

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
1 April 2004 (01.04.2004)

PCT

(10) International Publication Number  
**WO 2004/026903 A2**

- (51) International Patent Classification<sup>7</sup>: **C07K 14/45** (74) Agents: MESTROM, J., J., L. et al.; Intervet International B.V., P.O. Box 31, NL-5830 AA Boxmeer (NL).
- (21) International Application Number: PCT/EP2003/010696 (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (22) International Filing Date: 19 September 2003 (19.09.2003) (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 02078953.3 20 September 2002 (20.09.2002) EP
- (71) Applicant (*for all designated States except US*): **AKZO NOBEL N.V.** [NL/NL]; Velperweg 76, NL-6824 BM Arnhem (NL).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **VAN POPPEL, Nicole, Francisca, Johanna** [NL/NL]; Van Welderenstraat 105 A, NL-6511 MG Nijmegen (NL). **VERMEULEN, Arnoldus, Nicolaas** [NL/NL]; Korhoenderveld 34, NL-5431 HH Cuyk (NL). **SCHAAP, Theodorus, Cornelis** [NL/NL]; Van de Does de Willeboissingel 53, NL-5211 CE's-Hertogenbosch (NL).
- 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: LIVE ANTENUATED PARASITE VACCINE

(57) Abstract: The present invention relates *inter alia* to attenuated live parasites of the phylum Apicomplexa and the family of Trypanosomatidae and to the use of such attenuated live parasites in a vaccine and in the manufacturing of such a vaccine. Furthermore, the present invention relates to vaccines comprising such attenuated live parasites and to methods for the production of such vaccines. Finally, the invention relates to specific tet-repressor fusion proteins and to attenuated live parasites according to the invention comprising such tet-repressor fusion proteins.



WO 2004/026903 A2

## LIVE ATTENUATED PARASITE VACCINE

The present invention relates to attenuated live parasites of the phylum Apicomplexa and the order of Kinetoplastida, to the use of such attenuated live parasites in a vaccine and in  
5 the manufacturing of such a vaccine, to vaccines comprising such attenuated live parasites, to methods for the production of such vaccines, to specific tet-repressor fusion proteins and to attenuated live parasites comprising such tet-repressor fusion proteins.

Within the regnum protozoa, the phylum of the Apicomplexa and the order of the  
10 Kinetoplastida, more specifically the family of Trypanosomatidae, are known to harbour several notoriously pathogenic parasites.

The family of Trypanosomatidae harbours i.a. parasites belonging to the genus *Leishmania* and *Trypanosoma*.

15 *Leishmaniosis* is a term for a variety of disease manifestations caused by *Leishmania*. The disease occurs most commonly in dogs and humans. The parasite is transmitted by sand flies to a mammalian host and is prevalent in all tropical and subtropical zones of the world. In the host parasites are taken up by macrophages where they stay and multiply, causing chronic inflammatory processes. Clinically, the disease in dogs is characterised  
20 by loss of weight, anaemia, pyrexia and lymphadenopathy. Cutaneous lesions are frequently observed. In humans multiple *Leishmania* species are infective, of which the most pathogenic is *L. infantum*, causing severe, visceral *Leishmaniosis* (known as Kala azar), which affects spleen, liver and bone marrow, and is fatal if left untreated. Other pathogenic *Leishmania* species are i.e. *L. major* and *L. mexicana*.

25 Multiple species of trypanosomes are known, causing a variety of different diseases in both man and animal. Two trypanosome species in particular, are known to be pathogenic: *Trypanosoma brucei* and *Trypanosoma cruzi*.

*T. brucei* species are present in African countries and cause sleeping sickness in humans  
30 and Nagana in animals (cattle, horses, pigs). *T. brucei* is transmitted by the TseTse fly, delivering the trypomastigote form into the host.

*T. cruzi* species are mainly present in South America, the parasite has a broad host range (including domestic and wild animals), but is famous for causing Chagas disease in man. The parasites are transmitted by cone-nosed bugs (like *Rhodnius* spp. and *Triatoma* spp.). The metacyclic trypomastigote stage infects the host and unlike *T. brucei*, will  
35

multiply inside the host cytoplasm of different cell types. After rupture of the host cell new trypomastigote forms are released which can again be ingested by cone nosed bugs.

The phylum Apicomplexa, harbours i.a. parasites of the family Eimeriidae. Many different  
5 Eimeria species are present in a large variety of mammals and birds. Seven prevalent species infecting the gastrointestinal tract of chickens are *Eimeria tenella*, *E. necatrix*, *E. brunetti*, *E. maxima*, *E. acervulina*, *E. praecox* and *E. mitis*. These Eimeria species are all involved in Coccidiosis in poultry. This makes Eimeria the cause of the most important parasitic disease in poultry, causing great economically losses for farmers. Eimeria infect  
10 epithelial cells and submucosa of the intestines, causing severe hemorrhagic enteritis, which leads to high mortality in young birds.

This disease has a worldwide distribution and is the most frequently recorded disease affecting poultry kept in modern poultry industry.

15 The family of Sarcocystidae, comprising Toxoplasma, Sarcocystis and Neospora is also known to have pathogenic members.

Toxoplasma is a widespread parasitic infection, being present in almost all mammals, in particular in goat, sheep and pigs, but also in humans. Prevalence in human populations can be as high as 70% of the total population. Infection often occurs via eating of  
20 undercooked meat contaminated with the parasite, but can also occur by ingestion of oocysts, being spread in the faeces of cats, which are the final host. When animals or humans are infected during pregnancy, it can cause spontaneous abortion or congenital toxoplasmosis in the developing foetus. This can result in, neurological sequels or ocular disorders. Chronic and lethal infections (encephalitis) can occur in immune compromised  
25 patients.

Neospora, in particular *N. caninum* is a coccidian parasite very similar to Toxoplasma. However, in contrast to Toxoplasma, Neospora has the dog as final host. *N. caninum* induces abortions in its intermediate host, and can cause severe abortion storms in cattle.  
30 Another Neospora species, *N. hughesi*, is suspected to cause equine protozoal myeloencephalitis in horses.

Many Sarcocystis species are present in cattle, pigs, sheep, goats and horses. Economically, *Sarcocystis neurona* is recognized as the most common cause of clinical  
35 equine protozoal myeloencephalitis in horses. In the U.S. 50% of horses are seropositive for *S. neurona*.

Plasmodium belongs to the Haemosporida and is known i.a. as the cause of malaria, being transmitted by mosquitoes. In humans four Plasmodium species have been described, of which *P. falciparum* is the most pathogenic and deathly. 400 million people  
5 are estimated to be infected, causing two million deaths each year. Initial clinical symptoms are rhythmic fevers. After initial infection, Plasmodium parasitizes the red blood cells, often resulting in anaemia. Parasitized red blood cells are sequestered in capillaries of internal organs, thereby causing tissue anoxia. This is particularly serious in the brain, giving rise to multiple petechial haemorrhages, leading to oedema and coma,  
10 which may be fatal. Although Plasmodium species have mainly been described in man, other Plasmodium species can infect a large variety of vertebrates.

Babesia and Theileria, both belonging to the Piroplasmida harbour parasite species affecting many mammalian species, and causing a variety of different diseases.  
15 Babesia species are transmitted by ticks and can infect a wide range of vertebrates causing a disease referred to as Babesiosis. The disease is characterised by listlessness, anaemia and parasitemia leading to multi-organ dysfunction in infected animals. In advanced stages haemoglobinuria occurs. Important Babesia species in cattle include *B. bovis*, *B. divergens*, *B. major* and *B. bigemina*. In dogs *B. canis*, *B. rossii*, *B. microti* and *B.*  
20 *gibsoni* species are mainly causing Babesiosis and are a common cause of death. Some Babesia species, like *B. divergens* and *B. microti* have been reported to infect humans as well.

Theileria is a tick-transmitted disease, infecting ruminants and is mainly a problem in cattle. Theileria infects and develops in leukocytes and erythrocytes. Pathology is mainly  
25 attributable to the intra-leukocyte stage. Two major Theileria species should be discriminated in cattle, *T. parva* and *T. annulata*. *T. parva* causes East Coast Fever, a deathly cattle disease, being endemic in various African countries. East Coast Fever is characterised by high fever, lymphadenopathy, severe pulmonary oedema and wasting. *T. annulata* infects cattle and buffalo, first invading cells of the lymphatic system and later  
30 appearing in the peripheral blood as intra-erythrocyte forms. Infection with *T. annulata* is usually referred to as tropical Theileriosis. The disease starts with high fever and swelling of lymph nodes, followed by listlessness, accelerated pulse and respiration rates and anorexia. In the final stage of disease anaemia is observed and ultimately death occurs. In the horse *Babesia equi* (which has been re-named as *Theileria equi*) is also a major  
35 pathogen.

It is clear that different ways of attack against these parasites have been studied through the years.

One of the routes of combating parasitic infections is the use of pharmaceutical components, such as the extensive use of anticoccidials that nowadays is a very common therapy in the treatment of poultry Coccidiosis. Another route is undoubtedly vaccination. It is clear that, especially where there is an increasing reluctance against the use of antibiotics, there is a need for new and effective vaccines, especially vaccines that provide broad protection.

Currently, two different approaches are used in vaccination against parasitic infections: vaccination with a live attenuated vaccine and vaccination with inactivated (killed) vaccines. Both approaches have their advantages and disadvantages, as summarized below:

The main advantage of attenuated vaccines is that they closely mimic the natural infection: they activate all phases of the immune system, they can induce humoral IgG and local IgA, they raise immune responses to many protective antigens, they provide a more durable immunity and are more cross-reactive. Moreover they are low-cost and they provide a quick immunity in the majority of cases.

