATGCCAGCCTCCAAGGGCCGAACTGG GTTCTCCACGGCGCACATGAAGCCTGCAGCCTGGCTTATCCTCTTCCGTG GTGAAGAGGCAGGCCCGGGACTGGACGAGGGGCTAGCAGGGTGTGGTAGG
-------------	--

CACCTTGCGCCCCCACCCTGGCAGGAACCAGAGACCCTGGGGCTGAGAG
TGAGCCTCCAAACAGGATGCCCCACCCTTCAGGCCACCTTTCAATCCAGC
TACACTCCACCTGCCATTCTCCTCTGGGCACAGGGCCCAGCCCCTGGATC
TTGGCCTTGGCTCGACTTGCACCCACGCGCACACACACACTTCCTAACGT
GCTGTCCGCTCACCCCTCCCCAGCGTGGTCCATGGGCAGCACGGCAGTGC
GCGTCCGGCGGTAGTGAGTGCAGAGGTCCCTTCCCCTCCCCCAGGAGCCC
CAGGGGTGTGTGCAGATCTGGGGGCTCCTGTCCCTTACACCTTCATGCCC
CTCCCCTCATACCCACCCTCCAGGCGGGAGGCAGCGAGACCTTTGCCCAG
GGACTCAGCCAACGGGCACACGGGAGGCCAGCCCTCAGCAGCTGGCTCCC
AAAGAGGAGGTGGGAGGTAGGTCCACAGCTGCCACAGAGAGAAACCCTGA
CGGACCCACAGGGGCCACGCCAGCCGGAACCAGCTCCCTCGTGGGTGAG
CAATGGCCAGGGCCCCGCCGCCACCACGGCTGGCCTTGCGCCAGCTGAG
AACTCACGTCCAGTGCAGGGAGACTCAAGACAGCCTGTGCACACAGCCTC
GGATCTGCTCCCATTTCAAGCAGAAAAAGGAAACCGTGCAGGCAGCCCTC
AGCATTTCAAGGATTGTAGCAGCGGCCAACTATTCGTCGGCAGTGGCCGA
TTAGAATGACCGTGGAGAAGGGCGGAAGGGTCTCTCGTGGGCTCTGCGGC
CAACAGGCCCTGGCTCCACCTGCCCGCTGCCAGCCCGAGGGGCTTGGGCC
GAGCCAGGAACCACAGTGCTCACCGGGACCACAGTGACTGACCAAACCTCC
CGGCCAGAGCAGCCCCAGGCCAGCCGGGCTCTCGCCCTGGAGGACTCACC
ATCAGATGCACAAGGGGGCGAGTGTGGAAGAGACGTGTGCCCCGGGCCAT
TTGGGAAGGCGAAGGGACCTTCCAGGTGGACAGGAGGTGGGACGCACTCC
AGGCAAGGGACTGGGTCCCCAAGGCCTGGGGAAGGGGTACTGGCTTGGGG
GTTAGCCTGGCCAGGGAACGGGGAGCGGGGCGGGGGGCTGAGCAGGGAGG
ACCTGACCTCGTGGGAGCGAGGCAAGTCAGGCTTCAGGCAGCAGCCGCAC
ATCCCAGACCAGGAGGCTGAGGCAGGAGGGGCTTGCAGCGGGGCGGGGGC
CTGCCTGGCTCCGGGGGCTCCTGGGGGACGCTGGCTCTTGTTTCCGTGTC
CCGCAGCACAGGGGCCAGCTCGCTGGGCCTATGCTTACCTTGATGTCTGGG
GCCGGGGCGTCAAGGTCGTCTCCTCAGGGGAGAGTCCCCTGAGGCTA
CGCTGGGG*GGGGACTATGGCAGCTCCACCAGGGGCCTGGGGACCAGGGG
CCTGGACCAGGCTGCAGCCCGGAGGACGGGCAGGGCTCTGGCTCTCCAGC
ATCTGGCCCTCGGAAATGGCAGAACCCCTGGCGGGTGAGCGAGCTGAGAG
CGGGTCAGACAGACAGGGGGCCGGCCGGAAGGAGAAGTTGGGGGCAGAGC
CCGCCAGGGGCCAGGCCCAAGGTTCTGTGTGCCAGGGCCTGGGTGGGCAC
ATTGGTGTGGCCATGGCTACTTAGACGCGTGATCAAGGGCGAATTCCAGC
ACACTGGCGGCCGTTACTAGTggatcccgccgcccctaccgggtagggg
aggcgcttttccaagcagctctggagcatgcgctttagcagccccgctg
ggcacttggcgctacacaagtggcctctggcctcgcacacattccacac
caccggtaggcgccaaccggctccgttcttgggtggccccctcgcgccac
cttctactcctccccctagtcaggaagtccccccccgccccgcagctcgcg
tcgtgcaggacgtgacaaatggaagtagcacgtctactagtctcgtgca
gatggacagcaccgctgagcaatggaagcgggtaggccttggggcagcg
gccaatagcagcttggctccttcgcttctgggctcagaggctgggaag
gggtgggtccggggcgggctcagggcgggctcagggcgggcgggcg
cccgaaggtctccggaagcccggcattctgcagcttcaaaagcgcacg
tctgccgcgctgttctctctctctcctcctcctcgggcttgcacctgag
ccaatatgggatcgccattgaacaagatggattgcacgcaggtctccg
gccgcttgggtggagaggctattcggtatgactgggcacacagacaat
cggtgctctgatgcccggtgttcgggtgtcagcgagggcgccccg
ttcttttgaagaccgacctgtccggtgccctgaatgaactgcaggac
gaggcagcgcggtatcgtggctggccacgacggcggttcttgcgcagc
tgtgtcgcaggtgtactgaagcgggaaggactggctgctattgggcg
aagtccggggcaggatctctgtcatctcacctgtcctgccgagaaa
gtatccatcatggctgatgcaatgcggcggtgcatacgcttgatccggc
tacctgcccattcgaccaccaagcgaaacatcgcatcgagcgagcacgta
ctcggtggaagccggtcttgtcaatcaggatgatctggacgaagagcat
caggggctcgcgccagccgaactgttgcgaggtcaaggcgcgcatgcc

cgacggcgaggatctcgtcgtgacccatggcgatgcctgcttgcgaata
tcatgggtgaaaatggccgctttctggattcatcgactgtggccggctg
ggtgtggcgatcgctatcaggacatagcgttggctaccgtgatattgc
tgaagagcttggcggaatgggctgaccgcttctcgtgctttacgga
tcgccgctcccgattcgcagcgcatcgcccttatcgcccttctgacgag
ttctctgaggggatcaattcTCTAGATGCATGCTCGAGCGGCCGCCAGT
GTGATGGATATCTGCAGAAATTCGCCCTtCCAGGCGTTGAAGTCGTCTGT
CCTCAGGTAAGAACGGCCCTCCAGGGCCTTTAATTTCTGCTCTCGTCTGT
GGGCTTTTCTGACTCTGATCCTCGGGAGGCGTCTGTGCCCCCCCCGGGGA
TGAGGCCGGCTTGCCAGGAGGGGTCAGGGACCAGGAGCCTGTGGGAAGTT
CTGACGGGGGCTGCAGGCGGGAAGGGCCCCACCGGGGGGCGAGCCCCAGG
CCGCTGGGCGGCAGGAGACCCGTGAGAGTGCGCCTTGAGGAGGGTGTCTG
CGGAACCACGAACGCCCGCCGGGAAGGGCTTGCTGCAATGCGGTCTTCAG
ACGGGAGGCGTCTTCTGCCCTCACCGTCTTTCAAGCCCTTGTTGGGTCTGA
AAGAGCCATGTCCGAGAGAGAAGGGACAGGCCTGTCCCGACCTGGCCGAG
AGCGGGCAGCCCCGGGGGAGAGCGGGGCGATCGGCCTGGGCTCTGTGAGG
CCAGGTCCAAGGGAGGACGTGTGGTCCTCGTGACAGGTGCACTTGCGAAA
CCTTAGAAGACGGGGTATGTTGGAAGCGGCTCCTGATGTTTAAGAAAAGG
GAGACTGTAAAGTGAGCAGAGTCCTCAAGTGTGTTAAGGTTTTAAAGGTC
AAAGTGTTTTAAACCTTTGTGACTGCAGTTAGCAAGCGTGCGGGGAGTGA
ATGGGGTGCCAGGGTGGCCGAGAGGCAGTACGAGGGCCGTGCCGTCTCT
AATTCAGGGCTTAGTTTTGCAGAATAAAGTCGGCCTGTTTTCTAAAAGCA
TTGGTGCTGCTGAGCTGGTGGAGGAGGCCGCGGGCAGCCCTGGCCACCTG
CAGCAGGTGGCAGGAAGCAGGTCCGCCAAGAGGCTATTTTAGGAAGCCAG
AAAACACGGTCGATGAATTTATAGCTTCTGGTTTCCAGGAGGTGGTTGGG
CATGGCTTTGCGCAGCGCCACAGAACCGAAAGTGCCCACTGAGAAAAAAC
AACTCCTGCTTAATTTGCATTTTTCTAAAAGAAGAAACAGAGGCTGACGG
AACTGGAAAGTTCCTGTTTTAACTACTCGAATTGAGTTTTCGGTCTTAG
CTTATCAACTGCTCACTTAGATTCATTTTTCAAAGTAAACGTTTAAGAGCC
GAGGCATTCCTATCCTCTTCTAAGGCGTTATTCCTGGAGGCTCATTACCC
GCCAGCACCTCCGCTGCCTGCAGGCATTGCTGTCACCGTCACCGTGACGG
CGCGCACGATTTTCAGTTGGCCCCGCTTCCCCTCGTGATTAGGACAGACGC
GGGCACTCTGGCCCAGCCGTCTTGGCTCAGTATCTGCAGGCGTCCGTCTC
GGGACGGAGCTCAGGGGAAGAGCGTGACTCCAGTTGAACGTGATAGTCGG
TGCGTTGAGAGGAGACCCAGTCGGGTGTGAGTCAGAAGGGGGCCCGGGG
CCGAGGCCCTGGGCAGGACGGCCCGTGCCCTGCATCACGGGCCAGCGTC
CTAGAGGCAGGACTCTGGTGGAGAGTGTGAGGGTGCCTGGGGCCCCCTCCG
GAGCTGGGGCCGTGCGGTGCAGGTTGGGCTCTCGGCGCGGTGTTGGCTGT
TTCTGCGGGATTTGGAGGAATTCTTCCAGTGATGGGAGTCGCCAGTGACC
GGGCACCAGGCTGGTAAGAGGGAGGCCGCCGTGCTGGCCAGAGCAGCTGG
GAGGGTTCGGTAAAAGGCTCGCCCCGTTTCCTTTAATGAGGACTTTTCCTG
GAGGGCATTTAGTCTAGTCGGGACCGTTTTTCGACTCGGGAAGAGGGATGC
GGAGGAGGGCATGTGCCCAGGAGCCGAAGGCGCCGCGGGGAGAAGCCCAG
GGCTCTCCTGTCCCCACAGAGGCGACGCCACTGCCGCAGACAGAGGGC
CTTCCCTCTGATGACGGCAAAGGCGCCTCGGCTCTTGCGGGGTGCTGGG
GGGGAGTCGCCCCGAAGCCGCTCACCCAGAGGCCTGAGGGGTGAGACTGA
CCGATGCCTCTTGGCCGGGCCTGGGGCCGGACCGAGGGGGACTCCGTGGA
GGCAGGGCGATGGTGGCTGCGGGAGGGAACCGACCCTGGGCCGAGCCCGG
CTTGGCGATTCCCGGGGCGAGGGCCCTCAGCCGAGGCGAGTGGGTCCGGCG
GAACCAACCCTTTCTGGCCAGCGCCACAGGGCTCTCGGGACTGTCCGGGGC
GACGCTGGGCTGCCCGTGGCAGGCCTGGGCTGACCTGGACTTCACCAGAC
AGAACAGGGCTTTCAGGGCTGAGCTGAGCCAGGTTTAGCGAGGCCAAGTG
GGGCTGAACCAGGCTCAACTGGCCTGAGCTGGGTTGAGCTGGGCTGACCT
GGGCTGAGCTGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT
GGACTGGCTGAGCTGAGCTGGGTTGAGCTGAGCTGAGCTGGCCTGGGTTG
AGCTGGGCTGGGTTGAGCTGAGCTGGGTTGAGCTGGGTTGAGCTGGGTTG

ATCTGAGCTGAGCTGGGCTGAGCTGAGCTAGGCTGGGGTGAGCTGGGCTG
AGCTGGTTTGAGTTGGGTTGAGCTGAGCTGAGCTGGGCTGTGCTGGCTGA
GCTAGGCTGAGCTAGGCTAGGTTGAGCTGGGCTGGGCTGAGCTGAGCTAG
GCTGGGCTGATTTGGGCTGAGCTGAGCTGAGCTAGGCTGCGTTGAGCTGG
CTGGGCTGGATTGAGCTGGCTGAGCTGGCTGAGCTGGGCTGAGCTGGCCT
GGGTTGAGCTGAGCTGGACTGGTTTGAGCTGGGTCGATCTGGGTTGAGCT
GTCCTGGGTTGAGCTGGGCTGGGTTGAGCTGAGCTGGGTTGAGCTGGGCT
CAGCAGAGCTGGGTTGGGCTGAGCTGGGTTGAGCTGAGCTGGGCTGAGCT
GGCCTGGGTTGAGCTGGGCTGAGCTGAGCTGGGCTGAGCTGGCCTGTGTT
GAGCTGGGCTGGGTTGAGCTGGGCTGAGCTGGATTGAGCTGGGTTGAGCT
GAGCTGGGCTGGGCTGTGCTGACTGAGCTGGGCTGAGCTAGGCTGGGGTG
AGCTGGGCTGAGCTGATCCGAGCTAGGCTGGGCTGGTTTGGGCTGAGCTG
AGCTGAGCTAGGCTGGATTGATCTGGCTGAGCTGGGTTGAGCTGAGCTGG
GCTGAGCTGGTCTGAGCTGGCCTGGGTCGAGCTGAGCTGGACTGGTTTGA
GCTGGGTCGATCTGGGCTGAGCTGGCCTGGGTTGAGCTGGGCTGGGTTGA
GCTGAGCTGGGTTGAGCTGGGCTGAGCTGAGGGCTGGGGTGAGCTGGGCT
GAACTAGCCTAGCTAGGTTGGGCTGAGCTGGGCTGGTTTGGGCTGAGCTG
AGCTGAGCTAGGCTGCATTGAGCAGGCTGAGCTGGGCTGAGCAGGCCTGG
GGTGAGCTGGGCTAGGTGGAGCTGAGCTGGGTCGAGCTGAGTTGGGCTGA
GCTGGCCTGGGTTGAGGTAGGCTGAGCTGAGCTGAGCTAGGCTGGGTTGA
GCTGGCTGGGCTGGTTTGCCTGGGTCAGCTGGGCTGAGCTGGCCTGGG
TTGAGCTGGGCTCGGTTGAGCTGGGCTGAGCTGAGCCGACCTAGGCTGGG
ATGAGCTGGGCTGATTTGGGCTGAGCTGAGCTGAGCTAGGCTGCATTGAG
CAGGCTGAGCTGGGCCTGGAGCCTGGCCTGGGGTGAGCTGGGCTGAGCTG
CGCTGAGCTAGGCTGGGTTGAGCTGGCTGGGCTGGTTTGCCTGGGTCAA
GCTGGGCTGAGCTGGCCTGGGATGAGCTGGGCTGGGCTGGGCTGAGCTGA
GCTGAGCTAGGCTGCATTGAGCAGGCTGAGCTGGGCTGAGCTGGCCTGGG
GTGAGCTGGGCTGAGCTAAGCTGAGCTGGGCTGGTTTGGGCTGAGCTGGC
TGAGCTGGGTCCTGCTGAGCTGGGCTGAGCTGACCAGGGGTGAGCTGGGC
TGAGTTAGGCTGGGCTCAGCTAGGCTGGGTTGATCTGGCAGGGCTGGTTT
GCGCTGGGTCAAGCTCCCGGGAGATGGCCTGGGATGAGCTGGGCTGGTTT
GGGCTGAGCTGAGCTGAGCTGAGCTAGGCTGCATTGAGCAGGCTGAGCTG
GGCTGAGCTGGCCTGGGGTGAGCTGGGCTGGGTGGAGCTGAGCTGGGCTG
AACTGGGCTAAGCTGGCTGAGCTGGATCGAGCTGAGCTGGGCTGAGCTGG
CCTGGGGTTAGCTGGGCTGAGCTGAGCTGAGCTAGGCTGGGTTGAGCTGG
CTGGGCTGGTTTGCCTGGGTCAGCTGGGCTGGGCTGGGCTGGGCTGAG
CTGGGCTGGGCTGAGCTGAGCTAGGCTGGGTTGAGCTGGGCTGGGCTGAG
CTGAGCTAGGCTGCATTGAGCTGGCTGGGATGGATTGAGCTGGCTGAGCT
GGCTGAGCTGGCTGAGCTGGGCTGAGCTGGCCTGGGTTGAGCTGGGCTGG
GTTGAGCTGAGCTGGGCTGAGCTGGGCTCAGCAGAGCTGGGTTGAGCTGA
GCTGGGTTGAGCTGGGGTGAGCTGGGCTGAGCAGAGCTGGGTTGAGCTGA
GCTGGGTTGAGCTGGGCTCGAGCAGAGCTGGGTTGAGCTGAGCTGGGTTG
AGCTGGGCTCAGCAGAGCTGGGTTGAGCTGAGCTGGGTTGAGCTGGGCTG
AGCTAGCTGGGCTCAGCTAGGCTGGGTTGAGCTGAGCTGGGCTGAACTGG
GCTGAGCTGGGCTGAACTGGGCTGAGCTGGGCTGAGCTGGGCTGAGCAGA
GCTGGGCTGAGCAGAGCTGGGTTGGTCTGAGCTGGGTTGAGCTGGGCTGA
GCTGGGCTGAGCAGAGTTGGGTTGAGCTGAGCTGGGTTGAGCTGGGCTGA
GCTAGGCTGGGTTGAGCTGGGTTGAGTTGGGCTGAGCTGGGCTGGGTTGA
GCGGAGCTGGGCTGAACTGGGCTGAGCTGGGCTGAGCGGAACTGGGTTGA
TCTGAATTGAGCTGGGCTGAGCCGGGCTGAGCCGGGCTGAGCTGGGCTAG
GTTGAGCTTGGGTGAGCTTGCCTCAGCTGGTCTGAGCTAGGTTGGGTGGA
GCTAGGCTGGATTGAGCTGGGCTGAGCTGAGCTGATCTGGCCTCAGCTGG
GCTGAGGTAGGCTGAACTGGGCTGTGCTGGGCTGAGCTGAGCTGAGCCAG
TTTGAGCTGGGTTGAGCTGGGCTGAGCTGGGCTGTGTTGATCTTTCCTGA
ACTGGGCTGAGCTGGGCTGAGCTGGCCTAGCTGGATTGAACGGGGGTAAG
CTGGGCCAGGCTGGACTGGGCTGAGCTGAGCTAGGCTGAGCTGAGTTGAA

	TTGGGTAAAGCTGGGCTGAGATGGGCTGAGCTGGGCTGAGCTGGGTTGAG CCAGGTCGGACTGGGTTACCCTGGGCCACACTGGGCTGAGCTGGGCGGAG CTCGATTAACTGGTCAGGCTGAGTCGGGTCCAGCAGACATGCGCTGGCC AGGCTGGCTTGACCTGGACACGTTTCGATGAGCTGCCTTGGGATGGTTAC CTCAGCTGAGCCAGGTGGCTCCAGCTGGGCTGAGCTGGTGACCCTGGGTG ACCTCGGTGACCAGGTTGTCCTGAGTCCGGGCCAAGCCGAGGCTGCATCA GACTCGCCAGACCCAAGGCCTGGGCCCGGCTGGCAAGCCAGGGGCGGTG AAGGCTGGGCTGGCAGGACTGTCCCGGAAGGAGGTGCACGTGGAGCCGCC CGGACCCCGACCGGCAGGACCTGGAAAGACGCCTCTCACTCCCCTTTCTC TTCTGTCCCCTCTCGGGTCCTCAGAGAGCCAGTCTGCCCCGAATCTCTAC CCCCTCGTCTCTGCGTCAGCCCCCGTCCGATGAGAGCCTGGTGGCCCT GGGCTGCCTGGCCCGGACTTCCTGCCAGCTCCGTCACCTTCTCCTGGAA
--	---

Porcine Kappa Chain Targeting

In particular embodiments of the present invention, targeting vectors are provided to target the porcine kappa chain locus. In one particular embodiment, the targeting vector can contain 5' and 3' recombination arms that contain homologous sequence to the 3' and 5' flanking sequence of the constant region of the porcine immunoglobulin kappa chain locus. Since the present invention discovered that there is only one constant region of the porcine immunoglobulin kappa light chain locus, this will prevent the expression of a functional porcine kappa light chain immunoglobulin. In a specific embodiment, the targeting vector can contain a 5' recombination arm that contains sequence homologous to genomic sequence 5' of the constant region, optionally including the joining region, and a 3' recombination arm that contains sequence homologous to genomic sequence 3' of the constant region, optionally including at least part of the enhancer region (a "Kappa constant targeting construct"), see for example, Figure 2. Further, this kappa constant targeting construct can also contain a selectable marker gene that is located between the 5' and 3' recombination arms, see for example, Seq ID No 20 and Figure 2. In other embodiments, the targeting vector can contain a 5' recombination arm that contains sequence homologous to genomic sequence 5' of the joining region, and a 3' recombination arm that contains sequence homologous to genomic sequence 3' of the joining region of the porcine kappa light chain locus. In other embodiments, the 5' arm of the targeting vector can include Seq ID No 12 and/ or Seq ID No 25 or any contiguous sequence or fragment thereof. In another embodiment, the 3' arm of the targeting vector can include Seq ID No 15, 16 and/ or 19 or any contiguous sequence or fragment thereof.

In further embodiments, the targeting vector can include, but is not limited to any of the following sequences: the coding region of kappa light chain is represented, for example by residues 1-549 of Seq ID No 30 and 10026-10549 of Seq ID No 30, whereas the intronic sequence is represented, for example, by residues 550-10025 of Seq ID No 30, the Joining region of kappa light chain is represented, for example, by residues 5822- 7207 of Seq ID No 30 (for example, J1:5822-5859 of Seq ID No 30, J2:6180-6218 of Seq ID No 30, J3:6486-6523 of Seq ID No 30, J4:6826-6863 of Seq ID No 30, J5:7170-7207 of Seq ID No 30), the Constant Region is represented by the following residues: 10026- 10549 of Seq ID No 30 (C exon) and 10026-10354 of Seq ID No 30 (C coding), 10524-10529 of Seq ID No 30 (Poly(A) signal) and 11160-11264 of Seq ID No 30 (SINE element) or any fragment or combination thereof. Still further, any contiguous sequence at least about 17, 20, 30, 40, 50, 100, 150, 200 or 300 nucleotides of Seq ID No 30 or fragment and/ or combination thereof can be used as targeting sequence for the heavy chain targeting vector. It is understood that in general when designing a targeting construct one targeting arm will be 5' of the other targeting arm.