Disadvantages of attenuated vaccines are the difficulties in finding the right level of attenuation and the possibility of reversion to virulence (these are major disadvantages), the spread to contacts of the vaccinee and the problems in immuno-compromised humans and animals.

Advantages of inactivated vaccines are that they provide sufficient humoral immunity if boosters are given, they show no mutation or reversion (a big advantage), they can be used with immuno-deficient patients, and in principle they are safe.

Disadvantages of inactivated vaccines: they often do not raise (cellular) immunity, boosters are needed, they provide no local immunity (important), they are more expensive and their use is dangerous if inactivation is below 100%.

Development of vaccines against parasites however is complex, if only because of the complexity of the parasites as such, when compared to other microorganisms. Next to this, the various parasites even within the phylum Apicomplexa and within the family of Trypanosomatidae, although related, do not have sufficient similarity in their genetic

make-up to allocate a common attenuation site or inactivation method, equally applicable to all these parasites. Moreover, for the manufacture of attenuated live vaccines it is necessary to locate suitable attenuation targets for each and every parasite. For the production of killed vaccines, one needs to know which antigens must be left unaltered by the inactivation method for each and every parasite. And apart from this, so far, many inactivated parasite vaccines have been shown not to be effective. Finally, there is a variety of different infection routes, different hosts, different host cells within the host and often even host changes during the life cycle which is a characteristic of most parasites and which again differs from parasite to parasite. This also complicates the development of vaccines.

Therefore, the development of vaccines for combating parasitic infection so far has been difficult, time consuming and not very successful.

It is an objective of the present invention to provide vaccines for combating infections caused by parasites of the phylum Apicomplexa and the family of Trypanosomatidae, that combine most of the advantages of both killed and live attenuated vaccines almost completely without having the disadvantages of these vaccines. Moreover, the method for the production of such vaccines is universally applicable to parasites of the phylum Apicomplexa and the family of Trypanosomatidae.

In the life cycle of all parasites of the phylum Apicomplexa and the family of Trypanosomatidae, there is at least one moment in which a certain stage infects a cell of a host and starts dividing. It was now surprisingly found that if ribosome synthesis can be stopped at or around the moment of infection, the parasite nevertheless does enter the host cell and divides several times using the present pool of ribosomes, thereby perfectly mimicking natural infection. Finally however, after several rounds of dividing, the progeny parasites will die due to lack of ribosomes.

This has the advantage that the induction of the immune response after infection is triggered in the most natural way, as if a virulent infection occurred, whereas contrary to the natural situation the parasite will after some time unavoidably become extinct. This goal was attained by placing a ribosomal protein gene under the control of an inducible promoter.

An inducible promoter is a promoter that can deliberately be switched on and off. Examples of such promoters will be given below.

In principle, each ribosomal protein gene can be used as a target, since in principle all ribosomal proteins are needed for the synthesis of a stable, fully functional ribosome. All parasites of the phylum Apicomplexa and the family of Trypanosomatidae have  
5 cytoplasmatic ribosomes, and most of them have plastid ribosomes and/or mitochondrial ribosomes. All of these are necessary for the normal development of the parasite. Therefore, live attenuated parasites according to the invention can be obtained by placing a ribosomal protein gene under the control of an inducible promoter, regardless the fact if this ribosomal protein gene encodes a ribosomal protein to be incorporated in plastid-,  
10 mitochondrial or cytoplasmatic ribosomes.

Ribosomal protein sequences are highly conserved between the various parasites. Therefore, DNA probes of the ribosomal sequences provided below can be used for the detection of the analogous ribosomal proteins in each of the parasites of the phylum  
15 Apicomplexa and the family of Trypanosomatidae. Additionally, the sequences of many ribosomal protein genes for many different parasites can be found in the NCBI-protein database (<http://www.ncbi.nlm.nih.gov>).

The fact that the lack of one ribosomal protein can already disturb the formation of stable  
20 ribosomes has been demonstrated in various plants, animals and microorganisms. Merely as an example: in *Drosophila*, mutations in some of the eighty ribosomal proteins have been shown to result in a typical phenotype, e.g. thin and short bristles, prolonged development, and female semi-sterility in heterozygotes as well as homozygous lethality. This phenotype, termed Minute phenotype, has been observed with mutations in for  
25 example the ribosomal proteins S13, and L9, (Schmidt, A., Hollmann, M., Schäfer, U., Mol. Gen Genet. 251:381-387 (1996), Sæbøe-Larssen & S., Lambertsson, A., Genetics 143: 877-885 (1996)). Another example is the ribosomal protein gene YS3 of yeast, which encodes the yeast ribosomal protein S3. Its disruption yields non-viable haploid spores of *Saccharomyces cerevisiae* (Finken-Eigen, M., Domdey, H., Köhrer, K., Biochemical and  
30 Biophysical research communications 223, 397-403 (1996)). These studies demonstrated that down-regulating a single ribosomal protein can already affect the formation and/or proper functioning of ribosomal complexes.

The promoters to be used in parasites according to the invention for the control of  
35 transcription of the ribosomal protein gene need to fulfil only one prerequisite. They must be switched on during the propagation of the parasites. This is of course necessary in

order to provide the parasite according to the invention with the native amount of ribosomes necessary for normal propagation. The promoter must however be switched off in the recipient host that receives the parasite as a vaccine. A promoter is considered to be switched on if it initiates the transcription of the gene it controls. In the present invention this gene would be a ribosomal protein gene. A promoter is switched off if transcription of the gene that it controls is at least two times lower than in the on situation. Preferably, the level of transcription is at least 3, more preferably 4, still more preferably 5, 6 or even 7 times lower. It is stressed, that there is no need for a complete inhibition of transcription anyway. A low level of ribosomal protein transcription will finally result in an extended live span of the parasites, before they become extinct. Thus they will trigger the immune system for a somewhat longer period.

In principle, there are two different possibilities: either the promoter is switched on unless some condition is applied that switches the promoter off, or the promoter is switched off unless some condition is applied that switches the promoter on.

Preferably, the promoter is in the switched off status unless some condition is applied that is not present in the recipient host that receives the parasite as a vaccine.

If necessary, two or more ribosomal protein genes can be placed under the control of inducible promoters. This would be a preferred option if the inducible promoter used in a promoter cannot be sufficiently switched off, i.e. if the inducible promoter is a leaky promoter, or in the exceptional case that lack of one specific ribosomal protein is not sufficient to destabilize the ribosome.

The invention will be explained by the following examples.

*Toxoplasma gondii* uses the cat as a final host, and uses herbi- and omnivores respectively carnivores as subsequent intermediate hosts. In the case of *Toxoplasma*, it is the oocyst/tissue cyst stage of the parasite that ultimately infects humans. Humans and warm-blooded animals are the target mammals for vaccination, and therefore the *Toxoplasma* tachyzoite is the parasitic stage for which the live attenuated parasite is needed. Therefore, the tachyzoite is the parasitic stage in which, according to the invention, a ribosomal protein gene is brought under the control of an inducible promoter. The thus made recombinant parasite, further also referred to as the attenuated live parasite, can be propagated in the classical way under conditions under which the promoter is switched on. Under these circumstances, the number of ribosomes will be identical or close to that in the native situation. If sufficient parasites are grown for vaccine purposes, the live attenuated parasites are collected and administered as a vaccine. In the host to be vaccinated, the conditions under which the promoter is switched on are not



present and as a result the promoter will remain in the switched off situation. At the moment of vaccination, the parasite will behave as a wild-type parasite, because the pool of ribosomes is comparable to the native situation. Therefore, the process of infection, and of invasion of the host cell will perfectly mimic the process of natural infection. As soon as the parasite starts dividing in the host, it also divides the pool of ribosomes over its progeny. Since the promoter of (at least) one of the ribosomal protein genes is however in the switched off position when in the host cell, there will be either reduced or even no *de novo* synthesis of ribosomes. Therefore, the progeny will slowly become extinct. Nevertheless, in the meantime the process of infection, and therefore the triggering of the immune system has continued as in the case of a wild-type parasitic infection. Therefore, ultimately immunity will have build up as if an infection with a virulent wild-type parasite had taken place, whereas the live attenuated parasites used for the induction of immunity have become extinct after one or a few rounds of infection. The examples below provide further details.

The life cycle of *Neospora caninum* is comparable with that of *Toxoplasma* except for the fact that *Neospora* uses dogs as the final host, and causes abortions in i.a. cattle, dogs, sheep and horses. The approach for *Neospora* vaccines thus closely relates to that of *Toxoplasma* as described above. As for *Toxoplasma*, the tachyzoite is the parasitic stage in which, according to the invention, a ribosomal protein gene is brought under the control of an inducible promoter. The development of molecular genetics tools for *Neospora* has been described i.a. by Howe, D.K. and Sibley, L.D. METHODS: 13(2): 123-33 (1997))

For the production of a live attenuated *Eimeria* parasite, the merozoite is the parasitic stage in which, according to the invention, a ribosomal protein gene is brought under the control of an inducible promoter. In this case, the vaccine does not comprise the merozoite however, but the sporulated oocysts. This is due to the fact that the sporulated oocyst is the form in which the parasite is normally ingested by the chicken. For the replication of the first recombinant merozoites made according to the invention, it suffices however to introduce these into the digestive tract of the chicken. Recombinant oocysts will then be shed by the chicken and can be isolated and directly used as the live attenuated parasite in Coccidiosis vaccines, e.g. oral vaccines for administration to drinking water. Isolation of oocysts from chicken dung is a standard procedure well known in the art. Genetic engineering of *Eimeria* has i.a. been described by Kelleher, M. and Tomley, F.M. (Mol. Biochem. Parasitol. 97(1-2): 21-31 (1998)).

A live attenuated Malaria vaccine according to the invention can be made e.g. starting from erythrocyte stage plasmodium parasites. Plasmodium recombinant sporozoites. The sporozoite is the phase of the parasite that is injected into the (human) blood stream by the female mosquito. The sporozoite infects the liver within two minutes after injection, to produce schizonts and merozoites. The merozoites, in turn, infect erythrocytes and replicate there. It is at this moment in time that the pool of ribosomes must be divided over a large number of progeny parasites, and this is the moment at which the progeny parasites will become extinct. The whole immune defence system is already fully triggered at that moment in time. This example again illustrates the advantage of vaccines based upon recombinant parasites according to the present invention: they share all the advantages of live vaccines with the advantages of inactivated vaccines. Vaccination will preferably be done with either recombinant erythrocyte stage plasmodium parasites or (less practically) recombinant sporozoites. Recombinant DNA techniques for Plasmodium have been described i.a. by Crabb, B.S., et al., (Mol. Biochem. Parasitol. 90: 131-144 (1997)) and by Wu, Y. et al., (Proc. Natl. Acad. Sci., 93: 1130-1134 (1996), and Proc. Natl. Acad. Sci., 92: 973-977 (1995))

Live attenuated Theileria vaccines according to the invention can again be based upon recombinant merozoites. These merozoites can be grown and maintained in lymphocytes. It is in the lymphocyte that the merozoite starts dividing, synchronously with the division of the lymphocyte, while a few free progeny parasites will infect other lymphocytes, altogether leading to the induction of wild type like immunity, however leading, as in the other examples, to progeny that finally becomes extinct due to slowly increasing lack of ribosomes. Theileria can be propagated and cultured in primary lymphoid cells. See e.g. Shkap V. et al., Vet. Parasitol. 65: 11-20 (1996) and Hulliger, L. J. Protozool. 12: 649-655 (1965).

Live attenuated Babesia vaccines can be made using the merozoites and/or trophozoites for recombination. These can be cultured in erythrocytes. The whole approach is comparable to that described for Theileria above. See i.a. Levy, M.G and Ristic, M. Science 207: 1218-1220 (1980).

For Sarcocystis species such as *S. suihominis* and *S. neurona*, both the sporozoite and the merozoite are targets for recombination according to the invention. And again, the principle is the same: the recombinant sporozoite provides recombinant merozoites and these merozoites slowly become extinct due to lack of ribosomes in the absence of *de*

*novo* ribosome protein synthesis. The recombinant merozoites can be used directly in a vaccine. See e.g. Murphy, A.J. and Mansfield, L.S. J. Parasitol. 85: 979-981 (1999) and Ellison, S.P. et al., Vet. Parasitol. 95: 251-261 (2001).

- 5 As far as the order of Kinetoplastida is concerned, tetracycline regulated gene expression has been described for *Trypanosoma brucei* (Wirtz, E. and Clayton, C., Science 268: 1179-1183 (1995) and Biebinger, S. et al., Mol. & Biochem. Parasitol. 85: 99-112 (1997)); *Trypanosoma congolense* (Inoue N., et al., Mol. & Biochem. Parasitol. 120: 309-313 (2002)) and *Leishmania donovani* (Yan, S., et al., Mol. & Biochem. Parasitol. 112: 61-69 (2001)), and can be adjusted to regulate ribosomal protein gene transcription as follows: briefly, the procyclic form of the parasite is the target for transfections. The tetracycline repressor is integrated into the non-transcribed spacers of the ribosomal RNA repeats, so that heterologous genes (in this reference not a ribosomal gene) can be regulated in a tetracycline dependent manner. For the construction of live attenuated parasites
- 10 according to the invention of the order of Kinetoplastida, first an extra copy of a ribosomal protein gene is inserted together with a promoter containing one or more tetracycline operator elements. Subsequently, the endogenous gene copy is deleted from the parasite genome. This can easily be done by homologous recombination preferably in the presence of a marker for recombination. This is comparable to methods for Apicomplexa as described below. Direct targeting of the endogenous ribosomal protein genes is not feasible for *Leishmania* and *Trypanosoma*, because most genes in *Leishmania* and *Trypanosomes* are organized into large (> 100-500 kb) polycistronic clusters of adjacent genes on the same DNA strand. Thus inhibition of one gene would lead to inhibition of the transcription of all the genes localised downstream (Myler, P.J. et al., Med. Microbiol.
- 20 Immunol. 190: 9-12 (2001)).

- The examples given above are indeed merely examples. They by no means limit the scope of the invention. Examples of all kinds of parasites of the phylum Apicomplexa and the family of Trypanosomatidae and their life cycles can be found in the Encyclopaedic
- 30 Reference of Parasitology, Heinz Mehlhorn, Springer Verlag (2001) (ISBN 3-540-66829-2). Man skilled in the art is thus perfectly able, with the examples given above and using the Encyclopaedic Reference of Parasitology, to determine which stage would be the preferred stage as a starting point for making the live attenuated parasite according to the invention, for each parasite of the phylum Apicomplexa and the family of
- 35 Trypanosomatidae.

Many of the parasites belonging to the families mentioned above have a variety of different hosts. Merely as an example: there are Babesia species such as *B. canis* infecting dogs, *B. caballi* infecting horses, mules and donkeys, *B. divergens* infecting cattle, wild ruminants and humans. Nevertheless, in all cases the parasitic life cycle is comparable. Therefore, where it is indicated above that a vaccine according to the invention against e.g. Babesia can be based upon recombinant merozoites, this holds true for all Babesia species. Details concerning the life cycles of the various species of one family can also be found in the Encyclopaedic Reference of Parasitology, Heinz Mehlhorn, Springer Verlag (2001) (ISBN 3-540-66829-2), mentioned above.

10

Thus, one embodiment of the present invention relates to attenuated live parasites of the phylum Apicomplexa or the family of Trypanosomatidae that have as a characteristic that they comprise a ribosomal protein gene under the control of an inducible promoter.

- 15 The concept of inducible promoters has already been mentioned shortly above. An inducible promoter is a promoter that can be switched on and off under the influence of an external factor. Such a switching factor can be a physiological factor e.g. heat; the trigger of all of the many heat-shock promoters well known in the art for decades already. Such a factor can also be of chemical nature. Many such factors are again well known in the art.
- 20 There are too many inducible promoters known in the art to mention them all. A few examples will be mentioned here. The IPTG-inducible Lac-promoter is possibly one of the most frequently used inducible promoters. Alternative inducible promoter systems are e.g. the tetracycline-controlled transactivation system (Baron, U. et al., Oxford University Press 25: 2723-2729 (1995)) and the ecdysone-inducible expression system (Invitrogen) (Yao, T.P. et al., Cell 71: 63-72 (1992)).
- 25

In principle there are two kinds of inducible promoters: those that are switched on in the presence of a condition, and those that are switched off in the presence of a condition. This condition may be the presence of a chemical substance.

- 30 In a preferred form of this embodiment of the invention, the promoter to be used is switched on in the presence of a condition that is not naturally present in the host. The use of such promoters has the advantage that they automatically are in (or will switch to) the switched off position as soon as they are administered to the natural host of the parasite. This implies that a live attenuated parasite according to the invention is preferably grown under "artificial" conditions, i.e. conditions not present in the natural host,
- 35 in order to replicate.

A preferred type of inducible promoters is the type of inducible promoters based upon an operator site and a repressor capable of reversibly binding said operator site. The binding and detachment of the repressor protein can then be regulated by the "condition" applied as mentioned above, i.e. the presence or absence of heat, chemicals, etcetera.

A very suitable example of an inducible promoter, or more precisely; a promoter/operator/repressor complex, that can very efficiently be used in attenuated live parasites according to the invention, is the tet-promoter/tet-operator complex, further also referred to as the tetR-system.

The tetR-system as such has been described and proven to work in many different protozoan parasites, such as *T. brucei* (Wirtz et al., Science 268:1179-1183 (1995), Biebinger et al. (Mol. Biochem. Paras.85: 99-112 (1997)) and in *Entamoeba histolytica* (Hamann et al., Mol. Biochem. Paras. 84: 83-91 (1997)). The tetR-system was also successfully used in Toxoplasma to regulate expression of myosin A (Meissner M, et al., Nucleic Acids Res. 29(22): E115 (2001)). In addition, tetracycline regulated expression was also demonstrated in *Giardia lamblia* and *Leishmania donovani*, showing its universal applicability in parasites (Yan S, et al., Mol Biochem Parasitol. 112(1): 61-9 (2001), Sun, C.H. and Tai, Mol. Biochem. Parasitol. 105(1): 51-60 (2000)).

This complex operates as will be described shortly below and more extensively in the examples.

In principle, two steps must be made in order to generate tetracycline-regulated expression of ribosomal proteins: 1. integration and expression of a tetracycline repressor (tetR) gene and 2. integration of one or more tetracycline operator element(s) in the promoter of a ribosomal protein gene near the start of transcription.

The tet-repressor gene is a gene that encodes a protein capable of binding to the tet-operator site thus blocking transcription of the adjacent gene. This gene is now placed under the control of a constitutive promoter (i.e. constitutive in the recombinant parasite) and brought into the parasite using recombinant DNA techniques. Thus, the recombinant parasite will synthesize the tet-repressor protein. The tet-operator is preferably introduced in the vicinity of the transcription start site of one or more ribosomal protein genes, preferably in the endogenous promoter, upstream of the STS. The tet-repressor protein will consequently bind to the tet-operator, thus blocking the transcription of the downstream ribosomal protein gene. In the presence however of tetracycline, the

repressor will detach from the tet-operator site, thus enabling the transcription of the downstream gene. Therefore, in the presence of tetracycline, the recombinant parasite will be able to replicate as in the natural situation. If the recombinant parasite can be grown in vitro, as is the case for many parasites including most of the parasites of the examples  
5 given above, tetracycline can easily be added to the growth medium. If the growth of the parasites requires propagation in the natural host, which is e.g. the case for Eimeria parasites, tetracycline can easily be administered orally or by injection to the host (in this case the chicken). The following should be noted: tetracycline is taken up by extracellular and intracellular parasites. Cell rupture of the host cell is not required for the drug to have  
10 effects on the regulation of the expression of ribosomal proteins.