Seq ID No.20	ctcaaacgtaagtggcttttccgactgattcttgcgttttctaattgt tgggtggcttttgcattttcagtggtttcatcgaattagttgtcag ggaccaaacaaattgccttccagattaggtaccagggaggggacattgc tgcatgggagaccagaggggtggctaattttaacgtttccaagccaaat aactggggaagggggttgcgtgctgtgagggtaggtttatagaagt ggaagttaaggggaaatcgctatggttcacttttgctcggggaccaag tggagcccaaaattgagtacattttccatcaattttgtgagattttg tcctgtgtgctattgtgcaagttttgacattttggttgaatgagcca ttccagggacccaaaaggatgagaccgaaaagtagaaaagagccaactt ttaagctgagcagacagaccgaattgttgagttgtgaggagagtagggt ttgtaggagaaaaggggaacagatcgctggcttttctctgaattgcct ttctcatgggactggcttcagagggggtttgatgaggggaagtgttcta gagccttaactgtgggtgtgttcggtagcgggaccaagctgaaatcaa acgtaagtgcacttttctactccttttcttcttatacgggtgtgaaat tggggacttttcatgtttggagtatgagttgaggtcagttctgaagagag tgggactcatccaaaaatctgaggagtaagggtcgaacagagttgtctc atggaagaacaaagacctagttagttgatgaggcagctaaatgagtcagt tgacttgggatccaaatggccagacttcgtctgtaaccaacaatctaag agatgtagcagcaaaaagagatttccattgaggggaaagtaaaattgta atattgtgatcaccttgggtgaaggacatccgtggagattgaacgtaa gtatttttctctactaccttctgaaattgtctaaatgccagttgtgac tttagagggttaagtgtcagttttgtgaaaaatgggtaaacaagagcat ttcatatttattatcagttcaaaagttaaactcagctccaaaaatgaat ttgtagacaaaaagattaatttaagccaaattgaatgattcaaggaaaa aaaaattagtgtagatgaaaaaggaattcttacagctccaaagagcaaaa gcgaattaattttcttgaactttgccaaatcttgtaaatgattttgtt ctttacaatttaaaaagggttagagaaatgtatttcttagtctgtttctc tcttctgtctgataaattattatatgagataaaaatgaaaattaatagga tgtgctaaaaaatcagtaagaaggttagaaaaatatagtttatgttaaag
--------------	---

ttgccacttaattgagaatcagaagcaatgtatttttaagtctaaaat
 gagagataaactgtcaatacttaattctgcagagattctatacttgac
 agatatctccttttcaaaaatccaatttctatggtagactaaattgaa
 atgatcttctcataatggagggaagatggactgaccccaaaagctca
 gattt*aagaaaacctgtttaag*gaaagaaaataaaagaactgcattt
 ttaaaggcccatgaattttagaaaaataggaaatatttaataagtga
 ttctttatttctgttattacttgatgggtttttataccgccaagga
 ggccgtggcaccgtcagtgtgatctgtagaccccatggcggccttttct
 gcgattgaatgaccttggcgggtgggtcccagggtctgtgtggcagcgca
 ccagccgctaaaagccgctaaaaactgcccgtaaaggccacagcaacccc
 gcgaccgcccgttcaactgtgctgacacagtatacagataatgtcgcta
 acagaggagaatagaaatgatcgggcacacgctaattgtggggaaaagag
 ggagaagcctgattttatttttagagattctagagataaaattccag
 tattataccttttaataaaaaatttctattaggagattataaagaattt
 aaagctatttttaagtgggtgttaattcttctcagtagtctctgtcaa
 atggatttaagtaatagaggcttaatccaaatgagagaaatagacgcata
 acccttcaaggcaaaagctacaagagcaaaaattgaacacagcagccag
 ccactagccactcagatttgcagtttactgagttgaagtaata
 tcatgaaggataaattgtgataaaaaataagatacaggtgtgacacat
 cttaagttcagaaattaatggcttcagtaggattatattcacgtat
 acaaagtatctaagcagataaaaaatgccattaatgaaacttaatagaaa
 tatatttttaaatctctcattctgtgacagaaatttctaactgtgggtc
 ttttaacacctacccttgaaagagtttagtaattgtctatttgcctc
 gctgttactccagctaatttcaaaagtatacttgagaaagattattt
 tggtttgaaccacctggcaggactattttagggccattttaaactctt
 ttcaactaagtattttaaactgttctaaaccatttagggccttttaaaa
 atctttcatgaatttcaacttcgttaaaagtatttaaggtgtctggca
 agaacttcttatcaaatatgctaatagtttaatctgttaatgcaggata
 taaaattaaagtatcaaggcttgacccaacaggagtatcttcatagca
 tatttcccctccttttttctagaattcatatgatttgcgtccaaggct
 attttatataatctctggaaaaaaatagtaatgaaggtaaaagagaag
 aaaatatcagaacattaagaattcggattttactaactgcttggttaac
 atgaagggtttttattttaaggtttctatctttataaaaatctgttcc
 ctttctgctgatttctccaagcaaaagattcttgatttgttttaact
 cttacttcccacccaaggcctgaatgcccaaaaggggacttccagga
 ggccatctggcagctgctcaccgtcagaagtgaagccagccagttcctcc
 tgggcaggtggccaaaattacagttgacccctcctggtctggtgaacct
 tgccccatattggtgacagccatctggccagggccaggtctccctctgaa
 gcctttgggaggagaggagagtggctggcccgatcacagatgcggaagg
 ggctgactcctcaaccgggtgcagactctgcagggtgggtctgggcca
 acacacccaaagcacgcccaggaaggaaaggcagcttggtatcactgcc
 agagctaggagaggcaccgggaaaatgatctgtccaagaccgttctgc
 ttctaaactccgaggggggtcagatgaagtgtttgttcttggcctgaa
 gcatcgtgttccctgcaagaagcggggaacacagaggaaaggagagaaaag
 atgaactgaacaaagcatgcaaggcaaaaaggGGGTCTAGCCGCGGTCT
 AGGAAGCTTTCTAGGGTACCTCTAGGGATCCCGGCGCGCCCTACCGGGTA
 GGGGAGGCGCTTTTCCCAAGGCAGTCTGGAGCATGCGCTTTAGCAGCCCC
 GCTGGGCACTTGGCGCTACACAAGTGGCCTCTGGCCTCGCACACATTCCA
 CATCCACCGGTAGGCGCCAACCGGCTCCGTTCTTTGGTGGCCCCCTTCGCG
 CCACCTTCTACTCCTCCCCTAGTCAGGAAGTTCCCCCCCCGCCCCGAGCT
 CGCGTCGTGCAGGACGTGACAAATGGAAGTAGCACGTCTCACTAGTCTCG
 TGCAGATGGACAGCACCGCTGAGCAATGGAAGCGGGTAGGCCTTTGGGGC
 AGCGGCCAATAGCAGCTTTGGCTCCTTCGCTTTCTGGGCTCAGAGGCTGG
 GAAGGGGTGGGTCCGGGGGGCGGGCTCAGGGGGCGGGGCTCAGGGGGCGGGGCG
 GGCGCCCGAAGGTCCTCCGGAAGCCCGGCATTCTGCACGCTTCAAAGCG
 CACGTCTGCCGCGCTGTTCTCCTCTTCCTCATCTCCGGGCCTTTTCGACCT

GCAGCCAATATGGGATCGGCCATTGAACAAGATGGATTGCACGCAGGTTC
 TCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGA
 CAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGC
 CCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCA
 GGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCG
 CAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTG
 GGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGA
 GAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATC
 CGGCTACCTGCCCATTTCGACCACCAAGCGAAACATCGCATCGAGCGAGCA
 CGTACTCGGATGGAAGCCGGTCTTGTCAATCAGGATGATCTGGACGAAGA
 GCATCAGGGGCTCGCGCCAGCCGAACTGTTCCGCAAGGCTCAAGGCGCGCA
 TGCCCGACGGCGAGGATCTCGTCTGACCCATGGCGATGCCTGCTTGCCG
 AATATCATGGTGGAAAATGGCCGCTTTTCTGGATTTCATCGACTGTGGCCG
 GCTGGGTGTGGCGGATCGCTATCAGGACATAGCGTTGGCTACCCGTGATA
 TTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTAC
 GGTATCGCCGCTCCCGATTTCGACGCGCATCGCCTTCTATCGCCTTCTTGA
 CGAGTTCTTCTGAGGGGATCAATTCTCTAGAGCTCGCTGATCAGCCTCGA
 CTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCT
 TCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGA
 GGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTG
 GGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCAT
 GCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTG
 GGGGCGCGCCctcgagcgccgcccagtgatgatatctgcagaattc
 gcccttgatcaaacacgcctcatggacaatatgttggttcttagc
 ctgctgagacacaacaggaactccccggcaccacttagaggccagaga
 aacagcacagataaaattccctgccctcatgaagcttagtctagctgg
 ggagatatcataggcaagataaacacatacaatacatcatcttaggtaa
 taatatataactaaggagaaaattacaggggagaaaggagacaggaattgc
 tagggtaggattataagttcagatagttcatcaggaacactgttgctgag
 aagataacatttaggtaaagaccgaagtagtaaggaaatggaccgtgtgc
 ctaagtgggtaagaccattctaggcagcaggaacagcgatgaaagcactg
 aggtgggtgttactgcacagagttgttactgcacagagttgtgtggg
 aggggtaggtcttcagggctcttatggtcacaggaagaattgtttactc
 ccaccgagatgaaggttggtgattttgagcagaagaataattctgcctg
 gtttatataaacaggatttccctgggtgctctgatgagaataatctgtc
 aggggtgggtagggagagatatggcaataggagccttggttaggagccc
 acgacaataattccaagtgagaggtggtgctgattgaaagcaggactaa
 caagacctgctgacagtggtgtagtaaaaagatagaggagacgaaggt
 gcatctagggttttctgctgaggaattagaaagataaagctaaagctta
 tagaagatgcagcgctctggggagaaagaccagcagctcagtttgatcc
 atctggaattaatttggcataaagtatgaggtatgtgggttaacattat
 ttgtttttttttccatgtagctatccaactgtcccagcatcattat
 tttaaaagactttccttccccctattggattgtttggcaccttactga
 agatcaactgagcataaaattgggtctatttctaagctcttgattccatt
 ccatgacctattgttcatctttaccccagtagacactgccttgatgatt
 aaagccccgttaccatgtctgttttgacatggtaaatctgagatgcct
 attagccaaccaagcaagcacggcccttagagagctagatatgagagcct
 ggaattcagacgagaaaggctcctagagacatacatgtagtccatc
 accatgcggatggtgttaaaagccatcagactgcaacagactgtgagagg
 gtaccaagctagagagcatggatagagaacccaagcactgagctgggag
 gtgctctacattaagagattagttagatgaaggactgagaagattgatc
 agagaagaaggaaaatcaggaaaatggtgctgtcctgaaaatccaaggga
 agagatgttccaaaggagagaaaactgatcagttgtcagctagcgtcaat
 tgggatgaaaatggaccattggacagagggatgtagtgggtcatgggtga
 atagataagagcagcttctatagaatggcaggggcaaaattctcatctga
 tcggcatgggttctaaagaaaacgggaagaaaaaattgagtcatgacca

	gtccctcaagtagagaggtggaaaagggaaggaggaaaatgaggccacg acaacatgagagaaatgacagcatttttaaaatttttattttttttia ttatttttttgccttttagggctgcccctgcaacatatggaggtcc caggtaggggtctaatcagagctatagctgccagcctacaccacagcca tagcaatgccagatctacatgacctacaccacagctcacagcaacgccgg atccttaaccactgagtgaggccagagatcaaacccatctcttatgga tactagttaggttcattaccactgagccaaaatgggaaatcctgagtaat gacagcatttttaatgtgccaggaagcaaaacttgccaccccgaaatgt ctctcaggcatgttgatttttgagctgaaaacgattaaggcccaaaaa acacaagaagaaatgtggaccttcccccaacagcctaaaaaattagatt gagggcctgttcccagaatagagctattgccagactgtctacagaggct aagggttaggtgtggtgggaaaccctcagagatcagaggacgtttatg taccaagcattgacatttccatctccatgcgaatggccttcttccctct gtagcccaaacaccacccccaaaatcttcttctgtctttagctgaaga tggtgttgaggtagatgttcagccactttggcgagttcctcagtgtt ctgggtcttctctccTgatccacattattcgactgtgtttgatttctcc tgtttatctgtctcattggcaccatttcattcttagaccagcccaaaga acctagaagagtgaaggaaaatttctccaccctgacaaatgctaaatga gaatcaccgcagtagaggaaaatgatctggtgctgcgggagatagaagag aaaatcgttgagagatgtcactgagtaggtgagatgggaaagggtgac acaggtggaggtgttgcctcagctaggaagacagacagttcacagaaga gaagcgggtgtccgtggacatcttgcctcatggatgaggaaaccgaggct aagaaagactgcaaaagaaaggtaaggattgcagagaggtcgatccatga ctaaaatcacagtaaccaacccccaaaccaccatgttttctctagtctgg cacgtggcaggtactgtgtaggtttcaatattattggtttgaacagta cctattaggcctccatccccctctctactaactaactaaaagtgtgagactg gtcagtgaaaaatggtcttcttctctatgaatctttctcaagaagatac ataactttttatttatcataggttgaagagcaaatgagaacagcctc caacctatgacaccgtaacaaaatgttatgatcagtgaagggaagaaa caaacatacacagtaaaagacctccataatattgtgggtggcccaacac aggccaggtgtaaaagctttttattctttgatagaggaatggatagtaa tgtttcaacctggacagagatcatgttactgaatccttcaaaaattca tggttagttgaattataaggaaaataagacttaggataaatactttgtc caagatcccagagttaatgccaaaatcagtttcagactccaggcagcct gatcaagagcctaaactttaagacacagtccttaataactactattca cagttgcactttcaggcgcaaagactcattgaatcctacaatagaatga gtttagatatcaaatctctcagtaatagatgaggagactaaatagcgggc atgacctggtcacttaagacagaattgagattcaaggctagtgtcttt ctacctgtttgtttctacaagatgtagcaatgcgctaattacagacctc tcagggaaggaa
--	---

Porcine Lambda Chain Targeting

In particular embodiments of the present invention, targeting vectors are provided to target the porcine heavy chain locus. In one embodiment, lambda can be targeted by designing a targeting construct that contains a 5' arm containing sequence located 5' to the first JC cluster and a 3' arm containing sequence 3' to the last JC cluster, thus preventing functional expression of the lambda locus (see, Figures 3-4). In one embodiment, the targeting vector can contain any contiguous sequence (such as about 17, 20, 30, 40, 50, 75, 100, 200, 300 or 5000 nucleotides of contiguous sequence) or fragment thereof Seq ID No 28. In one embodiment, the 5' targeting

arm can contain Seq ID No. 32, which includes 5' flanking sequence to the first lambda J/C region of the porcine lambda light chain genomic sequence or any contiguous sequence (such as about 17, 20, 30, 40, 50, 75, 100, 200, 300 or 5000 nucleotides of contiguous sequence) or fragment thereof (see also, for example Figure 5). In another embodiment, the 3' targeting arm can contain, but is not limited to one or more of the following: Seq ID No. 33, which includes 3' flanking sequence to the J/C cluster region of the porcine lambda light chain genomic sequence, from approximately 200 base pairs downstream of lambda J/C; Seq ID No. 34, which includes 3' flanking sequence to the J/C cluster region of the porcine lambda light chain genomic sequence, approximately 11.8 Kb downstream of the J/C cluster, near the enhancer; Seq ID No. 35, which includes approximately 12 Kb downstream of lambda, including the enhancer region; Seq ID No. 36, which includes approximately 17.6 Kb downstream of lambda; Seq ID No. 37, which includes approximately 19.1 Kb downstream of lambda; Seq ID No. 38, which includes approximately 21.3 Kb downstream of lambda; and Seq ID No. 39, which includes approximately 27 Kb downstream of lambda, or any contiguous sequence (such as about 17, 20, 30, 40, 50, 75, 100, 200, 300 or 5000 nucleotides of contiguous sequence) or fragment thereof of Seq ID Nos 32-39 (see also, for example Figure 6). It is understood that in general when designing a targeting construct one targeting arm will be 5' of the other targeting arm.

In additional embodiments, the targeting constructs for the lambda locus can contain site specific recombinase sites, such as, for example, lox. In one embodiment, the targeting arms can insert the site specific recombinase site into the targeted region. Then, the site specific recombinase can be activated and/ or applied to the cell such that the intervening nucleotide sequence between the two site specific recombinase sites is excised (see, for example, Figure 6).

Selectable Marker Genes

The DNA constructs can be designed to modify the endogenous, target immunoglobulin gene. The homologous sequence for targeting the construct can have one or more deletions, insertions, substitutions or combinations thereof. The alteration can be the insertion of a selectable marker gene fused in reading frame with the upstream sequence of the target gene.

Suitable selectable marker genes include, but are not limited to: genes conferring the ability to grow on certain media substrates, such as the tk gene (thymidine kinase) or the hppt gene (hypoxanthine phosphoribosyltransferase) which confer the ability to grow on HAT

medium (hypoxanthine, aminopterin and thymidine); the bacterial gpt gene (guanine/xanthine phosphoribosyltransferase) which allows growth on MAX medium (mycophenolic acid, adenine, and xanthine). See, for example, Song, K-Y., et al. *Proc. Nat'l Acad. Sci. U.S.A.* 84:6820-6824 (1987); Sambrook, J., et al., *Molecular Cloning--A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989), Chapter 16. Other examples of selectable markers include: genes conferring resistance to compounds such as antibiotics, genes conferring the ability to grow on selected substrates, genes encoding proteins that produce detectable signals such as luminescence, such as green fluorescent protein, enhanced green fluorescent protein (eGFP). A wide variety of such markers are known and available, including, for example, antibiotic resistance genes such as the neomycin resistance gene (neo) (Southern, P., and P. Berg, *J. Mol. Appl. Genet.* 1:327-341 (1982)); and the hygromycin resistance gene (hyg) (*Nucleic Acids Research* 11:6895-6911 (1983), and Te Riele, H., et al., *Nature* 348:649-651 (1990)). Other selectable marker genes include: acetohydroxyacid synthase (AHAS), alkaline phosphatase (AP), beta galactosidase (LacZ), beta glucuronidase (GUS), chloramphenicol acetyltransferase (CAT), green fluorescent protein (GFP), red fluorescent protein (RFP), yellow fluorescent protein (YFP), cyan fluorescent protein (CFP), horseradish peroxidase (HRP), luciferase (Luc), nopaline synthase (NOS), octopine synthase (OCS), and derivatives thereof. Multiple selectable markers are available that confer resistance to ampicillin, bleomycin, chloramphenicol, gentamycin, hygromycin, kanamycin, lincomycin, methotrexate, phosphinothricin, puromycin, and tetracycline.

Methods for the incorporation of antibiotic resistance genes and negative selection factors will be familiar to those of ordinary skill in the art (see, e.g., WO 99/15650; U.S. Patent No. 6,080,576; U.S. Patent No. 6,136,566; Niwa et al., *J. Biochem.* 113:343-349 (1993); and Yoshida et al., *Transgenic Research* 4:277-287 (1995)).

Combinations of selectable markers can also be used. For example, to target an immunoglobulin gene, a neo gene (with or without its own promoter, as discussed above) can be cloned into a DNA sequence which is homologous to the immunoglobulin gene. To use a combination of markers, the HSV-tk gene can be cloned such that it is outside of the targeting DNA (another selectable marker could be placed on the opposite flank, if desired). After introducing the DNA construct into the cells to be targeted, the cells can be selected on the appropriate antibiotics. In this particular example, those cells which are resistant to G418 and

gancyclovir are most likely to have arisen by homologous recombination in which the neo gene has been recombined into the immunoglobulin gene but the tk gene has been lost because it was located outside the region of the double crossover.

Deletions can be at least about 50 bp, more usually at least about 100 bp, and generally not more than about 20 kbp, where the deletion can normally include at least a portion of the coding region including a portion of or one or more exons, a portion of or one or more introns, and can or can not include a portion of the flanking non-coding regions, particularly the 5'-non-coding region (transcriptional regulatory region). Thus, the homologous region can extend beyond the coding region into the 5'-non-coding region or alternatively into the 3'-non-coding region. Insertions can generally not exceed 10 kbp, usually not exceed 5 kbp, generally being at least 50 bp, more usually at least 200 bp.

The region(s) of homology can include mutations, where mutations can further inactivate the target gene, in providing for a frame shift, or changing a key amino acid, or the mutation can correct a dysfunctional allele, etc. The mutation can be a subtle change, not exceeding about 5% of the homologous flanking sequences. Where mutation of a gene is desired, the marker gene can be inserted into an intron or an exon.

The construct can be prepared in accordance with methods known in the art, various fragments can be brought together, introduced into appropriate vectors, cloned, analyzed and then manipulated further until the desired construct has been achieved. Various modifications can be made to the sequence, to allow for restriction analysis, excision, identification of probes, etc. Silent mutations can be introduced, as desired. At various stages, restriction analysis, sequencing, amplification with the polymerase chain reaction, primer repair, in vitro mutagenesis, etc. can be employed.

The construct can be prepared using a bacterial vector, including a prokaryotic replication system, e.g. an origin recognizable by *E. coli*, at each stage the construct can be cloned and analyzed. A marker, the same as or different from the marker to be used for insertion, can be employed, which can be removed prior to introduction into the target cell. Once the vector containing the construct has been completed, it can be further manipulated, such as by deletion of the bacterial sequences, linearization, introducing a short deletion in the homologous sequence. After final manipulation, the construct can be introduced into the cell.

The present invention further includes recombinant constructs containing sequences of immunoglobulin genes. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. The construct can also include regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example. Bacterial: pBs, pQE-9 (Qiagen), phagescript, PsiX174, pBluescript SK, pBsKS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSv2cat, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPv, pMSG, pSVL (Pharmacia), viral origin vectors (M13 vectors, bacterial phage 1 vectors, adenovirus vectors, and retrovirus vectors), high, low and adjustable copy number vectors, vectors which have compatible replicons for use in combination in a single host (pACYC184 and pBR322) and eukaryotic episomal replication vectors (pCDM8). Other vectors include prokaryotic expression vectors such as pcDNA II, pSL301, pSE280, pSE380, pSE420, pTrcHisA, B, and C, pRSET A, B, and C (Invitrogen, Corp.), pGEMEX-1, and pGEMEX-2 (Promega, Inc.), the pET vectors (Novagen, Inc.), pTrc99A, pKK223-3, the pGEX vectors, pEZZ18, pRIT2T, and pMC1871 (Pharmacia, Inc.), pKK233-2 and pKK388-1 (Clontech, Inc.), and pProEx-HT (Invitrogen, Corp.) and variants and derivatives thereof. Other vectors include eukaryotic expression vectors such as pFastBac, pFastBacHT, pFastBacDUAL, pSFV, and pTet-Splice (Invitrogen), pEUK-C1, pPUR, pMAM, pMAMneo, pBI101, pBI121, pDR2, pCMVEBNA, and pYACneo (Clontech), pSVK3, pSVL, pMSG, pCH110, and pKK232-8 (Pharmacia, Inc.), p3'SS, pXT1, pSG5, pPbac, pMbac, pMC1neo, and pOG44 (Stratagene, Inc.), and pYES2, pAC360, pBlueBacHis A, B, and C, pVL1392, pBlueBacIII, pCDM8, pcDNA1, pZeoSV, pcDNA3 pREP4, pCEP4, and pEBVHis (Invitrogen, Corp.) and variants or derivatives thereof. Additional vectors that can be used include: pUC18, pUC19, pBlueScript, pSPORT, cosmids, phagemids, YAC's (yeast artificial chromosomes), BAC's (bacterial artificial chromosomes), P1 (Escherichia coli phage), pQE70, pQE60, pQE9 (quagan), pBS vectors, PhageScript vectors, BlueScript vectors, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene), pcDNA3 (Invitrogen), pGEX, pTrsfus, pTrc99A, pET-5, pET-9, pKK223-3, pDR540, pRIT5 (Pharmacia), pSPORT1, pSPORT2, pCMVSPORT2.0 and pSV-SPORT1 (Invitrogen), pTrxFus, pThioHis, pLEX, pTrcHis, pTrcHis2, pRSET, pBlueBacHis2,

pcDNA3.1/His, pcDNA3.1(-)/Myc-His, pSecTag, pEBVHis, pPIC9K, pPIC3.5K, pAO815, pPICZ, pPICZ□, pGAPZ, pGAPZ□, pBlueBac4.5, pBlueBacHis2, pMelBac, pSinRep5, pSinHis, pIND, pIND(SP1), pVgRXR, pcDNA2.1, pYES2, pZErO1.1, pZErO-2.1, pCR-Blunt, pSE280, pSE380, pSE420, pVL1392, pVL1393, pCDM8, pcDNA1.1, pcDNA1.1/Amp, pcDNA3.1, pcDNA3.1/Zeo, pSe, SV2, pRc/CMV2, pRc/RSV, pREP4, pREP7, pREP8, pREP9, pREP 10, pCEP4, pEBVHis, pCR3.1, pCR2.1, pCR3.1-Uni, and pCRBac from Invitrogen; □ ExCell, □ gt11, pTrc99A, pKK223-3, pGEX-1□T, pGEX-2T, pGEX-2TK, pGEX-4T-1, pGEX-4T-2, pGEX-4T-3, pGEX-3X, pGEX-5X-1, pGEX-5X-2, pGEX-5X-3, pEZZ18, pRIT2T, pMC1871, pSVK3, pSVL, pMSG, pCH110, pKK232-8, pSL1180, pNEO, and pUC4K from Pharmacia; pSCREEN-1b(+), pT7Blue(R), pT7Blue-2, pCITE-4abc(+), pOCUS-2, pTag, pET-32LIC, pET-30LIC, pBAC-2cp LIC, pBACgus-2cp LIC, pT7Blue-2 LIC, pT7Blue-2, □SCREEN-1, □BlueSTAR, pET-3abcd, pET-7abc, pET9abcd, pET11abcd, pET12abc, pET-14b, pET-15b, pET-16b, pET-17b-pET-17xb, pET-19b, pET-20b(+), pET-21abcd(+), pET-22b(+), pET-23abcd(+), pET-24abcd(+), pET-25b(+), pET-26b(+), pET-27b(+), pET-28abc(+), pET-29abc(+), pET-30abc(+), pET-31b(+), pET-32abc(+), pET-33b(+), pBAC-1, pBACgus-1, pBAC4x-1, pBACgus4x-1, pBAC-3cp, pBACgus-2cp, pBACsurf-1, plg, Signal plg, pYX, Selecta Vecta-Neo, Selecta Vecta-Hyg, and Selecta Vecta-Gpt from Novagen; pLexA, pB42AD, pGBT9, pAS2-1, pGAD424, pACT2, pGAD GL, pGAD GH, pGAD10, pGilda, pEZM3, pEGFP, pEGFP-1, pEGFP-N, pEGFP-C, pEBFP, pGFPuv, pGFP, p6xHis-GFP, pSEAP2-Basic, pSEAP2-Contral, pSEAP2-Promoter, pSEAP2-Enhancer, p□gal-Basic, p□gal-Control, p□gal-Promoter, p□gal-Enhancer, pCMV□, pTet-Off, pTet-On, pTK-Hyg, pRetro-Off, pRetro-On, pIRES1neo, pIRES1hyg, pLXSN, pLNCX, pLAPSN, pMAMneo, pMAMneo-CAT, pMAMneo-LUC, pPUR, pSV2neo, pYEX4T-1/2/3, pYEX-S1, pBacPAK-His, pBacPAK8/9, pAcUW31, BacPAK6, pTriplEx, □gt10, □gt11, pWE15, and □TriplEx from Clontech; Lambda ZAP II, pBK-CMV, pBK-RSV, pBluescript II KS +/-, pBluescript II SK +/-, pAD-GAL4, pBD-GAL4 Cam, pSurfsript, Lambda FIX II, Lambda DASH, Lambda EMBL3, Lambda EMBL4, SuperCos, pCR-Script Amp, pCR-Script Cam, pCR-Script Direct, pBS +/-, pBC KS +/-, pBC SK +/-, Phagescript, pCAL-n-EK, pCAL-n, pCAL-c, pCAL-kc, pET-3abcd, pET-11abcd, pSPUTK, pESP-1, pCMVLacI, pOPRSVI/MCS, pOPI3 CAT, pXT1, pSG5, pPbac, pMbac, pMC1neo, pMC1neo Poly A, pOG44, pOG45, pFRT□GAL, pNEO□GAL, pRS403, pRS404, pRS405, pRS406, pRS413, pRS414, pRS415, and pRS416 from Stratagene and variants or derivatives

thereof. Two-hybrid and reverse two-hybrid vectors can also be used, for example, pPC86, pDBLeu, pDBTrp, pPC97, p2.5, pGAD1-3, pGAD10, pAct, pACT2, pGADGL, pGADGH, pAS2-1, pGAD424, pGBT8, pGBT9, pGAD-GAL4, pLexA, pBD-GAL4, pHISi, pHISi-1, placZi, pB42AD, pDG202, pJK202, pJG4-5, pNLexA, pYESTrp and variants or derivatives thereof. Any other plasmids and vectors may be used as long as they are replicable and viable in the host.

Techniques which can be used to allow the DNA construct entry into the host cell include, for example, calcium phosphate/DNA co precipitation, microinjection of DNA into the nucleus, electroporation, bacterial protoplast fusion with intact cells, transfection, or any other technique known by one skilled in the art. The DNA can be single or double stranded, linear or circular, relaxed or supercoiled DNA. For various techniques for transfecting mammalian cells, see, for example, Keown et al., *Methods in Enzymology* Vol. 185, pp. 527-537 (1990).

In one specific embodiment, heterozygous or homozygous knockout cells can be produced by transfection of primary fetal fibroblasts with a knockout vector containing immunoglobulin gene sequence isolated from isogenic DNA. In another embodiment, the vector can incorporate a promoter trap strategy, using, for example, IRES (internal ribosome entry site) to initiate translation of the Neor gene.

Site Specific Recombinases

In additional embodiments, the targeting constructs can contain site specific recombinase sites, such as, for example, lox. In one embodiment, the targeting arms can insert the site specific recombinase target sites into the targeted region such that one site specific recombinase target site is located 5' to the second site specific recombinase target site. Then, the site specific recombinase can be activated and/or applied to the cell such that the intervening nucleotide sequence between the two site specific recombinase sites is excised.

Site-specific recombinases include enzymes or recombinases that recognize and bind to a short nucleic acid site or sequence-specific recombinase target site, i.e., a recombinase recognition site, and catalyze the recombination of nucleic acid in relation to these sites. These enzymes include recombinases, transposases and integrases. Examples of sequence-specific recombinase target sites include, but are not limited to, lox sites, att sites, dif sites and frt sites. Non-limiting examples of site-specific recombinases include, but are not limited to,

bacteriophage P1 Cre recombinase, yeast FLP recombinase, Inti integrase, bacteriophage λ , phi 80, P22, P2, 186, and P4 recombinase, Tn3 resolvase, the Hin recombinase, and the Cin recombinase, *E. coli* xerC and xerD recombinases, *Bacillus thuringiensis* recombinase, TpnI and the β -lactamase transposons, and the immunoglobulin recombinases.

In one embodiment, the recombination site can be a lox site that is recognized by the Cre recombinase of bacteriophage P1. Lox sites refer to a nucleotide sequence at which the product of the cre gene of bacteriophage P1, the Cre recombinase, can catalyze a site-specific recombination event. A variety of lox sites are known in the art, including the naturally occurring loxP, loxB, loxL and loxR, as well as a number of mutant, or variant, lox sites, such as loxP511, loxP514, lox.DELTA.86, lox.DELTA.117, loxC2, loxP2, loxP3 and lox P23. Additional example of lox sites include, but are not limited to, loxB, loxL, loxR, loxP, loxP3, loxP23, lox Δ 86, lox Δ 117, loxP511, and loxC2.

In another embodiment, the recombination site is a recombination site that is recognized by a recombinases other than Cre. In one embodiment, the recombinase site can be the FRT sites recognized by FLP recombinase of the 2 pi plasmid of *Saccharomyces cerevisiae*. FRT sites refer to a nucleotide sequence at which the product of the FLP gene of the yeast 2 micron plasmid, FLP recombinase, can catalyze site-specific recombination. Additional examples of the non-Cre recombinases include, but are not limited to, site-specific recombinases include: att sites recognized by the Int recombinase of bacteriophage λ (e.g. att1, att2, att3, attP, attB, attL, and attR), the recombination sites recognized by the resolvase family, and the recombination site recognized by transposase of *Bacillus thuringiensis*.

In particular embodiments of the present invention, the targeting constructs can contain: sequence homologous to a porcine immunoglobulin gene as described herein, a selectable marker gene and/ or a site specific recombinase target site.

Selection of Homologously Recombined Cells

The cells can then be grown in appropriately-selected medium to identify cells providing the appropriate integration. The presence of the selectable marker gene inserted into the immunoglobulin gene establishes the integration of the target construct into the host genome. Those cells which show the desired phenotype can then be further analyzed by restriction

analysis, electrophoresis, Southern analysis, polymerase chain reaction, etc to analyze the DNA in order to establish whether homologous or non-homologous recombination occurred. This can be determined by employing probes for the insert and then sequencing the 5' and 3' regions flanking the insert for the presence of the immunoglobulin gene extending beyond the flanking regions of the construct or identifying the presence of a deletion, when such deletion is introduced. Primers can also be used which are complementary to a sequence within the construct and complementary to a sequence outside the construct and at the target locus. In this way, one can only obtain DNA duplexes having both of the primers present in the complementary chains if homologous recombination has occurred. By demonstrating the presence of the primer sequences or the expected size sequence, the occurrence of homologous recombination is supported.

The polymerase chain reaction used for screening homologous recombination events is known in the art, see, for example, Kim and Smithies, *Nucleic Acids Res.* 16:8887-8903, 1988; and Joyner et al., *Nature* 338:153-156, 1989. The specific combination of a mutant polyoma enhancer and a thymidine kinase promoter to drive the neomycin gene has been shown to be active in both embryonic stem cells and EC cells by Thomas and Capecchi, *supra*, 1987; Nicholas and Berg (1983) in *Teratocarcinoma Stem Cell*, eds. Siver, Martin and Strikland (Cold Spring Harbor Lab., Cold Spring Harbor, N.Y. (pp. 469-497); and Linney and Donerly, *Cell* 35:693-699, 1983.

The cell lines obtained from the first round of targeting are likely to be heterozygous for the targeted allele. Homozygosity, in which both alleles are modified, can be achieved in a number of ways. One approach is to grow up a number of cells in which one copy has been modified and then to subject these cells to another round of targeting using a different selectable marker. Alternatively, homozygotes can be obtained by breeding animals heterozygous for the modified allele, according to traditional Mendelian genetics. In some situations, it can be desirable to have two different modified alleles. This can be achieved by successive rounds of gene targeting or by breeding heterozygotes, each of which carries one of the desired modified alleles.

Identification Of Cells That Have Undergone Homologous Recombination

In one embodiment, the selection method can detect the depletion of the immunoglobulin gene directly, whether due to targeted knockout of the immunoglobulin gene by homologous recombination, or a mutation in the gene that results in a nonfunctioning or nonexpressed immunoglobulin. Selection via antibiotic resistance has been used most commonly for screening (see above). This method can detect the presence of the resistance gene on the targeting vector, but does not directly indicate whether integration was a targeted recombination event or a random integration. Certain technology, such as Poly A and promoter trap technology, increase the probability of targeted events, but again, do not give direct evidence that the desired phenotype, a cell deficient in immunoglobulin gene expression, has been achieved. In addition, negative forms of selection can be used to select for targeted integration; in these cases, the gene for a factor lethal to the cells is inserted in such a way that only targeted events allow the cell to avoid death. Cells selected by these methods can then be assayed for gene disruption, vector integration and, finally, immunoglobulin gene depletion. In these cases, since the selection is based on detection of targeting vector integration and not at the altered phenotype, only targeted knockouts, not point mutations, gene rearrangements or truncations or other such modifications can be detected.

Animal cells believed to lacking expression of functional immunoglobulin genes can be further characterized. Such characterization can be accomplished by the following techniques, including, but not limited to: PCR analysis, Southern blot analysis, Northern blot analysis, specific lectin binding assays, and/or sequencing analysis.

PCR analysis as described in the art can be used to determine the integration of targeting vectors. In one embodiment, amplimers can originate in the antibiotic resistance gene and extend into a region outside the vector sequence. Southern analysis can also be used to characterize gross modifications in the locus, such as the integration of a targeting vector into the immunoglobulin locus. Whereas, Northern analysis can be used to characterize the transcript produced from each of the alleles.

Further, sequencing analysis of the cDNA produced from the RNA transcript can also be used to determine the precise location of any mutations in the immunoglobulin allele.

In another aspect of the present invention, ungulate cells lacking at least one allele of a functional region of an ungulate heavy chain, kappa light chain and/or lambda light chain locus produced according to the process, sequences and/or constructs described herein are provided. These cells can be obtained as a result of homologous recombination. Particularly, by inactivating at least one allele of an ungulate heavy chain, kappa light chain or lambda light chain gene, cells can be produced which have reduced capability for expression of porcine antibodies. In other embodiments, mammalian cells lacking both alleles of an ungulate heavy chain, kappa light chain and/or lambda light chain gene can be produced according to the process, sequences and/or constructs described herein. In a further embodiment, porcine animals are provided in which at least one allele of an ungulate heavy chain, kappa light chain and/or lambda light chain gene is inactivated via a genetic targeting event produced according to the process, sequences and/or constructs described herein. In another aspect of the present invention, porcine animals are provided in which both alleles of an ungulate heavy chain, kappa light chain and/or lambda light chain gene are inactivated via a genetic targeting event. The gene can be targeted via homologous recombination. In other embodiments, the gene can be disrupted, i.e. a portion of the genetic code can be altered, thereby affecting transcription and/or translation of that segment of the gene. For example, disruption of a gene can occur through substitution, deletion ("knock-out") or insertion ("knock-in") techniques. Additional genes for a desired protein or regulatory sequence that modulate transcription of an existing sequence can be inserted.

In embodiments of the present invention, alleles of ungulate heavy chain, kappa light chain or lambda light chain gene are rendered inactive according to the process, sequences and/or constructs described herein, such that functional ungulate immunoglobulins can no longer be produced. In one embodiment, the targeted immunoglobulin gene can be transcribed into RNA, but not translated into protein. In another embodiment, the targeted immunoglobulin gene can be transcribed in an inactive truncated form. Such a truncated RNA may either not be translated or can be translated into a nonfunctional protein. In an alternative embodiment, the targeted immunoglobulin gene can be inactivated in such a way that no transcription of the gene occurs. In a further embodiment, the targeted immunoglobulin gene can be transcribed and then translated into a nonfunctional protein.

III. Insertion of Artificial Chromosomes Containing Human Immunoglobulin Genes

Artificial Chromosomes

One aspect of the present invention provides ungulates and ungulate cells that lack at least one allele of a functional region of an ungulate heavy chain, kappa light chain and/or lambda light chain locus produced according to the processes, sequences and/or constructs described herein, which are further modified to express at least part of a human antibody (i.e. immunoglobulin (Ig)) locus. This human locus can undergo rearrangement and express a diverse population of human antibody molecules in the ungulate. These cloned, transgenic ungulates provide a replenishable, theoretically infinite supply of human antibodies (such as polyclonal antibodies), which can be used for therapeutic, diagnostic, purification, and other clinically relevant purposes.

In one particular embodiment, artificial chromosome (ACs) can be used to accomplish the transfer of human immunoglobulin genes into ungulate cells and animals. ACs permit targeted integration of megabase size DNA fragments that contain single or multiple genes. The ACs, therefore, can introduce heterologous DNA into selected cells for production of the gene product encoded by the heterologous DNA. In a one embodiment, one or more ACs with integrated human immunoglobulin DNA can be used as a vector for introduction of human Ig genes into ungulates (such as pigs).

First constructed in yeast in 1983, ACs are man-made linear DNA molecules constructed from essential cis-acting DNA sequence elements that are responsible for the proper replication and partitioning of natural chromosomes (Murray et al. (1983), Nature 301:189-193). A chromosome requires at least three elements to function. Specifically, the elements of an artificial chromosome include at least: (1) autonomous replication sequences (ARS) (having properties of replication origins – which are the sites for initiation of DNA replication), (2) centromeres (site of kinetochore assembly that is responsible for proper distribution of replicated chromosomes at mitosis and meiosis), and (3) telomeres (specialized structures at the ends of linear chromosomes that function to both stabilize the ends and facilitate the complete replication of the extreme termini of the DNA molecule).

In one embodiment, the human Ig can be maintained as an independent unit (an episome) apart from the ungulate chromosomal DNA. For example, episomal vectors contain the

necessary DNA sequence elements required for DNA replication and maintenance of the vector within the cell. Episomal vectors are available commercially (see, for example, Maniatis, T. et al., *Molecular Cloning, A Laboratory Manual* (1982) pp. 368-369). The AC can stably replicate and segregate along side endogenous chromosomes. In an alternative embodiment, the human IgG DNA sequences can be integrated into the ungulate cell's chromosomes thereby permitting the new information to be replicated and partitioned to the cell's progeny as a part of the natural chromosomes (see, for example, Wigler et al. (1977), *Cell* 11:223). The AC can be translocated to, or inserted into, the endogenous chromosome of the ungulate cell. Two or more ACs can be introduced to the host cell simultaneously or sequentially.

ACs, furthermore, can provide an extra-genomic locus for targeted integration of megabase size DNA fragments that contain single or multiple genes, including multiple copies of a single gene operatively linked to one promoter or each copy or several copies linked to separate promoters. ACs can permit the targeted integration of megabase size DNA fragments that contain single or multiple human immunoglobulin genes. The ACs can be generated by culturing the cells with dicentric chromosomes (i.e., chromosomes with two centromeres) under such conditions known to one skilled in the art whereby the chromosome breaks to form a minichromosome and formerly dicentric chromosome.

ACs can be constructed from humans (human artificial chromosomes: "HACs"), yeast (yeast artificial chromosomes: "YACs"), bacteria (bacterial artificial chromosomes: "BACs"), bacteriophage P1-derived artificial chromosomes: "PACs") and other mammals (mammalian artificial chromosomes: "MACs"). The ACs derive their name (e.g., YAC, BAC, PAC, MAC, HAC) based on the origin of the centromere. A YAC, for example, can derive its centromere from *S. cerevisiae*. MACs, on the other hand, include an active mammalian centromere while HACs refer to chromosomes that include human centromeres. Furthermore, plant artificial chromosomes ("PLACs") and insect artificial chromosomes can also be constructed. The ACs can include elements derived from chromosomes that are responsible for both replication and maintenance. ACs, therefore, are capable of stably maintaining large genomic DNA fragments such as human Ig DNA..

In one embodiment, ungulates containing YACs are provided. YACs are genetically engineered circular chromosomes that contain elements from yeast chromosomes, such as *S. cerevisiae*, and segments of foreign DNAs that can be much larger than those accepted by

conventional cloning vectors (e.g., plasmids, cosmids). YACs allow the propagation of very large segments of exogenous DNA (Schlessinger, D. (1990), Trends in Genetics 6:248-253) into mammalian cells and animals (Choi et al. (1993), Nature Gen 4:117-123). YAC transgenic approaches are very powerful and are greatly enhanced by the ability to efficiently manipulate the cloned DNA. A major technical advantage of yeast is the ease with which specific genome modifications can be made via DNA-mediated transformation and homologous recombination (Ramsay, M. (1994), Mol Biotech 1:181-201). In one embodiment, one or more YACs with integrated human Ig DNA can be used as a vector for introduction of human Ig genes into ungulates (such as pigs).

The YAC vectors contain specific structural components for replication in yeast, including: a centromere, telomeres, autonomous replication sequence (ARS), yeast selectable markers (e.g., *TRP1*, *URA3*, and *SUP4*), and a cloning site for insertion of large segments of greater than 50 kb of exogenous DNA. The marker genes can allow selection of the cells carrying the YAC and serve as sites for the synthesis of specific restriction endonucleases. For example, the *TRP1* and *URA3* genes can be used as dual selectable markers to ensure that only complete artificial chromosomes are maintained. Yeast selectable markers can be carried on both sides of the centromere, and two sequences that seed telomere formation *in vivo* are separated. Only a fraction of one percent of a yeast cell's total DNA is necessary for replication, however, including the center of the chromosome (the centromere, which serves as the site of attachment between sister chromatids and the sites of spindle fiber attachment during mitosis), the ends of the chromosome (telomeres, which serve as necessary sequences to maintain the ends of eukaryotic chromosomes), and another short stretch of DNA called the ARS which serves as DNA segments where the double helix can unwind and begin to copy itself.

In one embodiment, YACs can be used to clone up to about 1, 2, or 3 Mb of immunoglobulin DNA. In another embodiment, at least 25, 30, 40, 50, 60, 70, 75, 80, 85, 90, or 95 kilobases.

Yeast integrating plasmids, replicating vectors (which are fragments of YACs), can also be used to express human Ig. The yeast integrating plasmid can contain bacterial plasmid sequences that provide a replication origin and a drug-resistance gene for growth in bacteria (e.g., *E. coli*), a yeast marker gene for selection of transformants in yeast, and restriction sites for inserting Ig sequences. Host cells can stably acquire this plasmid by integrating it directly into a

chromosome. Yeast replicating vectors can also be used to express human Ig as free plasmid circles in yeast. Yeast or ARS-containing vectors can be stabilized by the addition of a centromere sequence. YACs have both centromeric and telomeric regions, and can be used for cloning very large pieces of DNA because the recombinant is maintained essentially as a yeast chromosome.

YACs are provided, for example, as disclosed in U.S. Pat. Nos. 6,692,954, 6,495,318, 6,391,642, 6,287,853, 6,221,588, 6,166,288, 6,096,878, 6,015,708, 5,981,175, 5,939,255, 5,843,671, 5,783,385, 5,776,745, 5,578,461, and 4,889,806; European Patent Nos. 1 356 062 and 0 648 265; PCT Publication Nos. WO 03/025222, WO 02/057437, WO 02/101044, WO 02/057437, WO 98/36082, WO 98/12335, WO 98/01573, WO 96/01276, WO 95/14769, WO 95/05847, WO 94/23049, and WO 94/00569.