Step 1, the integration and expression of the tetracycline repressor gene (tetR), can be obtained as described in the literature mentioned above. A suitable and well-known selection marker that indicates the stable transformation and possibly integration of the tetracycline repressor gene is e.g. the CAT-gene (Kim, K., et al., Science 262(5135): 911-  
15 4 (1993)). Other suitable markers for selection of stable transfection are also known in the art, such as DHFR-TS (Donald, R.G. and Roos, D.S., Proc. Natl. Acad. Sci. U S A 90 (24): 11703-11707 (1993), Roos, D.S. et al., METHODS 13: 112-122 (1997)) and HXGPRT (Donald, R.G. et al., J. Biol. Chem. 271: 14010-14019 (1996), Donald, R.G. and Roos, D.S., Mol. Biochem. Parasitol. 91 (2): 295-305 (1998)).  
20 The Cre-lox system also provides a suitable selection system (see i.a. Hardy, S. et al., Journ. Virol. 71: 1842-1849 (1997)).

If the tetR-system is used as an inducible promoter system, the promoter upstream of the ribosomal protein gene can e.g. be the endogenous promoter, now made inducible by  
25 cloning the tet-operator in the vicinity of the start site of transcription (see below for details of the tet-operator sequence and preferred insertion sites). It goes without saying that any other promoter capable of providing a sufficiently high transcription level of the ribosomal protein gene is also suitable.

30 If another inducible promoter system is used, it would be easy to use that inducible promoter and delete the endogenous promoter. If however another regulatory element is used, of which the principle is comparable to the tet-operator, the promoter itself can equally well be the endogenous promoter. Again it goes without saying that any other promoter capable of providing a sufficiently high transcription level of the ribosomal  
35 protein gene cloned downstream, is also suitable.

Step 2, the replacement of a wild-type ribosomal protein gene with one containing one or more tetO sites (= tet-operator sites) in the vicinity of the STS requires the insertion of the tet-operator site between the promoter of the ribosomal protein gene of choice and the gene itself. The tet-operator has been described by Yan S, et al. (Mol. Biochem. Parasitol. 112(1): 61-9 (2001)), by Wirtz, E and Clayton, C. (Science 268(5214): 1179-83 (1995)) and by Meissner M, et al. (Nucleic Acids Res. 29(22): E115 (2001)).

The sequence of a single tet operator (tetO) site is

5'- TCCCTATCAGTGATAGAGATC -3'.

In principle, insertion of a single tet-operator site in front of the ribosomal protein gene of choice would suffice. The tetR-system is, as all biological systems, however not inducible from exactly 0% to 100% activity and vice versa. Therefore, if a stronger level of regulation is needed, preferably two or more operator sites are inserted.

The tet-operator interferes with the binding of the RNA-polymerase that transcribes the downstream gene. Therefore, the tet-operator is preferably inserted somewhere in the region that extends from nucleotide -100 to +3 relative to the site at which the transcription starts (herein referred to as the STS). Moreover, in the examples it is additionally described how to locate such STS.

The step of replacement of a wild-type ribosomal protein gene with a recombinant gene comprising one or more tet-operator sites can i.a. be performed with the hit-and-run strategy described by Donald, R.G. and Roos, D.S. (Mol. Biochem. Parasitol. 91(2): 295-305 (1998)).

The skilled artisan will be able to find alternative methods using other combinations of positive and negative selection markers. HSV Thymidine kinase can for example be used as a negative selection marker. (LeBowitz, J.H. et al., Mol. Biochem. Parasitol. 51(2): 321-5 (1992), Fox, B.A. et al, Mol. Biochem. Parasitol. 116(1): 85-8 (2001)).

The molecular tools used for the construction of recombinant Toxoplasma parasites according to the invention work equally well in Neospora (Howe, D.K. and Sibley, L.D. METHODS 13(2): 123-133 (1997)).

In Eimeria, the same methods are equally applicable. Merely as an example: it was shown that beta-galactosidase could be transiently expressed in *E. tenella* by Kelleher, M. and Tomley, F.M. (Mol Biochem Parasitol. 97(1-2): 21-31 (1998)).

For *Theileria*, methods have e.g. been developed to transiently transfect infective, uninucleate *Theileria annulata* sporozoites by Adamson, R. et al. (Mol. Biochem. Parasitol. 114(1): 53-61 (2001)).

5 In *Plasmodium*, dihydrofolate reductase-thymidylate synthase (dhfr-ts) coding sequences mutated to confer resistance to pyrimethamine, or Puromycin-N-acetyltransferase, or the blasticidin S deaminase (BSD) gene of *Aspergillus*, or the neomycin phosphotransferase II (NEO) gene from transposon Tn5 have been described as selectable markers (Wu, Y., et al., Proc. Natl. Acad. Sci. U S A. 93(3): 1130-4 (1996), Wang, P., et al., Mol. Biochem. Parasitol. 123(1): 1-10 (2002), de Koning-Ward, T.F., et al. (Mol. Biochem. Parasitol. 117  
10 (2):155-60. (2001))

Similar selection markers work in *Babesia* as well.

Therefore, man skilled in the art will be able to apply the present invention over the full range of parasites belonging to the phylum Apicomplexa and the family of Trypanosomatidae.

15

A preferred form of this embodiment relates to an attenuated live parasite according to the invention that belongs to the Coccidia, the Piroplasmida or the Haemosporida.

20

In a more preferred form of this embodiment, the attenuated live parasite belongs to the family Eimeriidae, Cryptosporidiidae or Sarcocystidae.

In an even more preferred form of this embodiment, the attenuated live parasite belongs to the genus *Eimeria*, *Cryptosporidium*, *Toxoplasma*, *Sarcocystis* or *Neospora*.

25

In another more preferred form of this embodiment, the attenuated live parasite belongs to the family of the Babesiidae or the Theileriidae.

In an even more preferred form of this embodiment, the attenuated live parasite belongs to the genus *Babesia* or *Theileria*.

30

In another more preferred form of this embodiment, the attenuated live parasite belongs to the genus *Plasmodium*.

35

In still another more preferred form of this embodiment, the attenuated live parasite belongs to the genus *Trypanosoma* or the genus *Leishmania*.



In an even more preferred form, the attenuated parasite belongs to the species *Leishmania mexicana*, *L. infantum* or *L. major* or the species *Trypanosoma brucei* or *T. cruzi*

- 5 In another preferred form of this embodiment, a ribosomal protein gene of the live attenuated parasite according to the invention is under the control of an inducible promoter that is inducible by antibiotics.

10 More preferably, these antibiotics are tetracycline or anhydrotetracyclin, or derivatives thereof.

In another preferred form of this embodiment, the ribosomal protein gene of choice is the gene encoding L9, S3, plastid-S9 or S13, preferably the L9, S3, plastid-S9 or S13 of *Toxoplasma gondii*.

15

The nucleotide sequence of the gene encoding Large subunit ribosomal protein number 9 (L9), as well as upstream sequences comprising the promoter region is depicted in SEQ ID NO: 1

	REGION	1	2296	promoter	promoter region
20	REGION	2297	2461	e	exon 1
	REGION	2416	2418	atg	atg start codon
	GENE	2416	4831	cds	coding sequence
	REGION	2462	3838	i	intron 1
	REGION	3839	4260	e	exon 2
25	REGION	4261	4727	i	intron 2
	REGION	4728	4834	e	exon 3
	REGION	4829	4831	stop	TAA stopcodon

30 The nucleotide sequence of the gene encoding plastid Small subunit ribosomal protein number 9 (S9), as well as upstream sequences comprising the promoter region is depicted in SEQ ID NO: 2

	REGION	1	3076	promoter	promoter region
	REGION	3077	3616	e	exon 1
	REGION	3156	3158	atg	ATG start codon
35	GENE	3156	4325	cds	coding sequence
	REGION	3617	3874	i	intron 1

	REGION	3875	4034	e	exon 2
	REGION	4035	4130	i	intron 2
	REGION	4131	4338	e	exon 3
	REGION	4323	4325	stop	TAG stop codon
5	REGION	4326	4338	3' utr	3' UTR

The nucleotide sequence of the gene encoding Small subunit ribosomal protein number 13 (S13), as well as upstream sequences comprising the promoter region is depicted in SEQ ID NO: 3

10	REGION	1	1289	promoter	promoter region
	REGION	1290	1594	e	exon 1
	REGION	1448	1450	atg	ATG start codon
	GENE	1448	3639	cds	coding sequence
	REGION	1595	2527	i	intron 1
15	REGION	2528	2615	e	exon 2
	REGION	2616	3489	i	intron 2
	REGION	3490	3639	e	exon 3

The nucleotide sequence of the gene encoding Small subunit ribosomal protein number 3 (S3), as well as upstream sequences comprising the promoter region is depicted in SEQ ID NO: 4

	REGION	1	1177	promoter	promoter region
	REGION	1178	1308	e	exon 1
	REGION	1291	1293	atg	ATG start codon
25	GENE	1291	2651	cds	coding sequence
	REGION	1309	1752	i	intron 1
	REGION	1753	2137	e	exon 2
	REGION	2138	2249	i	intron 2
	REGION	2250	2389	e	exon 3
30	REGION	2390	2486	i	intron 3
	GENE	2487	2748	e	exon 4
	REGION	2649	2651	stop	TAA stop codon
	REGION	2652	2748	3' utr	3' UTR

Attenuated live parasites according to the invention are very suitable for use in vaccines. This is, as extensively explained above, due to the fact that they combine the advantages of both live attenuated and inactivated vaccines, without suffering from the disadvantages. Therefore, another embodiment of the present invention relates to attenuated live  
5 parasites according to the invention for use in a vaccine.