In another embodiment, ungulates containing BACs are provided. BACs are F-based plasmids found in bacteria, such as *E. Coli*, that can transfer approximately 300 kb of foreign DNA into a host cell. Once the Ig DNA has been cloned into the host cell, the newly inserted segment can be replicated along with the rest of the plasmid. As a result, billions of copies of the foreign DNA can be made in a very short time. In a particular embodiment, one or more BACs with integrated human Ig DNA are used as a vector for introduction of human Ig genes into ungulates (such as pigs).

The BAC cloning system is based on the *E. coli* F-factor, whose replication is strictly controlled and thus ensures stable maintenance of large constructs (Willems, N., and R. Skurray (1987), Structure and function of the F-factor and mechanism of conjugation. In *Escherichia coli* and *Salmonella Typhimurium*: Cellular and Molecular Biology (F. C. Neidhardt, Ed) Vol.2 pp 1110-1133, Am. Soc. Microbiol., Washington, D.C.). BACs have been widely used for cloning of DNA from various eukaryotic species (Cai et al. (1995), Genomics 29:413-425; Kim et al. (1996), Genomics 34:213-218; Misumi et al. (1997), Genomics 40:147-150; Woo et al. (1994), Nucleic Acids Res 22:4922-4931; Zimmer, R. and Gibbins, A.M.V. (1997), Genomics 42:217-226). The low occurrence of the F-plasmid can reduce the potential for recombination between DNA fragments and can avoid the lethal overexpression of cloned bacterial genes. BACs can stably maintain the human immunoglobulin genes in a single copy vector in the host cells, even after 100 or more generations of serial growth.

BAC (or pBAC) vectors can accommodate inserts in the range of approximately 30 to 300 kb pairs. One specific type of BAC vector, pBeloBac11, uses a complementation of the lacZ gene to distinguish insert-containing recombinant molecules from colonies carrying the BAC vector, by color. When a DNA fragment is cloned into the lacZ gene of pBeloBac11, insertional activation results in a white colony on X-Gal/IPTG plates after transformation (Kim et al. (1996), Genomics 34:213-218) to easily identify positive clones.

For example, BACs can be provided such as disclosed in U.S. Pat. Nos. 6,713,281, 6,703,198, 6,649,347, 6,638,722, 6,586,184, 6,573,090, 6,548,256, 6,534,262, 6,492,577, 6,492,506, 6,485,912, 6,472,177, 6,455,254, 6,383,756, 6,277,621, 6,183,957, 6,156,574, 6,127,171, 5,874,259, 5,707,811, and 5,597,694; European Patent Nos. 0 805 851; PCT Publication Nos. WO 03/087330, WO 02/00916, WO 01/39797, WO 01/04302, WO 00/79001, WO 99/54487, WO 99/27118, and WO 96/21725.

In another embodiment, ungulates containing bacteriophage PACs are provided. In a particular embodiment, one or more bacteriophage PACs with integrated human Ig DNA are used as a vector for introduction of human Ig genes into ungulates (such as pigs). For example, PACs can be provided such as disclosed in U.S. Pat. Nos. 6,743,906, 6,730,500, 6,689,606, 6,673,909, 6,642,207, 6,632,934, 6,573,090, 6,544,768, 6,489,458, 6,485,912, 6,469,144, 6,462,176, 6,413,776, 6,399,312, 6,340,595, 6,287,854, 6,284,882, 6,277,621, 6,271,008, 6,187,533, 6,156,574, 6,153,740, 6,143,949, 6,017,755, and 5,973,133; European Patent Nos. 0 814 156; PCT Publication Nos. WO 03/091426, WO 03/076573, WO 03/020898, WO 02/101022, WO 02/070696, WO 02/061073, WO 02/31202, WO 01/44486, WO 01/07478, WO 01/05962, and WO 99/63103, .

In a further embodiment, ungulates containing MACs are provided. MACs possess high mitotic stability, consistent and regulated gene expression, high cloning capacity, and non-immunogenicity. Mammalian chromosomes can be comprised of a continuous linear strand of DNA ranging in size from approximately 50 to 250 Mb. The DNA construct can further contain one or more sequences necessary for the DNA construct to multiply in yeast cells. The DNA construct can also contain a sequence encoding a selectable marker gene. The DNA construct can be capable of being maintained as a chromosome in a transformed cell with the DNA construct. MACs provide extra-genomic specific integration sites for introduction of genes encoding proteins of interest and permit megabase size DNA integration so that, for example,

genes encoding an entire metabolic pathway, a very large gene [e.g., such as the cystic fibrosis (CF) gene (~ 600 kb)], or several genes [e.g., a series of antigens for preparation of a multivalent vaccine] can be stably introduced into a cell.

Mammalian artificial chromosomes [MACs] are provided. Also provided are artificial chromosomes for other higher eukaryotic species, such as insects and fish, produced using the MACS are provided herein. Methods for generating and isolating such chromosomes. Methods using the MACs to construct artificial chromosomes from other species, such as insect and fish species are also provided. The artificial chromosomes are fully functional stable chromosomes. Two types of artificial chromosomes are provided. One type, herein referred to as SATACs [satellite artificial chromosomes] are stable heterochromatic chromosomes, and the another type are minichromosomes based on amplification of euchromatin. As used herein, a formerly dicentric chromosome is a chromosome that is produced when a dicentric chromosome fragments and acquires new telomeres so that two chromosomes, each having one of the centromeres, are produced. Each of the fragments can be replicable chromosomes.

Also provided are artificial chromosomes for other higher eukaryotic species, such as insects and fish, produced using the MACS are provided herein. In one embodiment, SATACs [satellite artificial chromosomes] are provided. SATACs are stable heterochromatic chromosomes. In another embodiment, minichromosomes are provided wherein the minichromosomes are based on amplification of euchromatin.

In one embodiment, artificial chromosomes can be generated by culturing the cells with the dicentric chromosomes under conditions whereby the chromosome breaks to form a minichromosome and formerly dicentric chromosome. In one embodiment, the SATACs can be generated from the minichromosome fragment, see, for example, in U.S. Pat. No. 5,288,625. In another embodiment, the SATACs can be generated from the fragment of the formerly dicentric chromosome. The SATACs can be made up of repeating units of short satellite DNA and can be fully heterochromatic. In one embodiment, absent insertion of heterologous or foreign DNA, the SATACs do not contain genetic information. In other embodiments, SATACs of various sizes are provided that are formed by repeated culturing under selective conditions and subcloning of cells that contain chromosomes produced from the formerly dicentric chromosomes. These chromosomes can be based on repeating units 7.5 to 10 Mb in size, or megareplicons. These megareplicons can be tandem blocks of satellite DNA flanked by heterologous non-satellite

DNA. Amplification can produce a tandem array of identical chromosome segments [each called an amplicon] that contain two inverted megareplicons bordered by heterologous ["foreign"] DNA. Repeated cell fusion, growth on selective medium and/or BrdU [5-bromodeoxyuridine] treatment or other genome destabilizing reagent or agent, such as ionizing radiation, including X-rays, and subcloning can result in cell lines that carry stable heterochromatic or partially heterochromatic chromosomes, including a 150-200 Mb "sausage" chromosome, a 500-1000 Mb gigachromosome, a stable 250-400 Mb megachromosome and various smaller stable chromosomes derived therefrom. These chromosomes are based on these repeating units and can include human immunoglobulin DNA that is expressed. (See also US Patent No. 6,743,967)

In other embodiments, MACs can be provided, for example, as disclosed in U.S. Pat. Nos. 6,743,967, 6,682,729, 6,569,643, 6,558,902, 6,548,287, 6,410,722, 6,348,353, 6,297,029, 6,265,211, 6,207,648, 6,150,170, 6,150,160, 6,133,503, 6,077,697, 6,025,155, 5,997,881, 5,985,846, 5,981,225, 5,877,159, 5,851,760, and 5,721,118; PCT Publication Nos. WO 04/066945, WO 04/044129, WO 04/035729, WO 04/033668, WO 04/027075, WO 04/016791, WO 04/009788, WO 04/007750, WO 03/083054, WO 03/068910, WO 03/068909, WO 03/064613, WO 03/052050, WO 03/027315, WO 03/023029, WO 03/012126, WO 03/006610, WO 03/000921, WO 02/103032, WO 02/097059, WO 02/096923, WO 02/095003, WO 02/092615, WO 02/081710, WO 02/059330, WO 02/059296, WO 00/18941, WO 97/16533, and WO 96/40965.

In another aspect of the present invention, ungulates and ungulate cells containing HACs are provided. In a particular embodiment, one or more HACs with integrated human Ig DNA are used as a vector for introduction of human Ig genes into ungulates (such as pigs). In a particular embodiment, one or more HACs with integrated human Ig DNA are used to generate ungulates (for example, pigs) by nuclear transfer which express human Igs in response to immunization and which undergo affinity maturation.

Various approaches may be used to produce ungulates that express human antibodies ("human Ig"). These approaches include, for example, the insertion of a HAC containing both heavy and light chain Ig genes into an ungulate or the insertion of human B-cells or B-cell precursors into an ungulate during its fetal stage or after it is born (e.g., an immune deficient or immune suppressed ungulate) (see, for example, WO 01/35735, filed November 17, 2000, US

02/08645, filed March 20, 2002). In either case, both human antibody producing cells and ungulate antibody-producing B-cells may be present in the ungulate. In an ungulate containing a HAC, a single B-cell may produce an antibody that contains a combination of ungulate and human heavy and light chain proteins. In still other embodiments, the total size of the HAC is at least to approximately 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 Mb.

For example, HACs can be provided such as disclosed in U.S. Pat. Nos. 6,642,207, 6,590,089, 6,566,066, 6,524,799, 6,500,642, 6,485,910, 6,475,752, 6,458,561, 6,455,026, 6,448,041, 6,410,722, 6,358,523, 6,277,621, 6,265,211, 6,146,827, 6,143,566, 6,077,697, 6,025,155, 6,020,142, and 5,972,649; U.S. Pat. Application No. 2003/0037347; PCT Publication Nos. WO 04/050704, WO 04/044156, WO 04/031385, WO 04/016791, WO 03/101396, WO 03/097812, WO 03/093469, WO 03/091426, WO 03/057923, WO 03/057849, WO 03/027638, WO 03/020898, WO 02/092812, and WO 98/27200.

Additional examples of ACs into which human immunoglobulin sequences can be inserted for use in the invention include, for example, BACs (e.g., pBeloBAC11 or pBAC108L; see, e.g., Shizuya et al. (1992), Proc Natl Acad Sci USA 89(18):8794-8797; Wang et al. (1997), Biotechniques 23(6):992-994), bacteriophage PACs, YACs (see, e.g., Burke (1990), Genet Anal Tech Appl 7(5):94-99), and MACs (see, e.g., Vos (1997), Nat. Biotechnol. 15(12):1257-1259; Ascenzioni et al. (1997), Cancer Lett 118(2):135-142), such as HACs, see also, U.S. Pat. Nos. 6,743,967, 6,716,608, 6,692,954, 6,670,154, 6,642,207, 6,638,722, 6,573,090, 6,492,506, 6,348,353, 6,287,853, 6,277,621, 6,183,957, 6,156,953, 6,133,503, 6,090,584, 6,077,697, 6,025,155, 6,015,708, 5,981,175, 5,874,259, 5,721,118, and 5,270,201; European Patent Nos. 1 437 400, 1 234 024, 1 356 062, 0 959 134, 1 056 878, 0 986 648, 0 648 265, and 0 338 266; PCT Publication Nos. WO 04/013299, WO 01/07478, WO 00/06715, WO 99/43842, WO 99/27118, WO 98/55637, WO 94/00569, and WO 89/09219. Additional examples include those AC provided in, for example, PCT Publication No. WO 02/076508, WO 03/093469, WO 02/097059; WO 02/096923; US Publication Nos US 2003/0113917 and US 2003/003435; and US Patent No 6,025,155.

In other embodiments of the present invention, ACs transmitted through male gametogenesis in each generation. The AC can be integrating or non-integrating. In one embodiment, the AC can be transmitted through mitosis in substantially all dividing cells. In another embodiment, the AC can provide for position independent expression of a human

immunoglobulin nucleic acid sequence. In a particular embodiment, the AC can have a transmittal efficiency of at least 10% through each male and female gametogenesis. In one particular embodiment, the AC can be circular. In another particular embodiment, the non-integrating AC can be that deposited with the Belgian Coordinated Collections of Microorganisms--BCCM on Mar. 27, 2000 under accession number LMBP 5473 CB. In additional embodiments, methods for producing an AC are provided wherein a mitotically stable unit containing an exogenous nucleic acid transmitted through male gametogenesis is identified; and an entry site in the mitotically stable unit allows for the integration of human immunoglobulin genes into the unit.

In other embodiments, ACs are provided that include: a functional centromere, a selectable marker and/or a unique cloning site. In other embodiments, the AC can exhibit one or more of the following properties: it can segregate stably as an independent chromosome, immunoglobulin sequences can be inserted in a controlled way and can be expressed from the AC, it can be efficiently transmitted through the male and female germline and/or the transgenic animals can bear the chromosome in greater than about 30, 40, 50, 60, 70, 80 or 90% of its cells.

In particular embodiments, the AC can be isolated from fibroblasts (such as any mammalian or human fibroblast) in which it was mitotically stable. After transfer of the AC into hamster cells, a lox (such as loxP) site and a selectable marker site can be inserted. In other embodiments, the AC can maintain mitotic stability, for example, showing a loss of less than about 5, 2, 1, 0.5 or 0.25 percent per mitosis in the absence of selection. See also, US 2003/0064509 and WO 01/77357.

Xenogenous Immunoglobulin Genes

In another aspect of the present invention, transgenic ungulates are provided that expresses a xenogenous immunoglobulin loci or fragment thereof, wherein the immunoglobulin can be expressed from an immunoglobulin locus that is integrated within an endogenous ungulate chromosome. In one embodiment, ungulate cells derived from the transgenic animals are provided. In one embodiment, the xenogenous immunoglobulin locus can be inherited by offspring. In another embodiment, the xenogenous immunoglobulin locus can be inherited through the male germ line by offspring. In still further embodiments, an artificial chromosome

(AC) can contain the xenogenous immunoglobulin. In one embodiment, the AC can be a yeast AC or a mammalian AC. In a further embodiment, the xenogenous locus can be a human immunoglobulin locus or fragment thereof. In one embodiment, the human immunoglobulin locus can be human chromosome 14, human chromosome 2, and human chromosome 22 or fragments thereof. In another embodiment, the human immunoglobulin locus can include any fragment of a human immunoglobulin that can undergo rearrangement. In a further embodiment, the human immunoglobulin loci can include any fragment of a human immunoglobulin heavy chain and a human immunoglobulin light chain that can undergo rearrangement. In still further embodiment, the human immunoglobulin loci can include any human immunoglobulin locus or fragment thereof that can produce an antibody upon exposure to an antigen. In a particular embodiment, the exogenous human immunoglobulin can be expressed in B cells to produce xenogenous immunoglobulin in response to exposure to one or more antigens.

In other embodiments, the transgenic ungulate that lacks any expression of functional endogenous immunoglobulins can be further genetically modified to express an xenogenous immunoglobulin loci. In an alternative embodiment, porcine animals are provided that contain an xenogenous immunoglobulin locus. In one embodiment, the xenogenous immunoglobulin loci can be a heavy and/ or light chain immunoglobulin or fragment thereof. In another embodiment, the xenogenous immunoglobulin loci can be a kappa chain locus or fragment thereof and/ or a lambda chain locus or fragment thereof. In still further embodiments, an artificial chromosome (AC) can contain the xenogenous immunoglobulin. In one embodiment, the AC can be a yeast AC or a mammalian AC. In a further embodiment, the xenogenous locus can be a human immunoglobulin locus or fragment thereof. In one embodiment, the human immunoglobulin locus can be human chromosome 14, human chromosome 2, and human chromosome 22 or fragments thereof. In another embodiment, the human immunoglobulin locus can include any fragment of a human immunoglobulin that can undergo rearrangement. In a further embodiment, the human immunoglobulin loci can include any fragment of a human immunoglobulin heavy chain and a human immunoglobulin light chain that can undergo rearrangement. In still further embodiment, the human immunoglobulin loci can include any human immunoglobulin locus or fragment thereof that can produce an antibody upon exposure to an antigen. In a particular embodiment, the exogenous human immunoglobulin can be expressed in B cells to produce xenogenous immunoglobulin in response to exposure to one or more antigens.

In other embodiments, the transgenic ungulate that lacks any expression of functional endogenous immunoglobulins can be further genetically modified to express an xenogenous immunoglobulin loci. In an alternative embodiment, porcine animals are provided that contain an xenogenous immunoglobulin locus. In one embodiment, the xenogenous immunoglobulin loci can be a heavy and/ or light chain immunoglobulin or fragment thereof. In another embodiment, the xenogenous immunoglobulin loci can be a kappa chain locus or fragment thereof and/ or a lambda chain locus or fragment thereof. In still further embodiments, an artificial chromosome (AC) can contain the xenogenous immunoglobulin. In one embodiment, the AC can be a yeast AC or a mammalian AC. In a further embodiment, the xenogenous locus can be a human immunoglobulin locus or fragment thereof. In one embodiment, the human immunoglobulin locus can be human chromosome 14, human chromosome 2, and human chromosome 22 or fragments thereof. In another embodiment, the human immunoglobulin locus can include any fragment of a human immunoglobulin that can undergo rearrangement. In a further embodiment, the human immunoglobulin loci can include any fragment of a human immunoglobulin heavy chain and a human immunoglobulin light chain that can undergo rearrangement. In still further embodiment, the human immunoglobulin loci can include any human immunoglobulin locus or fragment thereof that can produce an antibody upon exposure to an antigen. In a particular embodiment, the exogenous human immunoglobulin can be expressed in B cells to produce xenogenous immunoglobulin in response to exposure to one or more antigens.

In another embodiment, porcine animals are provided that contain an xenogenous immunoglobulin locus. In one embodiment, the xenogenous immunoglobulin loci can be a heavy and/ or light chain immunoglobulin or fragment thereof. In another embodiment, the xenogenous immunoglobulin loci can be a kappa chain locus or fragment thereof and/ or a lambda chain locus or fragment thereof. In still further embodiments, an artificial chromosome (AC) can contain the xenogenous immunoglobulin. In one embodiment, the AC can be a yeast AC or a mammalian AC. In a further embodiment, the xenogenous locus can be a human immunoglobulin locus or fragment thereof. In one embodiment, the human immunoglobulin locus can be human chromosome 14, human chromosome 2, and human chromosome 22 or fragments thereof. In another embodiment, the human immunoglobulin locus can include any fragment of a human immunoglobulin that can undergo rearrangement. In a further embodiment, the human immunoglobulin loci can include any fragment of a human immunoglobulin heavy

chain and a human immunoglobulin light chain that can undergo rearrangement. In still further embodiment, the human immunoglobulin loci can include any human immunoglobulin locus or fragment thereof that can produce an antibody upon exposure to an antigen. In a particular embodiment, the exogenous human immunoglobulin can be expressed in B cells to produce xenogenous immunoglobulin in response to exposure to one or more antigens.

Human immunoglobulin genes, such as the Ig heavy chain gene (human chromosome 414), Ig kappa chain gene (human chromosome #2) and/or the Ig lambda chain gene (chromosome #22) can be inserted into Acs, as described above. In a particular embodiment, any portion of the human heavy, kappa and/or lambda Ig genes can be inserted into ACs. In one embodiment, the nucleic acid can be at least 70, 80, 90, 95, or 99% identical to the corresponding region of a naturally-occurring nucleic acid from a human. In other embodiments, more than one class of human antibody is produced by the ungulate. In various embodiments, more than one different human Ig or antibody is produced by the ungulate. In one embodiment, an AC containing both a human Ig heavy chain gene and Ig light chain gene, such as an automatic human artificial chromosome ("AHAC," a circular recombinant nucleic acid molecule that is converted to a linear human chromosome *in vivo* by an endogenously expressed restriction endonuclease) can be introduced. In one embodiment, the human heavy chain loci and the light chain loci are on different chromosome arms (i.e., on different side of the centromere). In one embodiment, the heavy chain can include the mu heavy chain, and the light chain can be a lambda or kappa light chain. The Ig genes can be introduced simultaneously or sequentially in one or more than one ACs.

In particular embodiments, the ungulate or ungulate cell expresses one or more nucleic acids encoding all or part of a human Ig gene which undergoes rearrangement and expresses more than one human Ig molecule, such as a human antibody protein. Thus, the nucleic acid encoding the human Ig chain or antibody is in its unrearranged form (that is, the nucleic acid has not undergone V(D)J recombination). In particular embodiments, all of the nucleic acid segments encoding a V gene segment of an antibody light chain can be separated from all of the nucleic acid segments encoding a J gene segment by one or more nucleotides. In a particular embodiment, all of the nucleic acid segments encoding a V gene segment of an antibody heavy chain can be separated from all of the nucleic acid segments encoding a D gene segment by one or more nucleotides, and/or all of the nucleic acid segments encoding a D gene segment of an

antibody heavy chain are separated from all of the nucleic acid segments encoding a J gene segment by one or more nucleotides. Administration of an antigen to a transgenic ungulate containing an unrearranged human Ig gene is followed by the rearrangement of the nucleic acid segments in the human Ig gene locus and the production of human antibodies reactive with the antigen.

In one embodiment, the AC can express a portion or fragment of a human chromosome that contains an immunoglobulin gene. In one embodiment, the AC can express at least 300 or 1300 kb of the human light chain locus, such as described in Davies et al. 1993 *Biotechnology* 11: 911-914.

In another embodiment, the AC can express a portion of human chromosome 22 that contains at least the λ light-chain locus, including V_λ gene segments, J_λ gene segments, and the single C_λ gene. In another embodiment, the AC can express at least one V_λ gene segment, at least one J_λ gene segment, and the C_λ gene. In other embodiment, ACs can contain portions of the lambda locus, such as described in Popov et al. *J Exp Med.* 1999 May 17;189(10):1611-20.

In another embodiment, the AC can express a portion of human chromosome 2 that contains at least the κ light-chain locus, including V_κ gene segments, J_κ gene segments and the single C_κ gene. In another embodiment, the AC can express at least one V_κ gene segment, at least one J_κ gene segment and the C_κ gene. In other embodiments, AC containing portions of the kappa light chain locus can be those describe, for example, in Li et al. 2000 *J Immunol* 164: 812-824 and Li S *Proc Natl Acad Sci U S A.* 1987 Jun;84(12):4229-33. In another embodiment, AC containing approximately 1.3 Mb of human kappa locus are provided, such as described in Zou et al *FASEB J.* 1996 Aug;10(10):1227-32.

In further embodiments, the AC can express a portion of human chromosome 14 that contains at least the human heavy-chain locus, including V_H , D_H , J_H and C_H gene segments. In another embodiment, the AC can express at least one V_H gene segment, at least one D_H gene segment, at least one J_H gene segment and at least one at least one C_H gene segment. In other embodiments, the AC can express at least 85 kb of the human heavy chain locus, such as described in Choi et al. 1993 *Nat Gen* 4:117-123 and/or Zou et al. 1996 *PNAS* 96: 14100-14105.

In other embodiments, the AC can express portions of both heavy and light chain loci, such as, at least 220, 170, 800 or 1020 kb, for example, as disclosed in Green et al. 1994 *Nat Gen* 7:13-22; Mendez et al 1995 *Genomics* 26: 294-307; Mendez et al. 1997 *Nat Gen* 15: 146-156;

Green et al. 1998 J Exp Med 188: 483-495 and/or Fishwild et al. 1996 Nat Biotech 14: 845-851. In another embodiment, the AC can express megabase amounts of human immunoglobulin, such as described in Nicholson J Immunol. 1999 Dec 15;163(12):6898-906 and Popov Gene. 1996 Oct 24;177(1-2):195-201. In addition, in one particular embodiment, MACs derived from human chromosome #14 (comprising the Ig heavy chain gene), human chromosome #2 comprising the Ig kappa chain gene) and human chromosome #22 (comprising the Ig lambda chain gene) can be introduced simultaneously or successively, such as described in US Patent Publication No. 2004/0068760 to Robl et al.. In another embodiments, the total size of the MAC is less than or equal to approximately 10, 9, 8, or 7 megabases.

In a particular embodiment, human Vh, human Dh, human Jh segments and human mu segments of human immunoglobulins in germline configuration can be inserted into an AC, such as a YAC, such that the Vh, Dh, Jh and mu DNA segments form a repertoire of immunoglobulins containing portions which correspond to the human DNA segments, for example, as described in U.S. Patent No. 5,545,807 to the Babraham Institute. Such ACs, after insertion into ungulate cells and generation of ungulates can produce heavy chain immunoglobulins. In one embodiment, these immunoglobulins can form functional heavy chain-light chain immunoglobulins. In another embodiment, these immunoglobulins can be expressed in an amount allowing for recovery from suitable cells or body fluids of the ungulate. Such immunoglobulins can be inserted into yeast artificial chromosome vectors, such as described by Burke, D T, Carle, G F and Olson, M V (1987) "Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors" Science, 236, 806-812, or by introduction of chromosome fragments (such as described by Richer, J and Lo, C W (1989) "Introduction of human DNA into mouse eggs by injection of dissected human chromosome fragments" Science 245, 175-177).

Additional information on specific ACs containing human immunoglobulin genes can be found in, for example, recent reviews by Giraldo & Montoliu (2001) Transgenic Research 10: 83-103 and Peterson (2003) Expert Reviews in Molecular Medicine 5: 1-25.

AC Transfer Methods

The human immunoglobulin genes can be first inserted into ACs and then the human-immunoglobulin-containing ACs can be inserted into the ungulate cells. Alternatively, the ACs

can be transferred to an intermediary mammalian cell, such as a CHO cell, prior to insertion into the ungulate cell. In one embodiment, the intermediary mammalian cell can also contain an AC and the first AC can be inserted into the AC of the mammalian cell. In particular, a YAC containing human immunoglobulin genes or fragments thereof in a yeast cell can be transferred to a mammalian cell that harbors an AC. The YAC can be inserted into the AC. The AC can then be transferred to an ungulate cell. The human Ig genes can be inserted into ACs by homologous recombination. The resulting AC containing human Ig genes, can then be introduced into ungulate cells. One or more ungulate cells can be selected by techniques described herein or those known in the art, which contain an AC containing a human Ig.

Suitable hosts for introduction of the ACs are provided herein, which include but are not limited to any animal or plant, cell or tissue thereof, including, but not limited to: mammals, birds, reptiles, amphibians, insects, fish, arachnids, tobacco, tomato, wheat, monocots, dicots and algae. In one embodiment, the ACs can be condensed (Marschall et al. *Gene Ther.* 1999 Sep;6(9):1634-7) by any reagent known in the art, including, but not limited to, spermine, spermidine, polyethylenimine, and/or polylysine prior to introduction into cells. The ACs can be introduced by cell fusion or microcell fusion or subsequent to isolation by any method known to those of skill in this art, including but not limited to: direct DNA transfer, electroporation, nuclear transfer, microcell fusion, cell fusion, spheroplast fusion, lipid-mediated transfer, lipofection, liposomes, microprojectile bombardment, microinjection, calcium phosphate precipitation and/or any other suitable method. Other methods for introducing DNA into cells, include nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells. Polycations, such as polybrene and polyornithine, may also be used. For various techniques for transforming mammalian cells, see e.g., Keown et al. *Methods in Enzymology* (1990) Vol.185, pp. 527-537; and Mansour et al. (1988) *Nature* 336:348-352.

The ACs can be introduced by direct DNA transformation; microinjection in cells or embryos, protoplast regeneration for plants, electroporation, microprojectile gun and other such methods known to one skilled in the art (see, e.g., Weissbach et al. (1988) *Methods for Plant Molecular Biology*, Academic Press, N.Y., Section VIII, pp. 421-463; Grierson et al. (1988) *Plant Molecular Biology*, 2d Ed., Blackie, London, Ch. 7-9; see, also U.S. Pat. Nos. 5,491,075; 5,482,928; and 5,424,409; see, also, e.g., U.S. Pat. No. 5,470,708,).

In particular embodiments, one or more isolated YACs can be used that harbor human Ig genes. The isolated YACs can be condensed (Marschall et al Gene Ther. 1999 Sep;6(9):1634-7) by any reagent known in the art, including, but not limited to spermine, spermidine, polyethylenimine, and/ or polylysine. The condensed YACs can then be transferred to porcine cells by any method known in the art (for example, microinjection, electroporation, lipid mediated transfection, etc). Alternatively, the condensed YAC can be transferred to oocytes via sperm-mediated gene transfer or intracytoplasmic sperm injection (ICSI) mediated gene transfer. In one embodiment, spheroplast fusion can be used to transfer YACs that harbor human Ig genes to porcine cells.

In other embodiments of the invention, the AC containing the human Ig can be inserted into an adult, fetal, or embryonic ungulate cell. Additional examples of ungulate cells include undifferentiated cells, such as embryonic cells (e.g., embryonic stem cells), differentiated or somatic cells, such as epithelial cells, neural cells epidermal cells, keratinocytes, hematopoietic cells, melanocytes, chondrocytes, B-lymphocytes, T-lymphocytes, erythrocytes, macrophages, monocytes, fibroblasts, muscle cells, cells from the female reproductive system, such as a mammary gland, ovarian cumulus, granulosa, or oviductal cell, germ cells, placental cell, or cells derived from any organ, such as the bladder, brain, esophagus, fallopian tube, heart, intestines, gallbladder, kidney, liver, lung, ovaries, pancreas, prostate, spinal cord, spleen, stomach, testes, thymus, thyroid, trachea, ureter, urethra, and uterus or any other cell type described herein.

Site Specific Recombinase Mediated Transfer

In particular embodiments of the present invention, the transfer of ACs containing human immunoglobulin genes to porcine cells, such as those described herein or known in the art, can be accomplished via site specific recombinase mediated transfer. In one particular embodiment, the ACs can be transferred into porcine fibroblast cells. In another particular embodiment, the ACs can be YACs.

In other embodiments of the present invention, the circularized DNA, such as an AC, that contain the site specific recombinase target site can be transferred into a cell line that has a site specific recombinase target site within its genome. In one embodiment, the cell's site specific recombinase target site can be located within an exogenous chromosome. The exogenous

chromosome can be an artificial chromosome that does not integrate into the host's endogenous genome. In one embodiment, the AC can be transferred via germ line transmission to offspring. In one particular embodiment, a YAC containing a human immunoglobulin gene or fragment thereof can be circularized via a site specific recombinase and then transferred into a host cell that contains a MAC, wherein the MAC contains a site specific recombinase site. This MAC that now contains human immunoglobulin loci or fragments thereof can then be fused with a porcine cell, such as, but not limited to, a fibroblast. The porcine cell can then be used for nuclear transfer.

In certain embodiments of the present invention, the ACs that contain human immunoglobulin genes or fragments thereof can be transferred to a mammalian cell, such as a CHO cell, prior to insertion into the ungulate cell. In one embodiment, the intermediary mammalian cell can also contain an AC and the first AC can be inserted into the AC of the mammalian cell. In particular, a YAC containing human immunoglobulin genes or fragments thereof in a yeast cell can be transferred to a mammalian cell that harbors a MAC. The YAC can be inserted in the MAC. The MAC can then be transferred to an ungulate cell. In particular embodiments, the YAC harboring the human Ig genes or fragments thereof can contain site specific recombinase target sites. The YAC can first be circularized via application of the appropriate site specific recombinase and then inserted into a mammalian cell that contains its own site specific recombinase target site. Then, the site specific recombinase can be applied to integrate the YAC into the MAC in the intermediary mammalian cell. The site specific recombinase can be applied in cis or trans. In particular, the site specific recombinase can be applied in trans. In one embodiment, the site specific recombinase can be expressed via transfection of a site specific recombinase expression plasmid, such as a Cre expression plasmid. In addition, one telomere region of the YAC can also be retrofitted with a selectable marker, such as a selectable marker described herein or known in the art. The human Ig genes or fragments thereof within the MAC of the intermediary mammalian cell can then be transferred to an ungulate cell, such as a fibroblast.

Alternatively, the AC, such as a YAC, harboring the human Ig genes or fragments thereof can contain site specific recombinase target sites optionally located near each telomere. The YAC can first be circularized via application of the appropriate site specific recombinase and then inserted into an ungulate cell directly that contains its own site specific recombinase target

site within its genome. Alternatively, the ungulate cell can harbor its own MAC, which contains a site specific recombinase target site. In this embodiment, the YAC can be inserted directly into the endogenous genome of the ungulate cell. In particular embodiments, the ungulate cell can be a fibroblast cell or any other suitable cell that can be used for nuclear transfer. See, for example, Figure 7; Call et al., Hum Mol Genet. 2000 Jul 22;9(12):1745-51.

In other embodiments, methods to circularize at least 100 kb of DNA are provided wherein the DNA can then be integrated into a host genome via a site specific recombinase. In one embodiment, at least 100, 200, 300, 400, 500, 1000, 2000, 5000, 10,000 kb of DNA can be circularized. In another embodiment, at least 1000, 2000, 5000, 10,000, or 20,000 megabases of DNA can be circularized. In one embodiment, the circularization of the DNA can be accomplished by attaching site specific recombinase target sites at each end of the DNA sequence and then applying the site specific recombinase to result in circularization of the DNA. In one embodiment, the site specific recombinase target site can be lox. In another embodiment, the site specific recombinase target site can be Flt. In certain embodiments, the DNA can be an artificial chromosome, such as a YAC or any AC described herein or known in the art. In another embodiment, the AC can contain human immunoglobulin loci or fragments thereof.

In another preferred embodiment, the YAC can be converted to, or integrated within, an artificial mammalian chromosome. The mammalian artificial chromosome is either transferred to or harbored within a porcine cell. The artificial chromosome can be introduced within the porcine genome through any method known in the art including but not limited to direct injection of metaphase chromosomes, lipid mediated gene transfer, or microcell fusion.

Site-specific recombinases include enzymes or recombinases that recognize and bind to a short nucleic acid site or sequence-specific recombinase target site, i.e., a recombinase recognition site, and catalyze the recombination of nucleic acid in relation to these sites. These enzymes include recombinases, transposases and integrases. Examples of sequence-specific recombinase target sites include, but are not limited to, lox sites, att sites, dif sites and frt sites. Non-limiting examples of site-specific recombinases include, but are not limited to, bacteriophage P1 Cre recombinase, yeast FLP recombinase, Inti integrase, bacteriophage λ , phi 80, P22, P2, 186, and P4 recombinase, Tn3 resolvase, the Hin recombinase, and the Cin recombinase, *E. coli* xerC and xerD recombinases, *Bacillus thuringiensis* recombinase, TpnI and the β -lactamase transposons, and the immunoglobulin recombinases.

In one embodiment, the recombination site can be a lox site that is recognized by the Cre recombinase of bacteriophage P1. Lox sites refer to a nucleotide sequence at which the product of the cre gene of bacteriophage P1, the Cre recombinase, can catalyze a site-specific recombination event. A variety of lox sites are known in the art, including the naturally occurring loxP, loxB, loxL and loxR, as well as a number of mutant, or variant, lox sites, such as loxP511, loxP514, lox.DELTA.86, lox.DELTA.117, loxC2, loxP2, loxP3 and lox P23. Additional example of lox sites include, but are not limited to, loxB, loxL, loxR, loxP, loxP3, loxP23, loxΔ86, loxΔ117, loxP511, and loxC2.

In another embodiment, the recombination site is a recombination site that is recognized by a recombinases other than Cre. In one embodiment, the recombinase site can be the FRT sites recognized by FLP recombinase of the 2 pi plasmid of *Saccharomyces cerevisiae*. FRT sites refer to a nucleotide sequence at which the product of the FLP gene of the yeast 2 micron plasmid, FLP recombinase, can catalyze site-specific recombination. Additional examples of the non-Cre recombinases include, but are not limited to, site-specific recombinases include: att sites recognized by the Int recombinase of bacteriophage λ (e.g. att1, att2, att3, attP, attB, attL, and attR), the recombination sites recognized by the resolvase family, and the recombination site recognized by transposase of *Bacillus thuringiensis*.

IV. Production of Genetically Modified Animals

In additional aspects of the present invention, ungulates that contain the genetic modifications described herein can be produced by any method known to one skilled in the art. Such methods include, but are not limited to: nuclear transfer, intracytoplasmic sperm injection, modification of zygotes directly and sperm mediated gene transfer.

In another embodiment, a method to clone such animals, for example, pigs, includes: enucleating an oocyte, fusing the oocyte with a donor nucleus from a cell in which at least one allele of at least one immunoglobulin gene has been inactivated, and implanting the nuclear transfer-derived embryo into a surrogate mother.

Alternatively, a method is provided for producing viable animals that lack any expression of functional immunoglobulin by inactivating both alleles of the immunoglobulin gene in embryonic stem cells, which can then be used to produce offspring.

In another aspect, the present invention provides a method for producing viable animals, such as pigs, in which both alleles of the immunoglobulin gene have been rendered inactive. In one embodiment, the animals are produced by cloning using a donor nucleus from a cell in which both alleles of the immunoglobulin gene have been inactivated. In one embodiment, both alleles of the immunoglobulin gene are inactivated via a genetic targeting event.

Genetically altered animals that can be created by modifying zygotes directly. For mammals, the modified zygotes can be then introduced into the uterus of a pseudopregnant female capable of carrying the animal to term. For example, if whole animals lacking an immunoglobulin gene are desired, then embryonic stem cells derived from that animal can be targeted and later introduced into blastocysts for growing the modified cells into chimeric animals. For embryonic stem cells, either an embryonic stem cell line or freshly obtained stem cells can be used.

In a suitable embodiment of the invention, the totipotent cells are embryonic stem (ES) cells. The isolation of ES cells from blastocysts, the establishing of ES cell lines and their subsequent cultivation are carried out by conventional methods as described, for example, by Doetchmann et al., *J. Embryol. Exp. Morph.* 87:27-45 (1985); Li et al., *Cell* 69:915-926 (1992); Robertson, E. J. "Tetracarcinomas and Embryonic Stem Cells: A Practical Approach," ed. E. J. Robertson, IRL Press, Oxford, England (1987); Wurst and Joyner, "Gene Targeting: A Practical Approach," ed. A. L. Joyner, IRL Press, Oxford, England (1993); Hogen et al., "Manipulating the Mouse Embryo: A Laboratory Manual," eds. Hogan, Beddington, Costantini and Lacy, Cold Spring Harbor Laboratory Press, New York (1994); and Wang et al., *Nature* 336:741-744 (1992). In another suitable embodiment of the invention, the totipotent cells are embryonic germ (EG) cells. Embryonic Germ cells are undifferentiated cells functionally equivalent to ES cells, that is they can be cultured and transfected in vitro, then contribute to somatic and germ cell lineages of a chimera (Stewart et al., *Dev. Biol.* 161:626-628 (1994)). EG cells are derived by culture of primordial germ cells, the progenitors of the gametes, with a combination of growth factors: leukemia inhibitory factor, steel factor and basic fibroblast growth factor (Matsui et al., *Cell* 70:841-847 (1992); Resnick et al., *Nature* 359:550-551 (1992)). The cultivation of EG cells can be carried out using methods described in the article by Donovan et al., "Transgenic Animals, Generation and Use," Ed. L. M. Houdebine, Harwood Academic Publishers (1997), and in the original literature cited therein.

Tetraploid blastocysts for use in the invention may be obtained by natural zygote production and development, or by known methods by electrofusion of two-cell embryos and subsequently cultured as described, for example, by James et al., *Genet. Res. Camb.* 60:185-194 (1992); Nagy and Rossant, "Gene Targeting: A Practical Approach," ed. A. L. Joyner, IRL Press, Oxford, England (1993); or by Kubiak and Tarkowski, *Exp. Cell Res.* 157:561-566 (1985).

The introduction of the ES cells or EG cells into the blastocysts can be carried out by any method known in the art. A suitable method for the purposes of the present invention is the microinjection method as described by Wang et al., *EMBO J.* 10:2437-2450 (1991).

Alternatively, by modified embryonic stem cells transgenic animals can be produced. The genetically modified embryonic stem cells can be injected into a blastocyst and then brought to term in a female host mammal in accordance with conventional techniques. Heterozygous progeny can then be screened for the presence of the alteration at the site of the target locus, using techniques such as PCR or Southern blotting. After mating with a wild-type host of the same species, the resulting chimeric progeny can then be cross-mated to achieve homozygous hosts.

After transforming embryonic stem cells with the targeting vector to alter the immunoglobulin gene, the cells can be plated onto a feeder layer in an appropriate medium, e.g., fetal bovine serum enhanced DMEM. Cells containing the construct can be detected by employing a selective medium, and after sufficient time for colonies to grow, colonies can be picked and analyzed for the occurrence of homologous recombination. Polymerase chain reaction can be used, with primers within and without the construct sequence but at the target locus. Those colonies which show homologous recombination can then be used for embryo manipulating and blastocyst injection. Blastocysts can be obtained from superovulated females. The embryonic stem cells can then be trypsinized and the modified cells added to a droplet containing the blastocysts. At least one of the modified embryonic stem cells can be injected into the blastocoel of the blastocyst. After injection, at least one of the blastocysts can be returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting litters screened for mutant cells having the construct. The blastocysts are selected for different parentage from the transformed ES cells. By providing for a different phenotype of the blastocyst and the ES cells, chimeric progeny can be readily detected, and then genotyping can be conducted to probe for the presence of the modified immunoglobulin gene.

In other embodiments, sperm mediated gene transfer can be used to produce the genetically modified ungulates described herein. The methods and compositions described herein to either eliminate expression of endogenous immunoglobulin genes or insert xenogenous immunoglobulin genes can be used to genetically modify the sperm cells via any technique described herein or known in the art. The genetically modified sperm can then be used to impregnate a female recipient via artificial insemination, intracytoplasmic sperm injection or any other known technique. In one embodiment, the sperm and/ or sperm head can be incubated with the exogenous nucleic acid for a sufficient time period. Sufficient time periods include, for example, about 30 seconds to about 5 minutes, typically about 45 seconds to about 3 minutes, more typically about 1 minute to about 2 minutes. In particular embodiments, the expression of xenogenous, such as human, immunoglobulin genes in ungulates as described herein, can be accomplished via intracytoplasmic sperm injection.

The potential use of sperm cells as vectors for gene transfer was first suggested by Brackett et al., *Proc., Natl. Acad. Sci. USA* 68:353-357 (1971). This was followed by reports of the production of transgenic mice and pigs after in vitro fertilization of oocytes with sperm that had been incubated by naked DNA (see, for example, Lavitrano et al., *Cell* 57:717-723 (1989) and Gandolfi et al. *Journal of Reproduction and Fertility Abstract Series* 4, 10 (1989)), although other laboratories were not able to repeat these experiments (see, for example, Brinster et al. *Cell* 59:239-241 (1989) and Gavora et al., *Canadian Journal of Animal Science* 71:287-291 (1991)). Since then, there have been several reports of successful sperm mediated gene transfer in chicken (see, for example, Nakanishi and Iritani, *Mol. Reprod. Dev.* 36:258-261 (1993)); mice (see, for example, Maione, *Mol. Reprod. Dev.* 59:406 (1998)); and pigs (see, for example, Lavitrano et al. *Transplant. Proc.* 29:3508-3509 (1997); Lavitrano et al., *Proc. Natl. Acad. Sci. USA* 99:14230-5 (2002); Lavitrano et al., *Mol. Reprod. Dev.* 64:284-91 (2003)). Similar techniques are also described in U.S. Pat. No. 6,376,743; issued Apr. 23, 2002; U.S. Patent Publication Nos. 20010044937, published Nov. 22, 2001, and 20020108132, published Aug. 8, 2002.

In other embodiments, intracytoplasmic sperm injection can be used to produce the genetically modified ungulates described herein. This can be accomplished by coinserting an exogenous nucleic acid and a sperm into the cytoplasm of an unfertilized oocyte to form a transgenic fertilized oocyte, and allowing the transgenic fertilized oocyte to develop into a transgenic embryo and, if desired, into a live offspring. The sperm can be a membrane-disrupted

sperm head or a demembranated sperm head. The coinsection step can include the substep of preincubating the sperm with the exogenous nucleic acid for a sufficient time period, for example, about 30 seconds to about 5 minutes, typically about 45 seconds to about 3 minutes, more typically about 1 minute to about 2 minutes. The coinsection of the sperm and exogenous nucleic acid into the oocyte can be via microinjection. The exogenous nucleic acid mixed with the sperm can contain more than one transgene, to produce an embryo that is transgenic for more than one transgene as described herein. The intracytoplasmic sperm injection can be accomplished by any technique known in the art, see, for example, US Patent No. 6,376,743. In particular embodiments, the expression of xenogenous, such as human, immunoglobulin genes in ungulates as described herein, can be accomplished via intracytoplasmic sperm injection.

Any additional technique known in the art may be used to introduce the transgene into animals. Such techniques include, but are not limited to pronuclear microinjection (see, for example, Hoppe, P. C. and Wagner, T. E., 1989, U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ lines (see, for example, Van der Putten et al., 1985, Proc. Natl. Acad. Sci., USA 82:6148-6152); gene targeting in embryonic stem cells (see, for example, Thompson et al., 1989, Cell 56:313-321; Wheeler, M. B., 1994, WO 94/26884); electroporation of embryos (see, for example, Lo, 1983, Mol Cell. Biol. 3:1803-1814); cell gun; transfection; transduction; retroviral infection; adenoviral infection; adenoviral-associated infection; liposome-mediated gene transfer; naked DNA transfer; and sperm-mediated gene transfer (see, for example, Lavitrano et al., 1989, Cell 57:717-723); etc. For a review of such techniques, see, for example, Gordon, 1989, Transgenic Animals, Intl. Rev. Cytol. 115:171-229. In particular embodiments, the expression of xenogenous, such as human, immunoglobulin genes in ungulates as described herein, can be accomplished via these techniques.

Somatic Cell Nuclear Transfer to Produce Cloned, Transgenic Offspring

In a further aspect of the present invention, ungulate, such as porcine or bovine, cells lacking one allele, optionally both alleles of an ungulate heavy chain, kappa light chain and/or lambda light chain gene can be used as donor cells for nuclear transfer into recipient cells to produce cloned, transgenic animals. Alternatively, ungulate heavy chain, kappa light chain and/or lambda light chain gene knockouts can be created in embryonic stem cells, which are then used to produce offspring. Offspring lacking a single allele of a functional ungulate heavy chain,

kappa light chain and/or lambda light chain gene produced according to the process, sequences and/or constructs described herein can be bred to further produce offspring lacking functionality in both alleles through mendelian type inheritance.