Still another embodiment of the invention relates to vaccines for combating parasitic infection that comprise a live attenuated parasite according to the invention and a pharmaceutically acceptable carrier.

10

A pharmaceutically acceptable carrier can be e.g. sterile water or a sterile physiological salt solution. In a more complex form the carrier can e.g. be a buffer such as PBS, well-known in the art.

15 Vaccines according to the present invention may in a preferred presentation also contain an immunostimulatory substance, a so-called adjuvant. Adjuvants in general comprise substances that boost the immune response of the host in a non-specific manner. A number of different adjuvants are known in the art. Examples of adjuvants frequently used in cow vaccines are muramyl dipeptides, lipopolysaccharides, several glucans and glycans  
20 and Carbopol® (a homopolymer).

The vaccine may also comprise a so-called "vehicle". A vehicle is a compound to which the protein adheres, without being covalently bound to it. Such vehicles are i.a. lipid vesicles, ISCOMs®, dendromers, niosomes, microparticles, especially chitosan-based microparticles, polysaccharide matrices and the like, bio-microcapsules, micro-alginates,  
25 liposomes and macrosols, all known in the art. Microparticles, more specifically those based upon chitosan, especially for use in oral vaccination are very suitable as vaccine vehicles.

A special form of such a vehicle, in which the antigen is partially embedded in the vehicle, is the so-called ISCOM® (EP 109.942, EP 180.564, EP 242.380)

30 In addition, the vaccine may comprise one or more suitable surface-active compounds or emulsifiers, e.g. Span® or Tween®. Also, the vaccine may comprise one or more immune stimulants such as cytokines, e.g. interferons.

Vaccines based upon live attenuated recombinant parasites described above can be  
35 administered in relatively low amounts, when compared to inactivated parasites, because they multiply themselves during the infection. Therefore, very suitable amounts would

range between  $10^2$  and  $10^7$  parasites per dose. Amounts below  $10^2$  parasites per dose may not always guarantee a sufficient level of protection in all vaccinated animals. Ranges from  $10^7$  up to  $10^8$  parasites per dose, although suitable, are not very practical, if only from an economic point of view.

5

Still another embodiment of the present invention relates to the use of an attenuated live parasite according to the invention for the manufacture of a vaccine for combating infection caused by a parasite of the phylum Apicomplexa or the family of Trypanosomatidae.

10

Again another embodiment of the present invention relates to methods for the production of a vaccine according to the invention that comprise the mixing of a live attenuated parasite according to the invention and a pharmaceutically acceptable carrier.

15 Vaccines according to the invention can be administered e.g. intradermally, subcutaneously, intramuscularly, intraperitoneally, intravenously, or at mucosal surfaces such as orally or intranasally.

The tet-repressor gene is a gene of prokaryotic origin. The codon usage for this gene is consequently sub-optimal in eukaryotic organisms such as the live attenuated parasites to which the present invention relates. Therefore, man skilled in the art would be motivated to adapt the coding sequence of the tet-repressor gene in such a way that it corresponds to the codon usage of the eukaryotic cell, thus arriving at a synthetic tet-repressor gene. This has been done by Meissner M, et al. (Nucleic Acids Res. 29(22): E115 (2001)).

25 Of course one would expect that this synthetic tet-repressor gene could not be further optimised, since it is already fully adapted to the eukaryotic cell. Moreover, one would expect this "synthetic" tet-repressor protein to be the best suitable repressor protein in the eukaryotic cell. This protein is in principle the same protein as the native protein, and thus by definition best fitted for interaction with the tet-operator site.

30

It was however surprisingly found now, that fusion proteins encoded by a recombinant gene comprising (part of) a heterologous gene fused to the N-terminal part of the native i.e. prokaryotic tet-repressor provide a significantly better regulation of the tet-operator than even the tet-repressor protein encoded by a fully eukaryote-adapted "synthetic" tet-repressor gene.

35

Thus, such fusion proteins would be the repressor proteins of choice to be used in the live attenuated parasites according to the present invention. This is even more an unexpected finding because 3D-structure studies of the tet-repressor protein would predict that N-terminal fusion would negatively interfere with DNA-binding. This was however  
5 surprisingly found not to be the case in practice.

A heterologous gene is any gene that encodes a protein other than the tet-repressor protein. A heterologous protein is any protein other than the tet-repressor protein. A recombinant gene is any artificially made gene that comprises (part of) a heterologous  
10 gene fused to that side of the tet-repressor gene that encodes the N-terminus of the tet-repressor protein.

The fusion protein must be able to reach the nucleus in order to interact with the tet-operator. Therefore there are a number of prerequisites to be fulfilled by the tet-repressor  
15 fusion protein: the final molecular weight of the monomeric tet-repressor fusion protein must be <60 kD, the heterologous part of the fusion protein must be on the N-terminal side of the tet-repressor protein, and the fusion protein must be free of GPI-anchors, secretion/excretion signals and trans-membrane regions. In principle, each and every protein or part thereof that meets with these prerequisites and (as a consequence) is  
20 capable of targeting the nucleus can be used for N-terminal fusion with the tet-repressor protein.

There is no need to use a full length heterologous protein for fusion. It suffices to use a part of such a heterologous protein. A part is considered to be a fragment of at least 10 amino acids, preferably a least 20 amino acids as the heterologous fusion protein.  
25 Preferably, the part originates from the N-terminal side of the heterologous protein. Heterologous proteins of choice are e.g. Green, Red and Yellow Fluorescent protein and the CAT-protein.

Therefore, another embodiment of the present invention relates to DNA-fragments  
30 encoding a tet-repressor fusion protein that has as a characteristic feature that it comprises the tet-repressor protein and a heterologous protein or a part thereof, that is fused at the N-terminal side of the tet-repressor protein wherein the monomeric form of the fusion protein has a size of <60 kD and the fusion protein is free of GPI-anchors, secretion/excretion signals and trans-membrane regions.

35

Still another embodiment of the present invention relates to a tet-repressor fusion protein as such, that has as a characteristic feature that it comprises the tet-repressor protein and a heterologous protein or a part thereof, that is fused at the N-terminal side of the tet-repressor protein wherein the monomeric form of the fusion protein has a size of <60 kD  
5 and the fusion protein is free of GPI-anchors, secretion/excretion signals and trans-membrane regions.

The membranes to which the wording "trans-membrane regions" refers, are those membranes that are located between the cytoplasm of the cell and the outside world.  
10 These membranes specifically exclude the membranes between the nucleus and the cytoplasm. Preferably, the tet-repressor fusion protein according to the invention does have some signals that specifically direct the fusion protein to the nucleus. This is clear, because the tet-repressor fusion protein (as is required for the native tet-repressor gene) has to enter the nucleus in order to be able to regulate the transcription of the gene it  
15 controls.

Due to its universal character, the combination of the tetR-system and the tet-repressor fusion protein can be used not only in live attenuated parasites according to the invention, but certainly also in other parasites and in other eukaryotic cells and organisms. It is  
20 universally applicable in eukaryotic cells, for the regulation of expression of any gene.

Attenuated live parasites according to the invention are thus even more suitable as a basis for vaccines, when such parasites comprise the tet-operator combined with (the genetic information encoding) the tet-repressor fusion protein described above. This  
25 allows an even better blocking and induction of the transcription of a ribosomal gene.

Therefore, in a more preferred form, attenuated live parasites according to the invention in which the induction of the gene is regulated by tetracycline, anhydrotetracyclin or derivatives thereof, comprise the tet-operator and the genetic information encoding a tet-repressor fusion protein as described above.  
30

As will be shown in the examples, the unexpected characteristics of the tet-repressor fusion protein as described above are even more significant if two or more tet-operator sites are cloned in tandem. The wording "in tandem" should be interpreted broadly, in the  
35 sense that tet-operator sites may be cloned directly adjacent to each other or with a spacer sequence in between the two or more tet-operator sites. As mentioned before, the

tet-operator sites are preferably cloned in the region between -100 and +3 relative to the STS.

- Thus, in an even more preferred form, such attenuated live parasites according to the  
5 invention comprise not only the tetR-system and a tet-repressor fusion protein as described above, but also two or more tet-operator sites, instead of one.

## EXAMPLES

## Example 1

*Primers used during the course of the experiments:*

Restriction sites that were inserted are underlined.