In another embodiment, the present invention provides a method for producing viable pigs that lack any expression of functional alpha-1,3-GT by breeding a male pig heterozygous for the alpha-1,3-GT gene with a female pig heterozygous for the alpha-1,3-GT gene. In one embodiment, the pigs are heterozygous due to the genetic modification of one allele of the alpha-1,3-GT gene to prevent expression of that allele. In another embodiment, the pigs are heterozygous due to the presence of a point mutation in one allele of the alpha-1,3-GT gene. In another embodiment, the point mutation can be a T-to-G point mutation at the second base of exon 9 of the alpha-1,3-GT gene. In one specific embodiment, a method to produce a porcine animal that lacks any expression of functional alpha-1,3-GT is provided wherein a male pig that contains a T-to-G point mutation at the second base of exon 9 of the alpha-1,3-GT gene is bred with a female pig that contains a T-to-G point mutation at the second base of exon 9 of the alpha-1,3-GT gene, or vice versa.

The present invention provides a method for cloning an animal, such as a pig, lacking a functional immunoglobulin gene via somatic cell nuclear transfer. In general, the animal can be produced by a nuclear transfer process comprising the following steps: obtaining desired differentiated cells to be used as a source of donor nuclei; obtaining oocytes from the animal; enucleating said oocytes; transferring the desired differentiated cell or cell nucleus into the enucleated oocyte, e.g., by fusion or injection, to form NT units; activating the resultant NT unit; and transferring said cultured NT unit to a host animal such that the NT unit develops into a fetus.

Nuclear transfer techniques or nuclear transplantation techniques are known in the art (Dai et al. *Nature Biotechnology* 20:251-255; Polejaeva et al *Nature* 407:86-90 (2000); Campbell et al, *Theriogenology*, 43:181 (1995); Collas et al, *Mol. Reprod Dev.*, 38:264-267 (1994); Keefer et al, *Biol. Reprod.*, 50:935-939 (1994); Sims et al, *Proc. Natl. Acad. Sci., USA*, 90:6143-6147 (1993); WO 94/26884; WO 94/24274, and WO 90/03432, U.S. Pat. Nos. 4,944,384 and 5,057,420).

A donor cell nucleus, which has been modified to alter the immunoglobulin gene, is transferred to a recipient oocyte. The use of this method is not restricted to a particular donor

cell type. The donor cell can be as described herein, see also, for example, Wilmut et al Nature 385 810 (1997); Campbell et al Nature 380 64-66 (1996); Dai et al., Nature Biotechnology 20:251-255, 2002 or Cibelli et al Science 280 1256-1258 (1998). All cells of normal karyotype, including embryonic, fetal and adult somatic cells which can be used successfully in nuclear transfer can be employed. Fetal fibroblasts are a particularly useful class of donor cells. Generally suitable methods of nuclear transfer are described in Campbell et al Theriogenology 43 181 (1995), Dai et al. Nature Biotechnology 20:251-255, Polejaeva et al Nature 407:86-90 (2000), Collas et al Mol. Reprod. Dev. 38 264-267 (1994), Keefer et al Biol. Reprod. 50 935-939 (1994), Sims et al Proc. Nat'l. Acad. Sci. USA 90 6143-6147 (1993), WO-A-9426884, WO-A-9424274, WO-A-9807841, WO-A-9003432, U.S. Pat. No. 4,994,384 and U.S. Pat. No. 5,057,420. Differentiated or at least partially differentiated donor cells can also be used. Donor cells can also be, but do not have to be, in culture and can be quiescent. Nuclear donor cells which are quiescent are cells which can be induced to enter quiescence or exist in a quiescent state in vivo. Prior art methods have also used embryonic cell types in cloning procedures (Campbell et al (Nature, 380:64-68, 1996) and Stice et al (Biol. Reprod., 20 54:100-110, 1996).

Somatic nuclear donor cells may be obtained from a variety of different organs and tissues such as, but not limited to, skin, mesenchyme, lung, pancreas, heart, intestine, stomach, bladder, blood vessels, kidney, urethra, reproductive organs, and a disaggregated preparation of a whole or part of an embryo, fetus, or adult animal. In a suitable embodiment of the invention, nuclear donor cells are selected from the group consisting of epithelial cells, fibroblast cells, neural cells, keratinocytes, hematopoietic cells, melanocytes, chondrocytes, lymphocytes (B and T), macrophages, monocytes, mononuclear cells, cardiac muscle cells, other muscle cells, extended cells, cumulus cells, epidermal cells or endothelial cells. In another embodiment, the nuclear donor cell is an embryonic stem cell. In a particular embodiment, fibroblast cells can be used as donor cells.

In another embodiment of the invention, the nuclear donor cells of the invention are germ cells of an animal. Any germ cell of an animal species in the embryonic, fetal, or adult stage may be used as a nuclear donor cell. In a suitable embodiment, the nuclear donor cell is an embryonic germ cell.

Nuclear donor cells may be arrested in any phase of the cell cycle (G0, G1, G2, S, M) so as to ensure coordination with the acceptor cell. Any method known in the art may be used to

manipulate the cell cycle phase. Methods to control the cell cycle phase include, but are not limited to, G0 quiescence induced by contact inhibition of cultured cells, G0 quiescence induced by removal of serum or other essential nutrient, G0 quiescence induced by senescence, G0 quiescence induced by addition of a specific growth factor; G0 or G1 quiescence induced by physical or chemical means such as heat shock, hyperbaric pressure or other treatment with a chemical, hormone, growth factor or other substance; S-phase control via treatment with a chemical agent which interferes with any point of the replication procedure; M-phase control via selection using fluorescence activated cell sorting, mitotic shake off, treatment with microtubule disrupting agents or any chemical which disrupts progression in mitosis (see also Freshney, R. I., "Culture of Animal Cells: A Manual of Basic Technique," Alan R. Liss, Inc, New York (1983).

Methods for isolation of oocytes are well known in the art. Essentially, this can comprise isolating oocytes from the ovaries or reproductive tract of an animal. A readily available source of oocytes is slaughterhouse materials. For the combination of techniques such as genetic engineering, nuclear transfer and cloning, oocytes must generally be matured in vitro before these cells can be used as recipient cells for nuclear transfer, and before they can be fertilized by the sperm cell to develop into an embryo. This process generally requires collecting immature (prophase I) oocytes from mammalian ovaries, e.g., bovine ovaries obtained at a slaughterhouse, and maturing the oocytes in a maturation medium prior to fertilization or enucleation until the oocyte attains the metaphase II stage, which in the case of bovine oocytes generally occurs about 18-24 hours post-aspiration. This period of time is known as the "maturation period". In certain embodiments, the oocyte is obtained from a gilt. A "gilt" is a female pig that has never had offspring. In other embodiments, the oocyte is obtained from a sow. A "sow" is a female pig that has previously produced offspring.

A metaphase II stage oocyte can be the recipient oocyte, at this stage it is believed that the oocyte can be or is sufficiently "activated" to treat the introduced nucleus as it does a fertilizing sperm. Metaphase II stage oocytes, which have been matured in vivo have been successfully used in nuclear transfer techniques. Essentially, mature metaphase II oocytes can be collected surgically from either non-superovulated or superovulated animal 35 to 48, or 39-41, hours past the onset of estrus or past the injection of human chorionic gonadotropin (hCG) or similar hormone. The oocyte can be placed in an appropriate medium, such as a hyaluronidase solution.

After a fixed time maturation period, which ranges from about 10 to 40 hours, about 16-18 hours, about 40-42 hours or about 39-41 hours, the oocytes can be enucleated. Prior to enucleation the oocytes can be removed and placed in appropriate medium, such as HECM containing 1 milligram per milliliter of hyaluronidase prior to removal of cumulus cells. The stripped oocytes can then be screened for polar bodies, and the selected metaphase II oocytes, as determined by the presence of polar bodies, are then used for nuclear transfer. Enucleation follows.

Enucleation can be performed by known methods, such as described in U.S. Pat. No. 4,994,384. For example, metaphase II oocytes can be placed in either HECM, optionally containing 7.5 micrograms per milliliter cytochalasin B, for immediate enucleation, or can be placed in a suitable medium, for example an embryo culture medium such as CR1aa, plus 10% estrus cow serum, and then enucleated later, such as not more than 24 hours later, or not more than 16-18 hours later.

Enucleation can be accomplished microsurgically using a micropipette to remove the polar body and the adjacent cytoplasm. The oocytes can then be screened to identify those of which have been successfully enucleated. One way to screen the oocytes is to stain the oocytes with 1 microgram per milliliter 33342 Hoechst dye in HECM, and then view the oocytes under ultraviolet irradiation for less than 10 seconds. The oocytes that have been successfully enucleated can then be placed in a suitable culture medium, for example, CR1aa plus 10% serum.

A single mammalian cell of the same species as the enucleated oocyte can then be transferred into the perivitelline space of the enucleated oocyte used to produce the NT unit. The mammalian cell and the enucleated oocyte can be used to produce NT units according to methods known in the art. For example, the cells can be fused by electrofusion. Electrofusion is accomplished by providing a pulse of electricity that is sufficient to cause a transient breakdown of the plasma membrane. This breakdown of the plasma membrane is very short because the membrane reforms rapidly. Thus, if two adjacent membranes are induced to breakdown and upon reformation the lipid bilayers intermingle, small channels can open between the two cells. Due to the thermodynamic instability of such a small opening, it enlarges until the two cells become one. See, for example, U.S. Pat. No. 4,997,384 by Prather et al. A variety of electrofusion media can be used including, for example, sucrose, mannitol, sorbitol and

phosphate buffered solution. Fusion can also be accomplished using Sendai virus as a fusogenic agent (Graham, Wister Inot. Symp. Monogr., 9, 19, 1969). Also, the nucleus can be injected directly into the oocyte rather than using electroporation fusion. See, for example, Collas and Barnes, Mol. Reprod. Dev., 38:264-267 (1994). After fusion, the resultant fused NT units are then placed in a suitable medium until activation, for example, CR1aa medium. Typically activation can be effected shortly thereafter, for example less than 24 hours later, or about 4-9 hours later, or optimally 1-2 hours after fusion. In a particular embodiment, activation occurs at least one hour post fusion and at 40-41 hours post maturation.

The NT unit can be activated by known methods. Such methods include, for example, culturing the NT unit at sub-physiological temperature, in essence by applying a cold, or actually cool temperature shock to the NT unit. This can be most conveniently done by culturing the NT unit at room temperature, which is cold relative to the physiological temperature conditions to which embryos are normally exposed. Alternatively, activation can be achieved by application of known activation agents. For example, penetration of oocytes by sperm during fertilization has been shown to activate prefusion oocytes to yield greater numbers of viable pregnancies and multiple genetically identical calves after nuclear transfer. Also, treatments such as electrical and chemical shock can be used to activate NT embryos after fusion. See, for example, U.S. Pat. No. 5,496,720, to Susko-Parrish et al. Fusion and activation can be induced by application of an AC pulse of 5 V for 5 s followed by two DC pulses of 1.5 kV/cm for 60 μ s each using an ECM2001 Electroculture Manipulator (BTX Inc., San Diego, CA). Additionally, activation can be effected by simultaneously or sequentially by increasing levels of divalent cations in the oocyte, and reducing phosphorylation of cellular proteins in the oocyte. This can generally be effected by introducing divalent cations into the oocyte cytoplasm, e.g., magnesium, strontium, barium or calcium, e.g., in the form of an ionophore. Other methods of increasing divalent cation levels include the use of electric shock, treatment with ethanol and treatment with caged chelators. Phosphorylation can be reduced by known methods, for example, by the addition of kinase inhibitors, e.g., serine-threonine kinase inhibitors, such as 6-dimethyl-aminopurine, staurosporine, 2-aminopurine, and sphingosine. Alternatively, phosphorylation of cellular proteins can be inhibited by introduction of a phosphatase into the oocyte, e.g., phosphatase 2A and phosphatase 2B.

The activated NT units, or "fused embryos", can then be cultured in a suitable in vitro culture medium until the generation of cell colonies. Culture media suitable for culturing and maturation of embryos are well known in the art. Examples of known media, which can be used for embryo culture and maintenance, include Ham's F-10+10% fetal calf serum (FCS), Tissue Culture Medium-199 (TCM-199)+10% fetal calf serum, Tyrodes-Albumin-Lactate-Pyruvate (TALP), Dulbecco's Phosphate Buffered Saline (PBS), Eagle's and Whitten's media, and, in one specific example, the activated NT units can be cultured in NCSU-23 medium for about 1-4 h at approximately 38.6°C in a humidified atmosphere of 5% CO₂.

Afterward, the cultured NT unit or units can be washed and then placed in a suitable media contained in well plates which can contain a suitable confluent feeder layer. Suitable feeder layers include, by way of example, fibroblasts and epithelial cells. The NT units are cultured on the feeder layer until the NT units reach a size suitable for transferring to a recipient female, or for obtaining cells which can be used to produce cell colonies. These NT units can be cultured until at least about 2 to 400 cells, about 4 to 128 cells, or at least about 50 cells.

Activated NT units can then be transferred (embryo transfers) to the oviduct of a female pig. In one embodiment, the female pig can be an estrus-synchronized recipient gilt. Crossbred gilts (large white/Duroc/Landrace) (280-400 lbs) can be used. The gilts can be synchronized as recipient animals by oral administration of 18-20 mg Regu-Mate (Altrenogest, Hoechst, Warren, NJ) mixed into the feed. Regu-Mate can be fed for 14 consecutive days. One thousand units of Human Chorionic Gonadotropin (hCG, Intervet America, Millsboro, DE) can then be administered i.m. about 105 h after the last Regu-Mate treatment. Embryo transfers can then be performed about 22-26 h after the hCG injection. In one embodiment, the pregnancy can be brought to term and result in the birth of live offspring. In another embodiment, the pregnancy can be terminated early and embryonic cells can be harvested.

Breeding for Desired Homozygous Knockout Animals

In another aspect, the present invention provides a method for producing viable animals that lack any expression of a functional immunoglobulin gene is provided by breeding a male heterozygous for the immunoglobulin gene with a female heterozygous for the immunoglobulin gene. In one embodiment, the animals are heterozygous due to the genetic modification of one allele of the immunoglobulin gene to prevent expression of that allele. In another embodiment,

the animals are heterozygous due to the presence of a point mutation in one allele of the alpha-immunoglobulin gene. In further embodiments, such heterozygous knockouts can be bred with an ungulate that expresses xenogenous immunoglobulin, such as human. In one embodiment, a animal can be obtained by breeding a transgenic ungulate that lacks expression of at least one allele of an endogenous immunoglobulin wherein the immunoglobulin is selected from the group consisting of heavy chain, kappa light chain and lambda light chain or any combination thereof with an ungulate that expresses an xenogenous immunoglobulin. In another embodiment, a animal can be obtained by breeding a transgenic ungulate that lacks expression of one allele of heavy chain, kappa light chain and lambda light chain with an ungulate that expresses an xenogenous, such as human, immunoglobulin. In a further embodiment, an animal can be obtained by breeding a transgenic ungulate that lacks expression of one allele of heavy chain, kappa light chain and lambda light chain and expresses an xenogenous, such as human, immunoglobulin with another transgenic ungulate that lacks expression of one allele of heavy chain, kappa light chain and lambda light chain with an ungulate and expresses an xenogenous, such as human, immunoglobulin to produce a homozygous transgenic ungulate that lacks expression of both alleles of heavy chain, kappa light chain and lambda light chain and expresses an xenogenous, such as human, immunoglobulin. Methods to produce such animals are also provided.

In one embodiment, sexually mature animals produced from nuclear transfer from donor cells that carrying a double knockout in the immunoglobulin gene, can be bred and their offspring tested for the homozygous knockout. These homozygous knockout animals can then be bred to produce more animals.

In another embodiment, oocytes from a sexually mature double knockout animal can be in vitro fertilized using wild type sperm from two genetically diverse pig lines and the embryos implanted into suitable surrogates. Offspring from these matings can be tested for the presence of the knockout, for example, they can be tested by cDNA sequencing, and/ or PCR. Then, at sexual maturity, animals from each of these litters can be mated. In certain methods according to this aspect of the invention, pregnancies can be terminated early so that fetal fibroblasts can be isolated and further characterized phenotypically and/or genotypically. Fibroblasts that lack expression of the immunoglobulin gene can then be used for nuclear transfer according to the

methods described herein to produce multiple pregnancies and offspring carrying the desired double knockout.

Additional Genetic Modifications

In other embodiments, animals or cells lacking expression of functional immunoglobulin, produced according to the process, sequences and/or constructs described herein, can contain additional genetic modifications to eliminate the expression of xenoantigens. The additional genetic modifications can be made by further genetically modifying cells obtained from the transgenic cells and animals described herein or by breeding the animals described herein with animals that have been further genetically modified. Such animals can be modified to eliminate the expression of at least one allele of the alpha-1,3-galactosyltransferase gene, the CMP-Neu5Ac hydroxylase gene (see, for example, USSN 10/863,116), the iGb3 synthase gene (see, for example, U.S. Patent Application 60/517,524), and/or the Forssman synthase gene (see, for example, U.S. Patent Application 60/568,922). In additional embodiments, the animals disclosed herein can also contain genetic modifications to express fucosyltransferase, sialyltransferase and/ or any member of the family of glucosyltransferases. To achieve these additional genetic modifications, in one embodiment, cells can be modified to contain multiple genetic modifications. In other embodiments, animals can be bred together to achieve multiple genetic modifications. In one specific embodiment, animals, such as pigs, lacking expression of functional immunoglobulin, produced according to the process, sequences and/or constructs described herein, can be bred with animals, such as pigs, lacking expression of alpha-1,3-galactosyl transferase (for example, as described in WO 04/028243).

In another embodiment, the expression of additional genes responsible for xenograft rejection can be eliminated or reduced. Such genes include, but are not limited to the CMP-NEUAc Hydroxylase Gene, the isoGloboside 3 Synthase gene, and the Forssman synthase gene. In addition, genes or cDNA encoding complement related proteins, which are responsible for the suppression of complement mediated lysis can also be expressed in the animals and tissues of the present invention. Such genes include, but are not limited to CD59, DAF, MCP and CD46 (see, for example, WO 99/53042; *Chen et al. Xenotransplantation*, Volume 6 Issue 3 Page 194 - August 1999, which describes pigs that express CD59/DAF transgenes; Costa C et al, *Xenotransplantation*. 2002 Jan;9(1):45-57, which describes transgenic pigs that express human

CD59 and H-transferase; Zhao L et al.; Diamond LE et al. Transplantation. 2001 Jan 15;71(1):132-42, which describes a human CD46 transgenic pigs.

Additional modifications can include expression of tissue factor pathway inhibitor (TFPI), heparin, antithrombin, hirudin, TFPI, tick anticoagulant peptide, or a snake venom factor, such as described in WO 98/42850 and US Patent No. 6,423,316, entitled "Anticoagulant fusion protein anchored to cell membrane"; or compounds, such as antibodies, which down-regulate the expression of a cell adhesion molecule by the cells, such as described in WO 00/31126, entitled "Suppression of xenograft rejection by down regulation of a cell adhesion molecules" and compounds in which co-stimulation by signal 2 is prevented, such as by administration to the organ recipient of a soluble form of CTLA-4 from the xenogeneic donor organism, for example as described in WO 99/57266, entitled "Immunosuppression by blocking T cell co-stimulation signal 2 (B7/CD28 interaction)".

Certain aspects of the invention are described in greater detail in the non-limiting Examples that follow.

EXAMPLES

EXAMPLE 1: Porcine Heavy Chain Targeting and Generation of Porcine Animals that Lack Expression of Heavy Chain

A portion of the porcine Ig heavy-chain locus was isolated from a 3X redundant porcine BAC library. In general, BAC libraries can be generated by fragmenting pig total genomic DNA, which can then be used to derive a BAC library representing at least three times the genome of the whole animal. BACs that contain porcine heavy chain immunoglobulin can then be selected through hybridization of probes selective for porcine heavy chain immunoglobulin as described herein.

Sequence from a clone (Seq ID 1) was used to generate a primer complementary to a portion of the J-region (the primer is represented by Seq ID No. 2). Separately, a primer was designed that was complementary to a portion of Ig heavy-chain mu constant region (the primer

is represented by Seq ID No. 3). These primers were used to amplify a fragment of porcine Ig heavy-chain (represented by Seq ID No. 4) that led the functional joining region (J-region) and sufficient flanking region to design and build a targeting vector. To maintain this fragment and subclones of this fragment in a native state, the *E. coli* (Stable 2, Invitrogen cat #1026-019) that harbored these fragments was maintained at 30°C. Regions of Seq. ID No. 4 were subcloned and used to assemble a targeting vector as shown in Seq. ID No. 5. This vector was transfected into porcine fetal fibroblasts that were subsequently subjected to selection with G418. Resulting colonies were screened by PCR to detect potential targeting events (Seq ID No. 6 and Seq ID No. 7, 5' screen primers; and Seq ID No. 8 and Seq ID No. 9, 3' screen primers). See Figure 1 for a schematic illustrating the targeting. Targeting was confirmed by southern blotting. Piglets were generated by nuclear transfer using the targeted fetal fibroblasts as nuclear donors.

Nuclear Transfer.

The targeted fetal fibroblasts were used as nuclear donor cells. Nuclear transfer was performed by methods that are well known in the art (see, e.g., Dai et al., *Nature Biotechnology* 20: 251-255, 2002; and Polejaeva et al., *Nature* 407:86-90, 2000).

Oocytes were collected 46-54 h after the hCG injection by reverse flush of the oviducts using pre-warmed Dulbecco's phosphate buffered saline (PBS) containing bovine serum albumin (BSA; 4 g l⁻¹) (as described in Polejaeva, I.A., et al. (*Nature* 407, 86-90 (2000))). Enucleation of *in vitro*-matured oocytes (BioMed, Madison, WI) was begun between 40 and 42 hours post-maturation as described in Polejaeva, I.A., et al. (*Nature* 407, 86-90 (2000))). Recovered oocytes were washed in PBS containing 4 g l⁻¹ BSA at 38°C, and transferred to calcium-free phosphate-buffered NCSU-23 medium at 38°C for transport to the laboratory. For enucleation, we incubated the oocytes in calcium-free phosphate-buffered NCSU-23 medium containing 5 µg ml⁻¹ cytochalasin B (Sigma) and 7.5 µg ml⁻¹ Hoechst 33342 (Sigma) at 38°C for 20 min. A small amount of cytoplasm from directly beneath the first polar body was then aspirated using an 18 µM glass pipette (Humagen, Charlottesville, Virginia). We exposed the aspirated karyoplast to ultraviolet light to confirm the presence of a metaphase plate.

For nuclear transfer, a single fibroblast cell was placed under the zona pellucida in contact with each enucleated oocyte. Fusion and activation were induced by application of an AC pulse of 5 V for 5 s followed by two DC pulses of 1.5 kV/cm for 60 µs each using an ECM2001 Electroculture Manipulator (BTX Inc., San Diego, CA). Fused embryos were cultured in

NCSU-23 medium for 1-4 h at 38.6°C in a humidified atmosphere of 5% CO₂, and then transferred to the oviduct of an estrus-synchronized recipient gilt. Crossbred gilts (large white/Duroc/landrace) (280-400 lbs) were synchronized as recipients by oral administration of 18-20 mg Regu-Mate (Altrenogest, Hoechst, Warren, NJ) mixed into their feed. Regu-Mate was fed for 14 consecutive days. Human chorionic gonadotropin (hCG, 1,000 units; Intervet America, Millsboro, DE) was administered intra-muscularly 105 h after the last Regu-Mate treatment. Embryo transfers were done 22-26 h after the hCG injection.

Nuclear transfer produced 18 healthy piglets from four litters. These animals have one functional wild-type Ig heavy-chain locus and one disrupted Ig heavy chain locus.

Seq ID 2: primer from Butler subclone to amplify J to C heavychain (637Xba5')	ggccagacttctcggaaacagctca
Seq ID 3: primer for C to amplify J to C heavychain (JM1L)	ttccaggagaaggtgacggagct
Seq ID 6: heavychain 5' primer for 5' screen (HCKOXba5'2)	tctagaagacgctggagagaggccag
Seq ID 7: heavychain 3' primer for 5' screen (5'arm5')	taaagcgcattgctccagactgcctt
Seq ID 8: heavychain 5' primer for 3' screen (NEO4425)	catgccttctatgccttctt
Seq ID 9: heavychain 3' primer for 3' screen (650+CA)	Aagtactgcccgcctctcagga

Southern blot analysis of cell and pig tissue samples. Cells or tissue samples were lysed overnight at 60°C in lysis buffer (10mM Tris, pH 7.5, 10mM EDTA, 10mM NaCl, 0.5% (w/v) Sarcosyl, 1 mg/ml proteinase K) and the DNA precipitated with ethanol. The DNA was then

digested with NcoI or XbaI, depending on the probe to be used, and separated on a 1% agarose gel. After electrophoresis, the DNA was transferred to a nylon membrane and probed with digoxigenin-labeled probe (SEQ ID No 41 for NcoI digest, SEQ ID No 40 for XbaI digest). Bands were detected using a chemiluminescent substrate system (Roche Molecular Biochemicals).

Probes for Heavy Chain Southern:

HC J Probe (used with Xba I digest)

CTCTGCACTCACTACCGCCGGACGCGCACTGCCGTGCTGCCCCATGGACCA
CGCTGGGGAGGGGTGAGCGGACAGCACGTTAGGAAGTGTGTGTGTGCGCG
TGGGTGCAAGTCGAGCCAAGGCCAAGATCCAGGGGCTGGGCCCTGTGCCC
AGAGGAGAATGGCAGGTGGAGTGTAGCTGGATTGAAAGGTGGCCTGAAGG
GTGGGGCATCCTGTTTGGAGGCTCACTCTCAGCCCCAGGGTCTCTGGTTC
CTGCCGGGGTGGGGGGCGCAAGGTGCCTACCACACCCTGCTAGCCCCTCG
TCCAGTCCCGGGCCTGCCTCTTCACCACGGAAGAGGATAAGCCAGGCTGC
AGGCTTCATGTGCGCCGTGGAGAACCCAGTTCGGCCCTTGAGAG (Seq ID No 40)

HC Mu Probe (used with NcoI digest)

GGCTGAAGTCTGAGGCCTGGCAGATGAGCTTGGACGTGCGCTGGGGAGTA
CTGGAGAAGGACTCCCGGGTGGGGACGAAGATGTTCAAGACGGGGGGGCTG
CTCCTCTACGACTGCAGGCAGGAACGGGGCGTCACTGTGCCGGCGGCACC
CGGCCCCGCCCCCGCCACAGCCACAGGGGGAGCCCAGCTCACCTGGCCCA
GAGATGGACACGGAAGTGGTGCCACTGGGGTGCTGGACCTCGCACACCAG
GAAGGCCTCTGGGTCCTGGGGGATGCTCACAGAGGGTAGGAGCACCCGGG
AGGAGGCCAAGTACTTGCCGCCTCTCAGGACGG (Seq ID No 41)

EXAMPLE 2: Porcine Kappa Light Chain Targeting and Generation of Porcine Lacking Expression of Kappa Light Chain

A portion of the porcine Ig kappa-chain locus was isolated from a 3X redundant porcine BAC library. In general, BAC libraries can be generated by fragmenting pig total genomic DNA, which can then be used to derive a BAC library representing at least three times the

genome of the whole animal. BACs that contain porcine kappa chain immunoglobulin can then be selected through hybridization of probes selective for porcine kappa chain immunoglobulin as described herein.

A fragment of porcine Ig light-chain kappa was amplified using a primer complementary to a portion of the J-region (the primer is represented by Seq ID No. 10) and a primer complementary to a region of kappa C-region (represented by Seq ID No. 11). The resulting amplicon was cloned into a plasmid vector and maintained in Stable2 cells at 30°C (Seq ID No. 12). See Figure 2 for a schematic illustration.

Separately, a fragment of porcine Ig light-chain kappa was amplified using a primer complementary to a portion of the C-region (Seq ID No. 13) and a primer complementary to a region of the kappa enhancer region (Seq ID No. 14). The resulting amplicon was fragmented by restriction enzymes and DNA fragments that were produced were cloned, maintained in Stable2 cells at 30 degrees C and sequenced. As a result of this sequencing, two non-overlapping contigs were assembled (Seq ID No. 15, 5' portion of amplicon; and Seq ID No. 16, 3' portion of amplicon). Sequence from the downstream contig (Seq ID No. 16) was used to design a set of primers (Seq ID No. 17 and Seq ID No. 18) that were used to amplify a contiguous fragment near the enhancer (Seq ID No. 19). A subclone of each Seq ID No. 12 and Seq ID No. 19 were used to build a targeting vector (Seq ID No. 20). This vector was transfected into porcine fetal fibroblasts that were subsequently subjected to selection with G418. Resulting colonies were screened by PCR to detect potential targeting events (Seq ID No. 21 and Seq ID No. 22, 5' screen primers; and Seq ID No. 23 and Seq ID No. 24, 3' screen primers, and Seq ID No. 24 and Seq ID No. 24, endogenous screen primers). Targeting was confirmed by southern blotting. Southern blot strategy design was facilitated by cloning additional kappa sequence, it corresponds to the template for germline kappa transcript (Seq ID No. 25). Fetal pigs were generated by nuclear transfer.

Nuclear Transfer.

The targeted fetal fibroblasts were used as nuclear donor cells. Nuclear transfer was performed by methods that are well known in the art (see, e.g., Dai et al., Nature Biotechnology 20: 251-255, 2002; and Polejaeva et al., Nature 407:86-90, 2000).

Oocytes were collected 46-54 h after the hCG injection by reverse flush of the oviducts using pre-warmed Dulbecco's phosphate buffered saline (PBS) containing bovine serum albumin

(BSA; 4 g l⁻¹) (as described in Polejaeva, I.A., *et al.* (*Nature* 407, 86-90 (2000))). Enucleation of *in vitro*-matured oocytes (BioMed, Madison, WI) was begun between 40 and 42 hours post-maturation as described in Polejaeva, I.A., *et al.* (*Nature* 407, 86-90 (2000))). Recovered oocytes were washed in PBS containing 4 g l⁻¹ BSA at 38°C, and transferred to calcium-free phosphate-buffered NCSU-23 medium at 38°C for transport to the laboratory. For enucleation, we incubated the oocytes in calcium-free phosphate-buffered NCSU-23 medium containing 5 µg ml⁻¹ cytochalasin B (Sigma) and 7.5 µg ml⁻¹ Hoechst 33342 (Sigma) at 38°C for 20 min. A small amount of cytoplasm from directly beneath the first polar body was then aspirated using an 18 µM glass pipette (Humagen, Charlottesville, Virginia). We exposed the aspirated karyoplast to ultraviolet light to confirm the presence of a metaphase plate.

For nuclear transfer, a single fibroblast cell was placed under the zona pellucida in contact with each enucleated oocyte. Fusion and activation were induced by application of an AC pulse of 5 V for 5 s followed by two DC pulses of 1.5 kV/cm for 60 µs each using an ECM2001 Electrocell Manipulator (BTX Inc., San Diego, CA). Fused embryos were cultured in NCSU-23 medium for 1-4 h at 38.6°C in a humidified atmosphere of 5% CO₂, and then transferred to the oviduct of an estrus-synchronized recipient gilt. Crossbred gilts (large white/Duroc/landrace) (280-400 lbs) were synchronized as recipients by oral administration of 18-20 mg Regu-Mate (Altrenogest, Hoechst, Warren, NJ) mixed into their feed. Regu-Mate was fed for 14 consecutive days. Human chorionic gonadotropin (hCG, 1,000 units; Intervet America, Millsboro, DE) was administered intra-muscularly 105 h after the last Regu-Mate treatment. Embryo transfers were done 22-26 h after the hCG injection.

Nuclear transfer using kappa targeted cells produced 33 healthy pigs from 5 litters. These pigs have one functional wild-type allele of porcine Ig light-chain kappa and one disrupted Ig light-chain kappa allele.

Seq ID 10: kappa J to C 5' primer (kjc5'1)	caaggaqaccaagctggaactc
Seq ID 11: kappa J to C 3' primer (kjc3'2)	tgatcaagcacaccacagagacag
Seq ID 13: 5' primer for Kappa C to E (porKCS1)	gatgccaagccatccgtcttcac

Seq ID 14: 3' primer for Kappa C to E (porKCA1)	tgaccaaagcagtgtgacggtgc
Seq ID 17: kappa 5' primer for amplification of enhancer region (K3'arm1S)	ggatcaaacacgcacccatggac
Seq ID 18: kappa 3' primer for amplification of enhancer region (K3'arm1A)	ggtgattggggcatggtgagg
Seq ID 21: kappa screen, 5' primer, 5' (kappa5armS)	cgaacccctgtgtatagtt
Seq ID 22: kappa screen, 3' primer, 5' (kappaNeoA)	gagatgaggaagaggagaaca
Seq ID 23: kappa screen, 5' primer, 3' (kappaNeoS)	gcattgtctgagtaggtgcatt
Seq ID 24: kappa screen, 3' primer, 5' (kappa5armProbe3')	cgttcttcaggggaacacgat
Seq ID No 43, Kappa screen, 3' primer (kappa3armA2)	GTCTTTGGTTTTTGCTGAGGGTT

Southern blot analysis of cell and pig tissue samples. Cells or tissue samples were lysed overnight at 60°C in lysis buffer (10mM Tris, pH 7.5, 10mM EDTA, 10mM NaCl, 0.5% (w/v) Sarcosyl, 1 mg/ml proteinase K) and the DNA precipitated with ethanol. The DNA was then digested with SacI and separated on a 1% agarose gel. After electrophoresis, the DNA was transferred to a nylon membrane and probed with digoxigenin-labeled probe (SEQ ID No 42). Bands were detected using a chemiluminescent substrate system (Roche Molecular Biochemicals).

Probe for Kappa Southern:

Kappa5ArmProbe 5'/3'

gaagtgaagccagccagttcctcctgggcaggtggccaaaattacagttg
 acccctcctggtctggctgaaccttgcccatatggtgacagccatctgg
 ccagggcccaggtctccctctgaagccttgggaggagaggagagtggtgc
 tggcccgatcacagatgcggaaggggtgactcctcaaccggggtgcaga
 ctctgcaggggtgggtctgggccaacacacccaaagcacgccaggaagg
 aaaggcagcttggtatcactgccagagctaggagaggcaccgggaaaat
 gatctgtccaagaccggtcttcttctaaactccgaggggggtcagatga
 agtggttttgttcttggcctgaagcatcgtgtccctgcaagaagcgg (SEQ ID No 42)

EXAMPLE 3**Characterization of the Porcine Lambda Gene Locus**

To disrupt or disable porcine lambda, a targeting strategy has been devised that allows for the removal or disruption of the region of the lambda locus that includes a concatamer of J to C expression cassettes. BAC clones that contain portions of the porcine genome can be generated. A portion of the porcine Ig lambda-chain locus was isolated from a 3X redundant porcine BAC library. In general, BAC libraries can be generated by fragmenting pig total genomic DNA, which can then be used to derive a BAC library representing at least three times the genome of the whole animal. BACs that contain porcine lambda chain immunoglobulin can then be selected through hybridization of probes selective for porcine lambda chain immunoglobulin as described herein.

BAC clones containing a lambda J-C flanking region (see Figure 3), can be independently fragmented and subcloned into a plasmid vector. Individual subclones have been screened by PCR for the presence of a portion of the J to C intron. We have cloned several of these cassettes by amplifying from one C region to the next C region. This amplification was accomplished by using primers that are oriented to allow divergent extension within any one C region (Seq ID 26 and Seq ID 27). To obtain successful amplification, the extended products converge with extended products originated from adjacent C regions (as opposed to the same C region). This strategy produces primarily amplimers that extend from one C to the adjacent C. However, some amplimers are the result of amplification across the adjacent C and into the next C which lies beyond the adjacent C. These multi-gene amplimers contain a portion of a C, both the J and C region of the next J-C unit, the J region of the third J-C unit, and a portion of the C

region of the third J-C unit. Seq ID 28 is one such amplimer and represents sequence that must be removed or disrupted.

Other porcine lambda sequences that have been cloned include: Seq ID No. 32, which includes 5' flanking sequence to the first lambda J/C region of the porcine lambda light chain genomic sequence; Seq ID No. 33, which includes 3' flanking sequence to the J/C cluster region of the porcine lambda light chain genomic sequence, from approximately 200 base pairs downstream of lambda J/C; Seq ID No. 34, which includes 3' flanking sequence to the J/C cluster region of the porcine lambda light chain genomic sequence, approximately 11.8 Kb downstream of the J/C cluster, near the enhancer; Seq ID No. 35, which includes approximately 12 Kb downstream of lambda, including the enhancer region; Seq ID No. 36, which includes approximately 17.6 Kb downstream of lambda; Seq ID No. 37, which includes approximately 19.1 Kb downstream of lambda; Seq ID No. 38, which includes approximately 21.3 Kb downstream of lambda; and Seq ID No. 39, which includes approximately 27 Kb downstream of lambda.

Seq ID 26: 5' primer for lambda C to C amplimer (lamC5')	ccttcctcctgcacctgtcaac
Seq ID 27: 3' primer for lambda C to C amplimer (lamC3')	tagacacaccagggtggccttg

Example 4

Production of Targeting Vectors for the Lambda Gene

In one example, a vector has been designed and built with one targeting arm that is homologous to a region upstream of J1 and the other arm homologous to a region that is downstream of the last C (see Figure 4). One targeting vector is designed to target upstream of J1. This targeting vector utilizes a selectable marker that can be selected for or against. Any combination of positive and negative selectable markers described herein or known in the art can be used. A fusion gene composed of the coding region of Herpes simplex thymidine kinase (TK) and the Tn5 aminoglycoside phosphotransferase (Neo resistance) genes is used. This fusion

gene is flanked by recognition sites for any site specific recombinase (SSRRS) described herein or known in the art, such as lox sites. Upon isolation of targeted cells through the use of G418 selection, Cre is supplied in trans to delete the marker gene (See Figure 5). Cells that have deleted the marker gene are selected by addition of any drug known in the art that can be metabolized by TK into a toxic product, such as ganciclovir. The resulting genotype is then targeted with a second vector. The second targeting vector (Figure 6) is designed to target downstream of last C and uses a positive/negative selection system that is flanked on only one side by a specific recombination site (lox). The recombination site is placed distally in relation to the first targeting event. Upon isolation of the targeted genotype, Cre is again supplied in trans to mediate deletion from recombination site provided in the first targeting event to the recombination site delivered in the second targeting event. The entire J to C cluster will be removed. The appropriate genotype is again selected by administration of ganciclovir.

In another example, insertional targeting vectors are used to disrupt each C regions independently. An insertional targeting vector will be designed and assembled to disrupt individual C region genes. There are at least 3 J to C regions in the J-C cluster. We will begin the process by designing vectors to target the first and last C regions and will include in the targeting vector site-specific recombination sites. Once both insertions have been made, the intervening region will be deleted with the site-specific recombinase.

Example 5: Crossbreeding of Heavy chain single knockout with Kappa single knockout pigs.

To produce pigs that have both one disrupted Ig heavy chain locus and one disrupted Ig light-chain kappa allele, single knockout animals were crossbred. The first pregnancy yielded four fetuses, two of which screened positive by both PCR and Southern for both heavy-chain and kappa targeting events (see examples 1 and 2 for primers). Fetal fibroblasts were isolated, expanded and frozen. A second pregnancy resulting from the mating of a kappa single knockout with a heavy chain single knockout produced four healthy piglets.

Fetal fibroblast cells that contain a heavy chain single knockout and a kappa chain single knockout will be used for further targeting. Such cells will be used to target the lambda locus via the methods and compositions described herein. The resulting offspring will be heterozygous

knockouts for heavy chain, kappa chain and lambda chain. These animals will be further crossed with animals containing the human Ig genes as described herein and then crossbred with other single Ig knockout animals to produce porcine Ig double knockout animals with human Ig replacement genes.

This invention has been described with reference to its preferred embodiments. Variations and modifications of the invention, will be obvious to those skilled in the art from the foregoing detailed description of the invention.

DEMANDES OU BREVETS VOLUMINEUX

**LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS
COMPREND PLUS D'UN TOME.**

CECI EST LE TOME __1__ DE __2__

NOTE: Pour les tomes additionels, veuillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

**THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE
THAN ONE VOLUME.**

THIS IS VOLUME __1__ OF __2__

NOTE: For additional volumes please contact the Canadian Patent Office.

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. A targeting vector comprising: (a) a first nucleotide sequence comprising at least 17, 20, 25 or 30 contiguous nucleic acids identical to a nucleotide sequence encoding a porcine kappa light chain immunoglobulin; (b) a selectable marker gene; and (c) a second nucleotide sequence comprising at least 17, 20, 25 or 30 contiguous nucleic acids identical to a nucleotide sequence encoding a porcine kappa light chain immunoglobulin, which does not overlap with the first nucleotide sequence, wherein the kappa light chain sequence comprises SEQ ID No 30.
2. The targeting vector of claim 1 wherein the selectable marker comprises an antibiotic resistance gene.
3. The targeting vector of claim 1 wherein the first nucleotide sequence comprises a 5' recombination arm.
4. The targeting vector of claim 1 wherein the second nucleotide sequence comprises a 3' recombination arm.
5. A cell transfected with the targeting vector of claim 1.
6. The cell of claim 5 wherein at least one allele of a kappa chain immunoglobulin locus has been rendered inactive.
7. A targeting vector comprising: (a) a first nucleotide sequence comprising at least 17, 20, 25 or 30 contiguous nucleic acids identical to a nucleotide sequence encoding a porcine lambda light chain immunoglobulin; (b) a selectable marker gene; and (c) a second nucleotide sequence comprising at least 17, 20, 25 or 30 contiguous nucleic acids identical to a nucleotide sequence encoding a porcine lambda light chain immunoglobulin, which does not overlap with the first nucleotide sequence, wherein the lambda light chain sequence comprises SEQ ID No 28.
8. A cell transfected with the targeting vector of claim 7.

9. The cell of claim 8 wherein at least one allele of a lambda chain immunoglobulin locus has been rendered inactive.
10. An isolated nucleic acid molecule comprising a nucleotide sequence encoding a porcine kappa light chain immunoglobulin locus comprising at least 17, 20, 25 or 30 contiguous nucleotides of SEQ ID No 30.
11. The nucleic acid molecule of claim 10 wherein the kappa light chain immunoglobulin locus comprises at least one joining region, one constant region and/or one enhancer region.
12. The nucleic acid molecule of claim 10 wherein the nucleotide sequence comprises at least five joining regions, one constant region and/or one enhancer region.
13. The nucleic acid molecule of claim 10 comprising SEQ ID No 30.
14. An isolated nucleic acid molecule comprising a nucleotide sequence encoding a porcine lambda light chain immunoglobulin locus or fragment thereof, wherein said nucleotide sequence comprises at least 17, 20, 25 or 30 contiguous nucleotides of SEQ ID No 28.
15. The nucleic acid molecule of claim 14 wherein the lambda light chain immunoglobulin locus comprises a concatamer of J to C units.
16. The nucleic acid molecule of claim 14 wherein the lambda light chain immunoglobulin locus comprises at least one joining region-constant region pair and/or at least one variable region.
17. The nucleic acid molecule of claim 14 wherein said nucleotide sequence comprises SEQ ID No 28.
18. An isolated nucleic acid molecule comprising a nucleotide sequence encoding a porcine kappa light chain immunoglobulin thereof wherein the sequence is at least 80, 85, 90, 95, 98 or 99% identical to SEQ ID No 30 over the entire length of SEQ ID No 30.

19. An isolated nucleic acid molecule comprising a nucleotide sequence encoding a porcine lambda light chain immunoglobulin wherein the sequence is at least 80, 85, 90, 95, 98 or 99% identical to SEQ ID No 28 over the entire length of SEQ ID No 28.
20. An isolated nucleic acid molecule comprising a nucleotide sequence encoding a porcine kappa light chain immunoglobulin locus or fragment thereof, wherein said nucleotide sequence comprises at least 17, 20, 25 or 30 contiguous nucleotides of SEQ ID No 12.
21. The nucleic acid molecule of claim 20 comprising SEQ ID No 12.
22. An isolated nucleic acid molecule comprising a nucleotide sequence encoding a porcine kappa light chain immunoglobulin locus or fragment thereof, wherein said nucleotide sequence comprises at least 17, 20, 25 or 30 contiguous nucleotides of a nucleotide sequence selected from the group consisting of: SEQ ID No 15, SEQ ID No 16, SEQ ID No 19, and SEQ ID No 25.
23. The nucleic acid molecule of claim 22 wherein said the nucleotide sequence is selected from the group consisting of: SEQ ID No 15, SEQ ID No 16, SEQ ID No 19, and SEQ ID No 25.
24. An isolated nucleic acid molecule comprising a nucleotide sequence encoding a porcine lambda light chain immunoglobulin locus or fragment thereof comprising at least 17, 20, 25 or 30 contiguous nucleotides of a nucleotide sequence selected from the group consisting of: SEQ ID No 32, SEQ ID No 33, SEQ ID No 34, SEQ ID No 35, SEQ ID No 36, SEQ ID No 37, SEQ ID No 38 and SEQ ID No 39.
25. The nucleic acid molecule of claim 24 wherein the nucleotide sequence is selected from the group consisting of: SEQ ID No 32, SEQ ID No 33, SEQ ID No 34, SEQ ID No 35, SEQ ID No 36, SEQ ID No 37, SEQ ID No 38 and SEQ ID No 39.
26. A cell transfected with the targeting vector of claim 1 and the targeting vector of claim 7.