SEQ ID NO:	#	NAME	SEQUENCE 5' → 3'
5	1	SAG3-FW	CGATA <u>AAGCTT</u> CGAATCTCTGAACGGATGTGT
6	2	TUB5-RV	CGAGATCTGGGAATTCAAGAAAAAATGCCAACG
7	3	TETAVR5-FW	CGAT <u>CCTAGGAT</u> GTCTAGATTAGATAAAAG
8	4	TETPST3-RV	CGTCTGCAGTTAAGACCCACTTTCACATTTAAG
9	5	T3	ATTAACCCCTCACTAAAGGGAA
10	6	SAG1/1634-RV	CGATA <u>AAGCTT</u> TCGGGGGGGCAAGAATTGTGT
11	7	REV 13A	GCGCCCCATGGTGACGGAGAAAAATCG
12	8	REV 13B (nested primer)	GGGAACCGCAAGGTGGGAGCGGAGAAC
13	9	S13PROMFUS FW	GCATA <u>AAGCTT</u> CCTCGCAGAGATTGTCTAGTG
14	10	S13PROMFUS RV	GCATTCTAGAGGCAGACATGCCCTTTCCAGG
15	11	LACZ-AVR11 FW	CGAT <u>CCTAGGAT</u> GACCATGATTACGGATTCAC
16	12	LACZ-PST1 RV	CGATCTGCAGTTATTTTGTACACCAGACCAA
17	13	S13INSTETO+3FW	GGTTCTCCCCTCAATCCCTATCAGTGATAGAGATCTC TCTTCCTTTCTCT
18	14	S13INSTETO+3RV	AGAGAAAGGAAGAGAGATCTCTATCACTGATAGGGAT TGAGGGGAGAACC
19	15	S13SUBTETO-23FW	CTACGCGGCCGACGGTCCCTATCAGTGATAGAGATCT TCCTCGACGGGTTC
20	16	S13SUBTETO-23RV	GAACCCGTCGAGGAAGATCTCTATCACTGATAGGGAC CGTCGCGCCGCTAG
21		S13NOT1-FW	CGATGCGGCCGCGTCAGTGCATGACACAACCG
22		S13SAC1-RV	GCTAGAGCTCCTGTAAGTCGCCAGAGAAGCAC
23		M13-REV	AACAGCTATGACCATGATTACGC
24		S13CL FW3	CGATAGTGTGCAATAACAGG



25	HRCHECK II 5 S13-FW	GTCGAGTCCTGTAGGTTTCATC
26	HRCHECK II S13-RV	CTCCGAAGGAGTCTCTCAGTG
27	T7	AATACGACTCACTATAG
28	HXGPRT/BGLII-FW	CGATAGATCTAAAATGGCGTCCAAACCCATTG
29	HXGPRT/PSTI-RV	CGATCTGCAGTTACTTCTCGAACTTTTTCG

*Construction of TubYFP/TR-sagCAT (9332 bp).*

Plasmid ptubYFP/TR-sagCAT was engineered stepwise as described below. First the construct ptubCAT/GFP was made by amplifying the *Toxoplasma gondii* tubuline A (tub) promoter from the ptubYFP/YFP-sagCAT construct (Llopis, J. et al., PNAS 97(8): 4363-4368 (2000)) using the primers SAG3FW (#1, SEQ ID NO: 5) and TUB5RV (#2, SEQ ID NO: 6). The PCR product as well as the plasmid pdhfrCAT/GFP (Striepen, B. et al., Molecular and Biochemical Parasitology 92: 325-338 (1998)) were digested with HindIII and BglII, and ligated with each other. This resulted in ptubCAT/GFP where the dhfr promoter has been replaced by the tub promoter. The resulting plasmid is based on Bluescript pKS+® (Stratagene, La Jolla, CA) and contains the  $\alpha$ -tubuline promoter separated from the fusion of chloramphenicol acetyl transferase (CAT) coding sequence with green fluorescent protein coding sequence by a BglII site.

To obtain the ptubYFP/TR construct the CAT coding sequence was exchanged for yellow fluorescent protein (YFP) and the GFP coding sequence was exchanged for tet-repressor coding sequence (tetR). The YFP gene was cut out of the ptubYFP/YFP-sagCAT construct by BglII and AvrII, and ligated between BglII and AvrII site of the ptubCAT/GFP construct replacing the CAT coding sequence. The tetR coding sequence was amplified by PCR from *E. coli* Tn10 (Hillen, W. and Berens, C., Annu. Rev. Microbiol. 48: 345-369 (1994)) using the primers TETAVR5-FW (#3, SEQ ID NO: 7) and TETPST3-RV (#4, SEQ ID NO: 8), digested by AvrII and PstI, and ligated in the construct by exchanging GFP coding sequence for the tetR coding sequence. The resulting plasmid was named ptubYFP/TR.

Finally a CAT selection cassette was inserted upstream of the tub promoter, resulting in the ptubYFP/TR-sagCAT plasmid. This was done by amplification of the CAT-cassette from the ptubYFP/YFP-sagCAT construct mentioned before using the primers T3 (#5, SEQ ID NO: 9) and SAG1/1634 RV (#6, SEQ ID NO: 10), digested with HindIII and ligated into the unique HindIII site of the ptubYFP/TR construct.

The construction of TubYFP/TR-sagCAT and its full sequence are presented in Figure 1.

## Example 2

### *Determination of the start transcription site of the ribosomal protein gene S13 of Toxoplasma gondii*

In order to determine the start of transcription of the ribosomal protein gene S13, RNA was isolated from *Toxoplasma gondii* RH $\Delta$ HXGPRT tachyzoites grown in Vero cells. Using the GeneRacer® kit (Invitrogen) gene specific full-length cDNA was obtained from the total RNA. With this kit an RNA oligo was ligated to the ends of full-length mRNA. After reverse transcription by oligo dT had taken place, amplification by PCR with a GeneRacer primer binding to the RNA oligo together with a gene specific primer resulted in a product. Then the start of transcription (STS) could be determined. This was done for the ribosomal protein gene S13 using the following primers: REV13A (#7, SEQ ID NO: 11) and REV13B (#8, SEQ ID NO: 12). Primer #7 was used together with the GeneRacer primer to get a product after which primer #8 was used for the nested PCR. The PCR product showed three bands; two weak bands and a strong band. The band showing the highest amount of product has been isolated and the STS was determined and indicated as position 0. In Figure 3 A and 3 B, the STS is also represented in relation to the startcodon.

## Example 3

### *S13/LZ constructs*

In order to test inducible expression by the tet repressor several reporter constructs were made with the lacZ gene under control of the S13 promoter with or without the presence of a single tetO site. First the plasmid S13/lacZ was made (see Figure 2 for the structure and sequence of the final construct) and subsequently this plasmid was used to insert or substitute sequences for a tetO site as described below.

The promoter region of S13 was amplified by PCR from the genomic DNA of the *Toxoplasma gondii* RH $\Delta$ HXGPRT strain with the primers S13PROMFUS FW (#9, SEQ ID NO: 13) and S13PROMFUS RV (#10, SEQ ID NO: 14). The lacZ coding sequence was amplified by PCR from the genomic DNA of BL21 bacteria with the primers LACZ-AVR11

FW (#11, SEQ ID NO: 15) and LACZ-PSTI RV (#12, SEQ ID NO: 16). Subsequently the S13 PCR product was digested by HindIII and XbaI while the lacZ PCR product was digested by AvrII and PstI. The plasmid ptubYFP/YFP-sagCAT was used to exchange the ptubYFP part together with the CAT selection cassette for the S13 promoter part. The remaining YFP gene was exchanged for the lacZ gene, resulting in S13/lacZ plasmid. The S13/lacZ plasmid was used to insert or substitute sequences for a single tet operator (tetO) site

(5'-TCCCTATCAGTGATAGAGATC-3') by site-directed mutagenesis. This was done using the QuickChange® site-directed mutagenesis kit (Stratagene). The tetO was inserted or substituted in the vicinity of the determined STS. The primers S13INSTETO+3 FW (#13, SEQ ID NO: 17) and S13INSTETO+3 RV (#14, SEQ ID NO: 18) were used to insert a tetO site at position +3 related to STS, which is indicated as 0. The primers S13SUBTETO-23 FW (#15, SEQ ID NO: 19) and S13SUBTETO-23 RV (#16, SEQ ID NO: 20) were used to substitute sequences for a tetO site between -43 and -23 related to STS. These two constructs, S13instetO+3/lacZ and S13subtetO-23/lacZ together with the S13/lacZ construct have been tested in the *Toxoplasma gondii* strains RHΔHXGPRT, REP1/2 (Meissner, M. et al., Nucleic Acids Research 29 (22), E115 (2001)) and tubYFP/TR by a CPRG assay (Seeber, F. et al., Gene 169: 39-45 (1996)) in the absence or presence of tetR and (anhydro)tetracycline.

The S13/lacZ construct is shown in Figure 2 and the sites of substitution or insertion of the tet operator in the S13/lacZ construct are shown in Figure 3A.

#### *L9/LZ constructs*

tetO insertions / substitutions into the rp-L9 promoter are presented in Figure 3B

#### **Example 4**

##### *Selection of stable transfectant Toxoplasma parasites carrying pTub-YFP-TR-sagCAT.*

Electroporation of *Toxoplasma* parasites was done as described by Roos, D.S. et al. ("Methods in Microbial Pathogenesis" In Methods in Cell Biology (1994), D.G. Russell, editor).

Selection of the stable transfectants was done according to Kim, K., et al. (Science 262(5135): 911-4 (1993)).

Electroporation of S13/LZ, S13i+3/lacZ and S13s-23/lacZ constructs was again done according to Roos, D.S. et al. (1994, supra).

#### Results of Example 4

*Determination of LacZ expression driven by an S13 promoter containing a single tet-operator, electroporated into the tub-YFP-TR strain.*

The following constructs have been tested:

- a) S13/LZ: This is the tub-YFP-TR transfectant Toxoplasma strain, transiently transfected with the LacZ gene under the control of the S13 ribosomal protein gene promoter. There is no tet-operator-site present in this construct.
- b) S13i+3/lacZ: This is the tub-YFP-TR transfectant Toxoplasma strain, transiently transfected with the LacZ gene under the control of the S13 ribosomal protein gene promoter, which additionally carries a tet-operator-site inserted at site +3 relative to the STS (see Figure 3A).
- c) S13s-23/lacZ: This is the tub-YFP-TR transfectant Toxoplasma strain, transiently transfected with the LacZ gene under the control of the S13 ribosomal protein gene promoter which additionally carries a tet-operator-site has been substituted at site -23 relative to the STS (see Figure 3A).

Similar constructs of tetO insertions/substitutions into rp-L9 promoters are presented in Figure 3B.

As can be seen in Figure 4, tub-YFP-TR produces the same level of LacZ in both the presence and absence of anhydro-tetracycline and tetracycline, as expected.

Transfection with construct S13i+3/lacZ results in the production of an amount of LacZ in the absence of anhydro-tetracycline and tetracycline, that is half the amount of LacZ produced in the presence of anhydro-tetracycline and tetracycline.

This clearly shows the inducibility of LacZ-transcription in this strain.

Transfection with construct S13s-23/lacZ results in the production of an amount of LacZ in the absence of anhydro-tetracycline and tetracycline, that is about 1/3 of the amount of LacZ produced in the presence of anhydro-tetracycline and tetracycline.

This again clearly shows the inducibility of LacZ-transcription in this strain.

These results moreover prove that the site at which the tet-operator site is located relative to the STS, is not very critical. It additionally proves that the tet-operator site may be introduced by both insertion and substitution.

*CPRG-assay of transient transfectants electroporated with a construct comprising a LacZ gene driven by an S13 promoter comprising a single tet-operator or a double tet-operator.*

In this assay the following constructs were compared:

- a) S13/LZ as described above
- b) S13s-23/lacZ(I) as described above (= S13s-23/lacZ)
- c) S13s-23/lacZ(II) which equals S13s-23/lacZ except for the fact that an additional tet-operator site has been cloned immediately downstream of the first tet-operator. The construct was assembled using similar techniques as for S13s-23/lacZ(I).

As follows from Figure 5, both the synthetic tet-repressor gene (Meissner) mentioned above and a fusion tet-repressor gene (tub-YFP-TR) according to the invention are capable of blocking the transcription of LacZ in the absence of tetracycline. More strikingly, it clearly follows that the blocking of expression is between 3 and 4 times better when two adjacent tet-operator sites are used compared to the use of a single tet-operator.

*CPRG-assay of transient transfectants comparing LacZ expression in a strain comprising the synthetic tet-repressor gene (Meissner) as described above, and a strain comprising a fusion tet-repressor gene according to the invention.*

As follows surprisingly from Figure 5, a fusion tet-repressor protein according to the invention gives a significantly better blocking of the transcription of LacZ when compared to the blocking found with synthetic tet-repressor protein (Meissner) as described above. Also, surprisingly, a much better induction of LacZ transcription is found with a fusion tet-repressor gene according to the invention when compared to the induction found with synthetic tet-repressor gene (Meissner) mentioned above.

## Example 5

*Insertion of tet operator elements in the ribosomal protein S13 locus using homologous recombination with the hit-and-run mutagenesis procedure.*

To integrate a tet operator site on the genome at a specific locus, in this case the ribosomal protein S13 locus (S13), homologous recombination is required. For homologous recombination a large sequence part (in this case ~1200 bp) upstream and downstream of the integration site is needed to obtain a homologous recombination instead of a non-homologous recombination. As described by Donald et al. (Mol. Biochem. Paras. 91: 295-305 (1998)), it is possible to integrate a sequence element at a specific locus in two steps using the hypoxanthine-xanthine-guanine phosphoribosyltransferase (HXGPRT) gene as a selection marker.

In detail, a transfected plasmid containing part of the S13 locus near the integration site, which is preceded by an HXGPRT cassette will homologously recombine once with the homologous genomic DNA S13 locus, creating a pseudodiploid type I or II (Figure 6). This is performed under positive selection for HXGPRT by mycophenolic acid and xanthine as described (Donald et al. 1998, supra). Subsequently the second homologous recombination is selected with 6-thioxanthine against HXGPRT which results in loss of the pseudodiploid and creation of a tachyzoite with or without a tet operator site integrated at the S13 locus (~1:1 ratio). This procedure is called hit-and-run mutagenesis.

To perform this procedure first a plasmid was made containing an HXGPRT selection cassette under the control of a DHFR promoter. RNA was isolated from *Toxoplasma gondii* RH tachyzoites. This RNA was used for making cDNA using SUPERScript™ II RnaseH- Reverse Transcriptase (Gibco BRL) and standard molecular biological procedures (Sambrook & Russell: "Molecular cloning: a laboratory manual" (2001), Cold Spring Harbor Laboratory Press; ISBN: 0879695773). The HXGPRT coding sequence was amplified from the *T. gondii* RH tachyzoites' cDNA using primers HXGPRT/BGLII-FW (SEQ ID NO: 28) and HXGPRT/PSTI-RV (SEQ ID NO: 29) and splice variant-I was selected for further use (Donald et al., J. Biol. Chem. 271: 14010-14019 (1996)). Both the PCR product and plasmid pdhfrCAT/GFP (Striepen, B. et al., Mol. Biochem. Paras. 92: 325-338 (1998)) were digested with BglII and PstI after which the CAT/GFP coding sequence was exchanged for the HXGPRT coding sequence, resulting in a dhfrHXGPRT construct named pminiHXGPRT.

Subsequently a DNA part containing the area both upstream (~1200 bp) and downstream (~1200 bp) of the tet operator integration site (-43/-23 relative to STS) was PCR amplified

from genomic DNA of *T. gondii* RH tachyzoites using primers S13NOTI-FW (SEQ ID NO: 21) and S13SACI-RV (SEQ ID NO: 22). Both this PCR product and the pminiHXGPRT were digested by NotI and SacI after which the PCR product was ligated downstream of the HXGPRT cassette. Finally, a tet operator was inserted with primer S13SUBTETO-23FW (SEQ ID NO: 19) and primer S13SUBTETO-23RV (SEQ ID NO: 20) by substitution using site-directed mutagenesis as described in Example 3, creating pS13s-23/pminiHXGPRT.

Circular pS13s-23/pminiHXGPRT plasmid was electroporated as described previously (Example 4) into RH $\Delta$ HXGPRT tachyzoites. After infection into Vero cell monolayers, mycophenolic acid / xanthine selection was started as described by Donald et al. (1998, supra).

After stable transfectants were generated according to Kim, K., et al. (Science 262 (5135): 911-4 (1993)), several clonal parasite lines were picked. Genomic DNA was isolated from each of these clones. PCR analysis was performed on these genomic DNA samples to check for the presence or absence of the pseudodiploid form in these transfectants using the primers M13-REV (SEQ ID NO:23), S13CL FW3 (SEQ ID NO: 24), HRCHECK II 5 S13-FW (SEQ ID NO: 25), HRCHECK II S13-RV (SEQ ID NO: 26), and T7 (SEQ ID NO: 27). Four clones (c4, c5, c6 and c9) were analysed in detail and the genomic DNA of the strain RH $\Delta$ HXGPRT and of Vero cells was used as a negative control. Different primer combinations (Figure 6) were used to amplify by PCR the genomic DNA of these samples, these are listed as: 23/24, 25/26, 23/26, and 25/27, meaning the combination of primers of SEQ ID NO: 23 and 24, etc.. Results are presented in Figure 7.

Primer combination 23/24 shows the presence of the plasmid in the different clones.

Primer M13-REV (SEQ ID NO: 23) anneals to the vector part which is absent in untransfected parasites (RH $\Delta$ HXGPRT). All transfected clones show bands of the correct size (2.8 kb), indicating that all stable transfectants have taken up the plasmid after electroporation and kept it during selection. Subsequently, the primer combination 25/26 shows whether the pseudodiploid form is present in the clones. On the genome, both primers are located upstream (primer HRCHECK II 5 S13-FW (SEQ ID NO: 25)) or downstream (primer HRCHECK II S13-RV (SEQ ID NO: 26)) of the S13 part present in the vector. If the pseudodiploid is not present, the "wild type" S13 situation will be PCR amplified which results in a product of ~2.6 kb as can be observed with clone c4 and the wild type parasite RH $\Delta$ HXGPRT. This shows that clone c4 is a stable transfectant without a pseudodiploid, suggesting that non-homologous recombination occurred. The absence of the 2.6 kb PCR product for the clones c5, c6 and c9 indicates that these clones do contain the pseudodiploid form. In addition, a product of nearly 10 kb can be observed for

clones c5 and c9, which is as expected when the pseudodiploid is present. No 10 kb product was detectable for clone c6. The PCRs with primer combination 23/26 and primer combination 25/27 were performed to demonstrate that the p13s-23/pminiHXGPRT vector is juxtaposed at both sides by the S13 locus. Primer M13-REV (SEQ ID NO: 23) is located in the vector sequence and primer HRCHECK II S13-RV (SEQ ID NO: 26) is located on the genome downstream of the homologous S13 part of the vector. Primer T7 (SEQ ID NO: 27) is located in the vector sequence and primer HRCHECK II 5 S13-FW (SEQ ID NO: 25) is located on the genome upstream of the homologous S13 part of the vector. In the wild type situation the primer combination 23/27 does not anneal to DNA and for that reason no PCR product can be amplified. In case of a pseudodiploid the primer combination 23/26 results in a product of 4.6 kb and the combination 25/27 results in a PCR product of 2.6 kb. The data presented in Figure 7 demonstrate that indeed the positive clones show the right bands for both combinations whereas for the negative samples no products were observed.

This PCR analysis is therefore used to verify that the homologous recombination into e.g. the S13 locus by the hit-and-run mutagenesis procedure is performed successfully.



### Legend to the Figures.

Figure 1: Description of the TubYFP/TR-sagCAT construct:

Figure 1A: Full sequence: relevant regions are indicated below the sequence; restriction enzyme recognition sites are indicated above the sequence,

Figure 1B: List of relevant features and regions of the TubYFP/TR-sagCAT construct,

Figure 1C: Graphical map of the TubYFP/TR-sagCAT construct.