FIGURE 1

Heavy-chain strategy

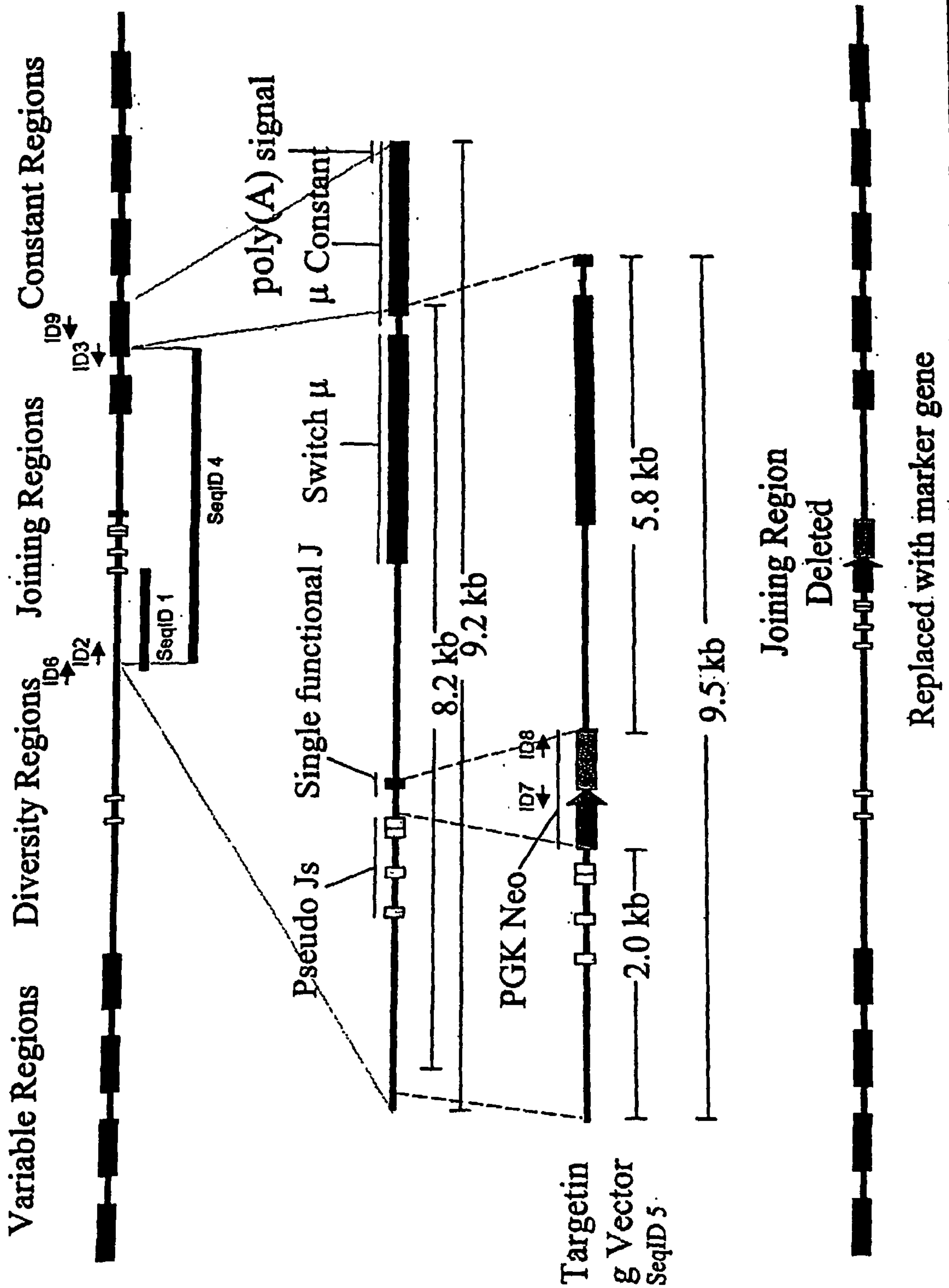


FIGURE 2

Light-chain Kappa strategy

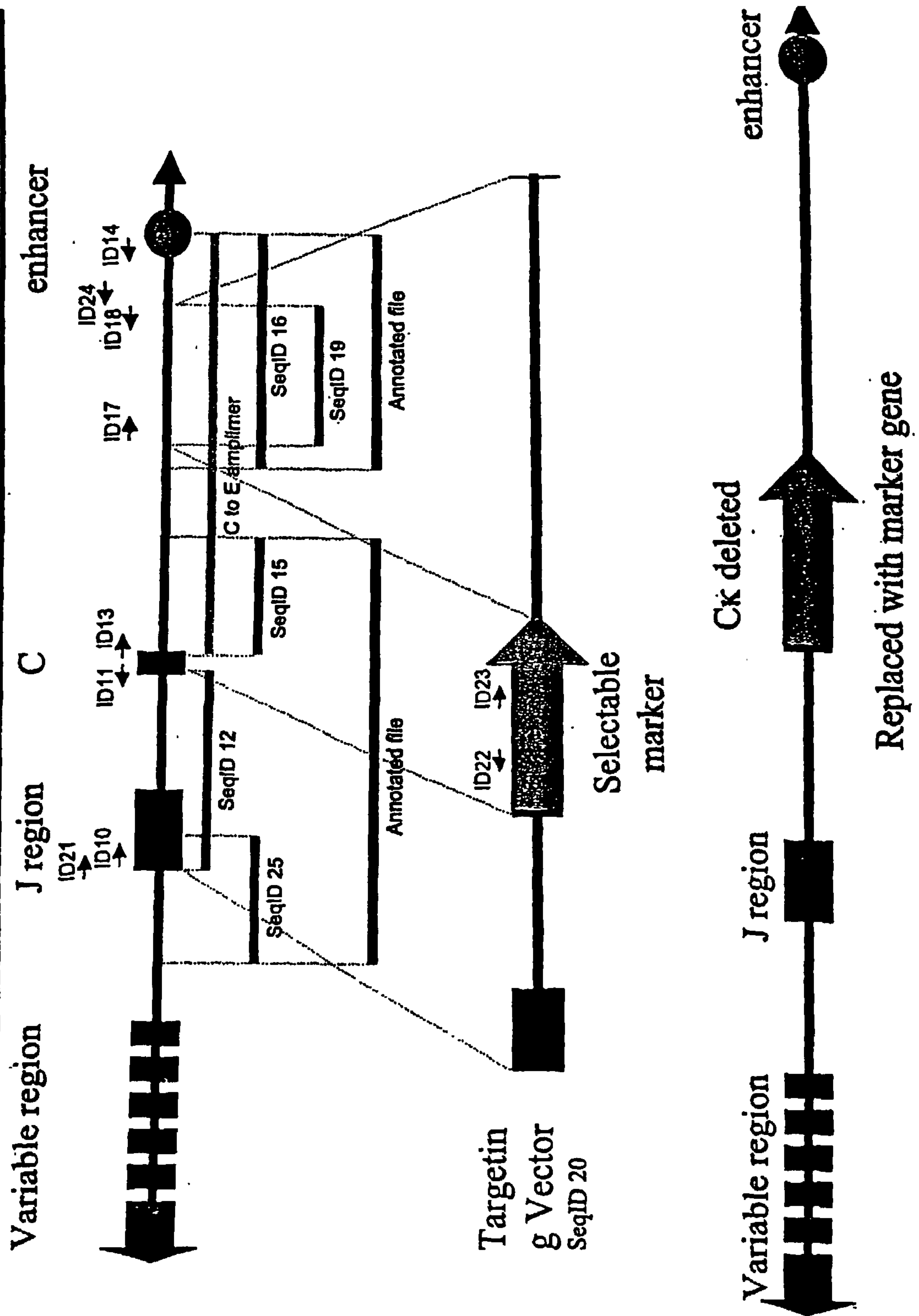
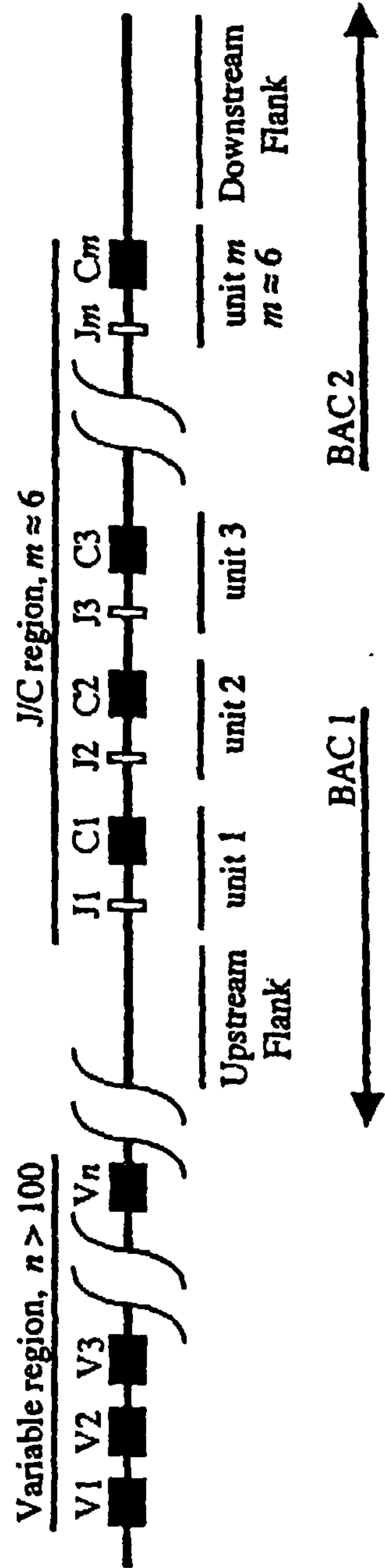
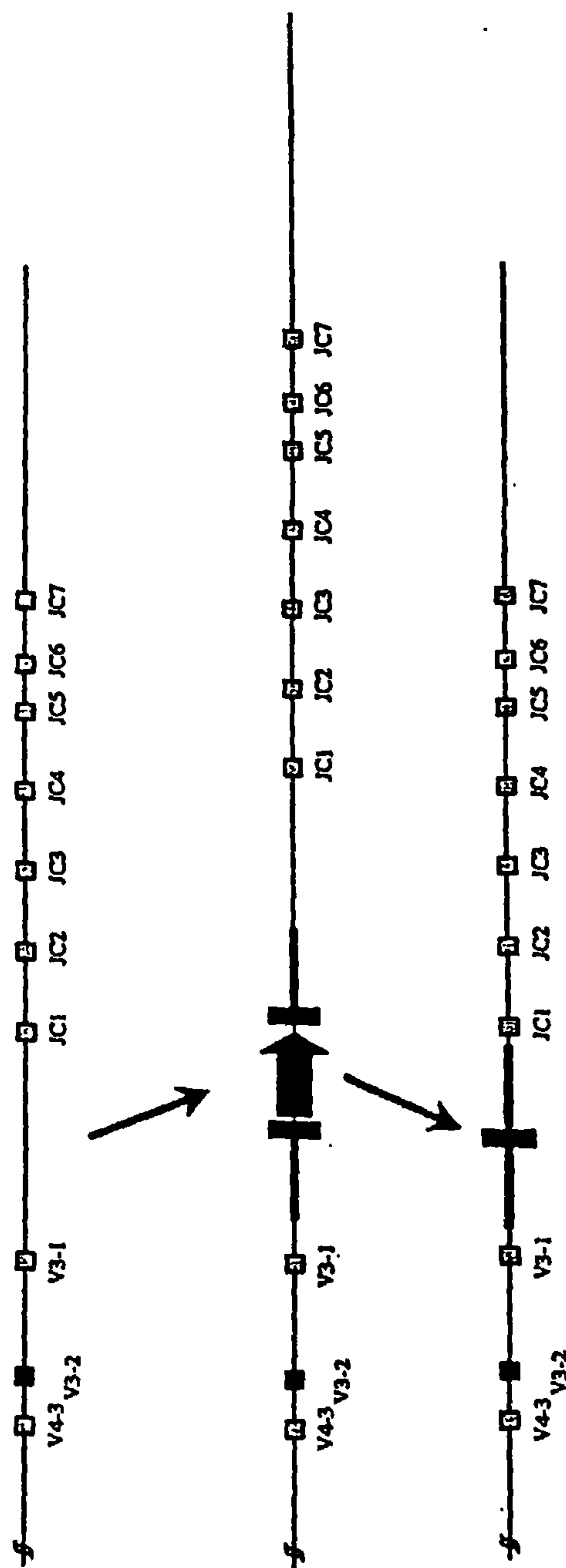
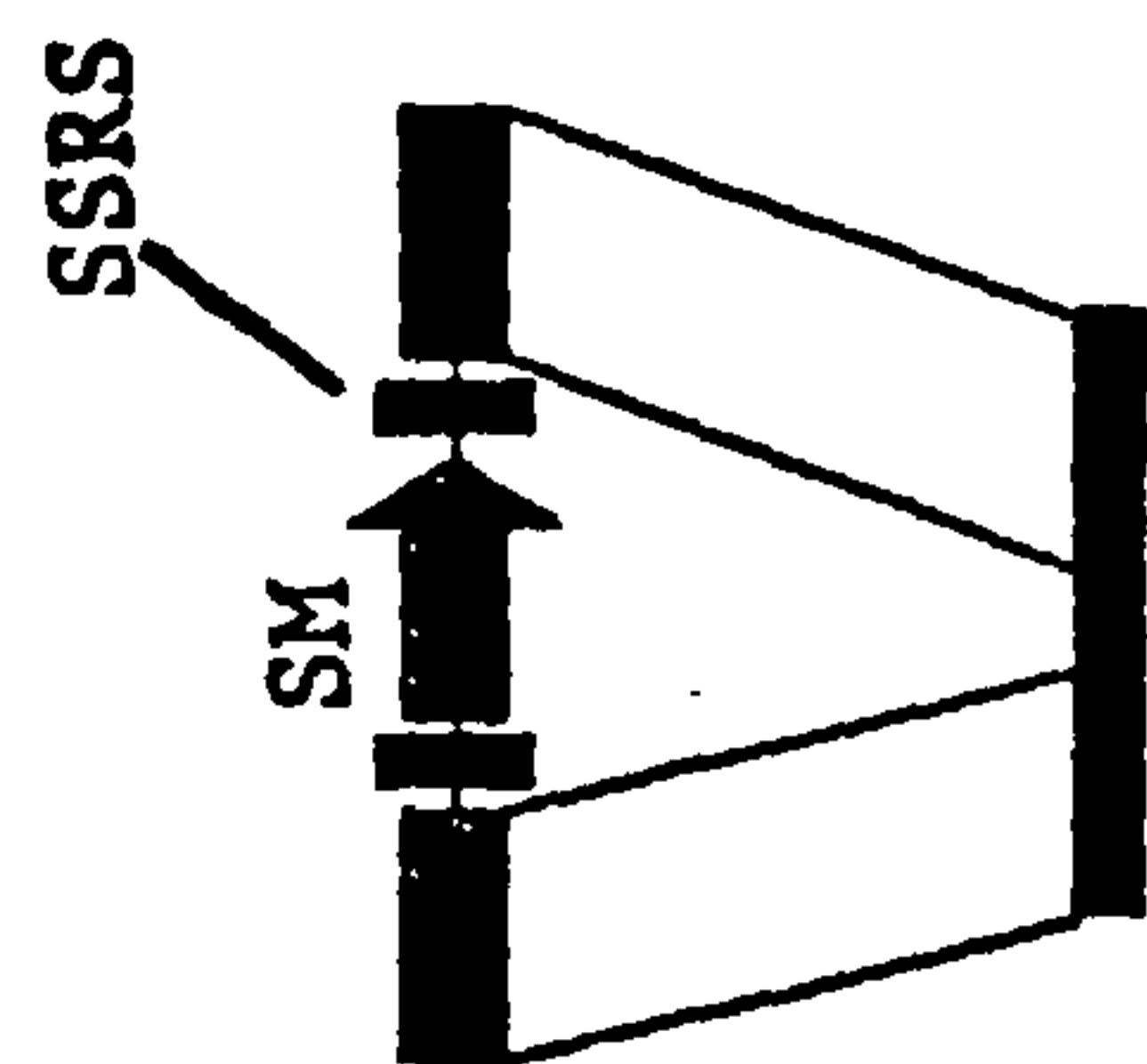


Figure 3



Graphic view of Human sequence AC002060 and NG_000002 with Porcine lambda contigs that flank the JC cluster.

Figure 5



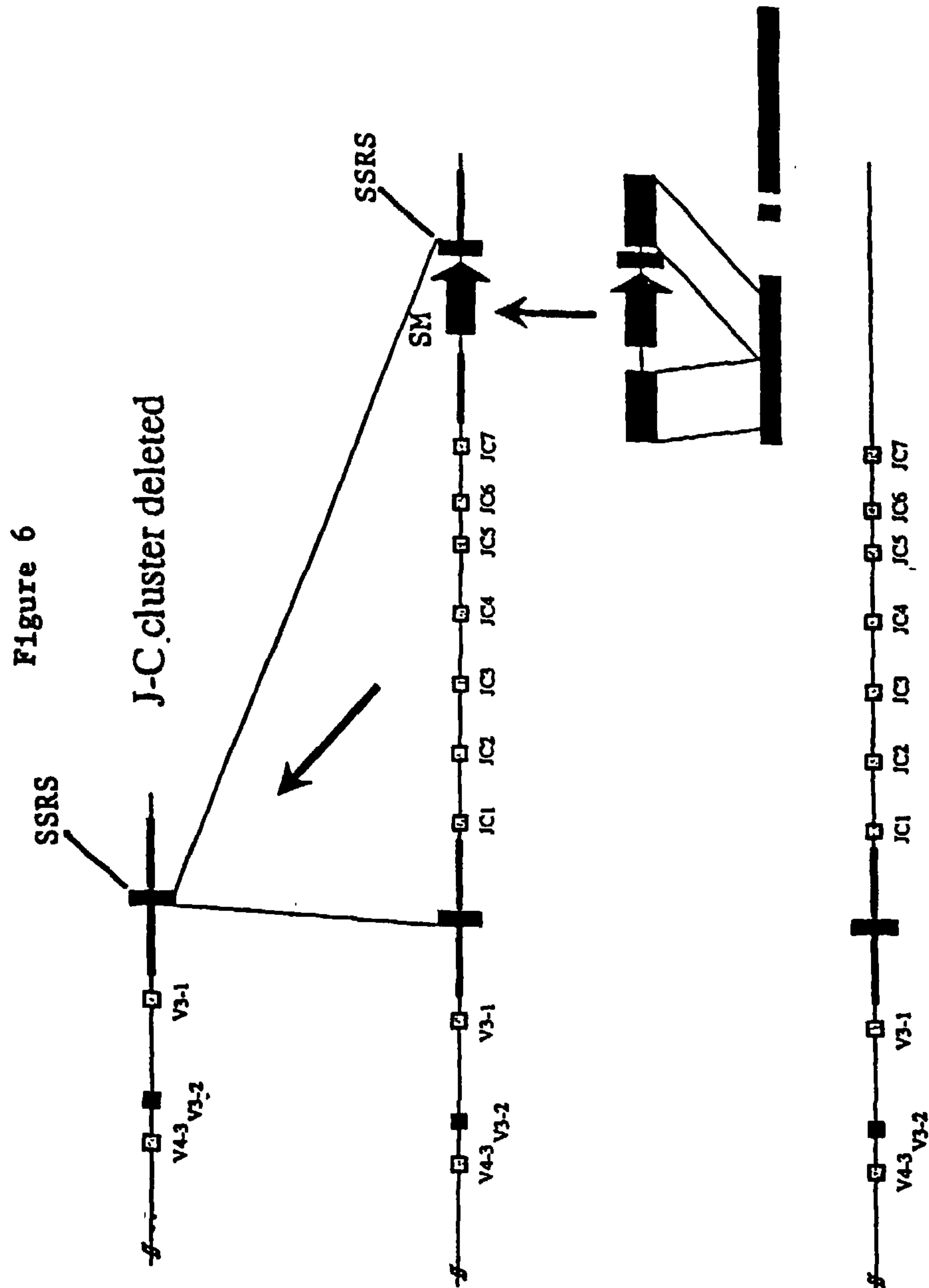
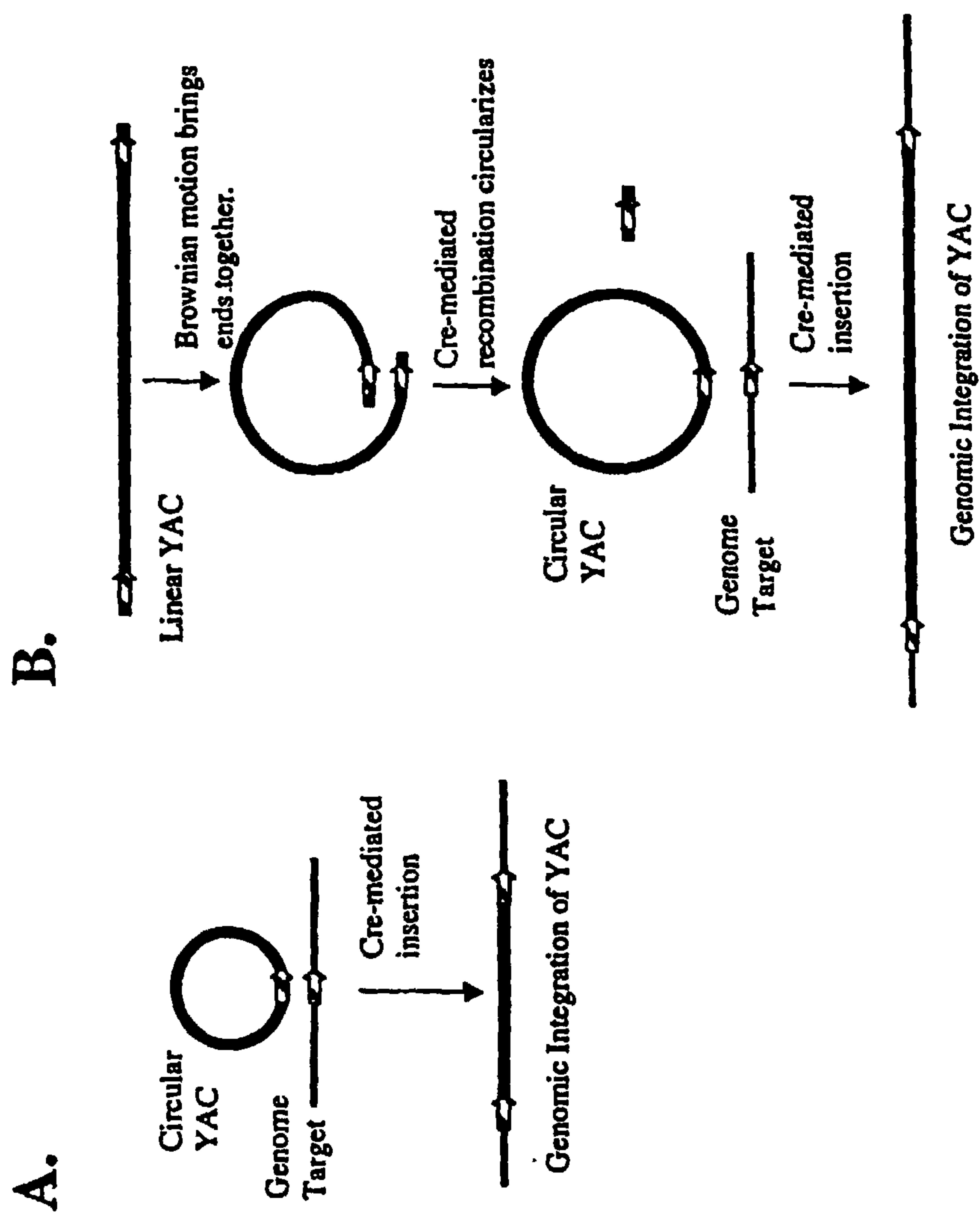


Figure 7



Heavy-chain strategy