Figure 2: Description of the S13/lacZ construct:

Figure 1A: Full sequence: relevant regions are indicated below the sequence; restriction enzyme recognition sites are indicated above the sequence,

Figure 1B: List of relevant features and regions of the S13/lacZ construct,

Figure 1C: Graphical map of the S13/lacZ construct.

Figure 3:

Figure 3A: tetO insertions/substitutions in rp-S13 promoter:

Sequence of part of the ribosomal protein S13-promoter, also indicating the site of the +3 insertion and the -23 substitution, relative to the STS. Also indicated are the first three amino acids of the coding region.

Figure 3B: tetO insertions/substitutions in rp-L9 promoter:

Figure 4: Determination of the level of LacZ expression by tubYFP/TR stable transfectants electroporated with the constructs S13/LZ, S13i+3/lacZ and S13s-23/lacZ without antibiotics, in the presence of 1 µg/ml anhydro-tetracycline (Atc) or in the presence of 1 µg/ml tetracycline (Tc). The OD is an indication for the level of LacZ expression. The labels of the horizontal axis indicate that  $1.25 \times 10^6$  tachyzoites were used (50 % of originally made amount).

Figure 5: Determination of the LacZ expression level in different strains (RH, REP, tubYFP/TR) electroporated with the constructs S13/LZ, S13s-23/lacZ(I) and S13s-23/lacZ(II)

RH represents the strain without tet-repressor gene. REP represents the strain carrying the synthetic tet-repressor gene (Meissner). TYT represents the strain carrying the fusion tet-repressor gene (tub-YFP-TR). Equal amounts of cells have been used in these

comparative experiments. Experiments have been done in the absence or presence of tetracycline as indicated in the figure.

Figure 6: Formation of type I and II pseudodiploid forms after first step of hit-and-run mutagenesis:

Figure 7: PCR on genomic DNA of different clones to determine presence of pseudodiploid forms

**Claims**

- 1) Attenuated live parasite of the phylum Apicomplexa or the family of Trypanosomatidae, characterised in that said parasite comprises a ribosomal protein gene under the control of an inducible promoter.
- 2) Attenuated live parasite according to claim 1, characterised in that said parasite belongs to the Coccidia, the Piroplasmida or the Haemosporida.
- 3) Attenuated live parasite according to claim 2, characterised in that said parasite belongs to the family of the Eimeriidae, Cryptosporidiidae or Sarcocystidae.
- 4) Attenuated live parasite according to claim 3, characterised in that said parasite belongs to the genus Eimeria, Cryptosporidium, Toxoplasma, Sarcocystis or Neospora.
- 5) Attenuated live parasite according to claim 2, characterised in that said parasite belongs to the family of the Babesiidae or the Theileriidae.
- 6) Attenuated live parasite according to claim 5, characterised in that said parasite belongs to the genus Babesia or Theileria.
- 7) Attenuated live parasite according to claim 2, characterised in that said parasite belongs to the genus Plasmodium.
- 8) Attenuated live parasite according to claim 1, characterised in that said parasite belongs to the genus Trypanosoma or the genus Leishmania
- 9) Attenuated live parasite according to claims 1-8, characterised in that said inducible promoter is based upon an operator site and a repressor protein capable of reversibly binding said operator site.
- 10) Attenuated live parasite according to claims 1-9, characterised in that said inducible promoter is inducible by antibiotics.

- 11) Attenuated live parasite according to claim 10, characterised in that said inducible promoter is inducible by tetracycline or anhydrotetracyclin, or a derivative thereof.
- 12) Attenuated live parasite according to claim 11, characterised in that a tetR-system is used as the inducible promoter.
- 13) Attenuated live parasite according to claims 1-12, characterised in that said ribosomal protein gene is the gene encoding L9, S3, plastid-S9 or S13, preferably L9, S3, plastid-S9 or S13 of *Toxoplasma gondii*.
- 14) Attenuated live parasite according to claims 1-13 for use in a vaccine.
- 15) Vaccine for combating parasitic infection characterised in that said vaccine comprises a live attenuated parasite according to claims 1-13 and a pharmaceutically acceptable carrier.
- 16) Use of an attenuated live parasite according to claims 1-13 for the manufacture of a vaccine for combating infection caused by a parasite of the phylum Apicomplexa or the family of Trypanosomatidae.
- 17) Method for the production of a vaccine according to claim 15, said method comprising the mixing of a live attenuated parasite according to claims 1-13 and a pharmaceutically acceptable carrier.
- 18) DNA-fragment encoding a tet-repressor fusion protein comprising the tet-repressor protein and a heterologous protein or a part thereof, said heterologous protein or a part thereof being fused to the N-terminal side of the tet-repressor protein, the monomeric form of said fusion protein having a molecular weight of less than 60 kD and being free of GPI-anchors, secretion/excretion signals and trans-membrane regions.
- 19) Attenuated live parasite according to claims 1-13, characterised in that said parasite comprises the tet-operator site and a DNA fragment encoding a tet-repressor fusion protein according to claim 18.

- 20) Attenuated live parasite according to claim 19, characterised in that said parasite comprises two or more tet-operator sites.

## Figures

Figure 1 A

```
1  gtggcacttt tcggggaaat gtgcgcgga cccctatttg tttatttttc taaatacatt
61  caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa
121 ggaagagtat gagtattcaa catttccgtg tcgccccttat tccctttttt gcggcatttt
181 gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt
241 tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt
301 ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg
361 tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga
421 atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa
481 gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac ttacttctga
541 caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa
601 ctgcgcctga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca
661 ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta
721 ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac
781 ttctgcgctc ggcccttcog gctggctggt ttattgctga taaatctgga gccggtgagc
841 gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag
901 ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga
961 taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt
1021 agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata
1081 atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag
1141 aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa
```

```

1201 caaaaaaacc accgctacca gcggtgggtt gtttgccgga tcaagagcta ccaactcttt
1261 ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc
1321 cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa
1381 tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa
1441 gacgatagtt accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc
1501 ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa
1561 gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa
1621 caggagagcg cagcagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg
1681 ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctgcgcaggg gggcggagcc
1741 tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg
1801 ctcacatggt ctttcctgcg ttatccctg attctgtgga taaccgtatt accgcctttg
1861 agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg
1921 aagcggaaga gcgccaata cgcaaacgc ctctccccgc gcgttgggcg attcattaat
1981 gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg
2041 tgagttagct cactcattag gcacccagc ctttacactt tatgcttcg gctcgtatgt
2101 tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg
                                     KpnI
                                     -----
2161 ccaagcgcgc aattaaccct cactaaaggg aacaaaagct gggtagcggg cccccctcgc
      >>.....T3 .....>>
                                     HindIII
                                     -----
                                     PstI
                                     -----
2221 aggtcgacgg tatcgataag cttgatatcg aattcctgca gccccgaga cgcgtgttct
                                     >>.pSAG1.>
2281 aaccacaaac cttgagacgc gtgttccaac cagcaccct gacacgcgtg ttccaaccac
      >.....pSAG1.....>
2341 gcaccctgag acgcgtgttc tawccacgca ccctgagacg cgtgttctaa ccacgcaccc
      >.....pSAG1.....>

```

```

2401  tgagacgcgt gttctgccgc acaatgtgca cctgtaggaa gctgtagtca ctgctgattc
>.....pSAG1.....>

2461  tcaactgttct cggcaagggc cgacgaccgg agtacagttt ttgtgggcag agccgttgtg
>.....pSAG1.....>

2521  cagctttccg ttcttctcgg ttgtgtcaca tgtgtcattg tcgtgtaaac acacggttgt
>.....pSAG1.....>>

2581  atgtcggttt cgctgcacca cttcattatt tcttctgggt ttttgacgag tatgcatgag
>>>
>>.....SAG1 CDS.....>>
>>>
>>>

2641  aaaaaaatca ctggatatac caccgttgat atatcccaat ggcatcgtaa agaacatttt
>.....CAT CDS.....>

2701  gaggcatttc agtcagttgc tcaatgtacc tataaccaga ccgttcagct ggatattacg
>.....CAT CDS.....>

2761  gccttttttaa agaccgtaaa gaaaaataag cacaagtttt atccggcctt tattcacatt
>.....CAT CDS.....>

2821  cttgcccgcc tgatgaatgc tcatccggaa ttccgtatgg caatgaaaga cggtgagctg
>.....CAT CDS.....>

2881  gtgatatggg atagtgttca cccttgttac accgttttcc atgagcaaac tgaaacgttt
>.....CAT CDS.....>

2941  tcatcgctct ggagtgaata ccacgacgat ttccggcagt ttctacacat atattcgcaa
>.....CAT CDS.....>

3001  gatgtggcgt gttacggtga aaacctggcc tatttcccta aagggtttat tgagaatatg
>.....CAT CDS.....>

3061  tttttcgtct cagccaatcc ctgggtgagt ttcaccagtt ttgatttaaa cgtggccaat
>.....CAT CDS.....>

3121  atggacaact tcttcgcccc cgttttcacc atgggcaaatt attatacgca aggcgacaag
>.....CAT CDS.....>

3181  gtgctgatgc cgtggcgat tcaggttcat catgccgttt gtgatggctt ccatgtcggc
>.....CAT CDS.....>

3241  agaatgctta atgaattaca acagtactgc gatgagtggc agggcggggc ktaaktratr
>.....CAT CDS.....>>
>>....>

3301  accgttgtgc tcaattctca aatcgacaaa ggaaacacac ttcgtgcagc atgtgcccc
>.....SAG1-I.....>

```



```
3361 ttataaagaa actgagttgt tccgctgtgg cttgcaggtg tcacatccac aaaaaccggc
>.....SAG1-I.....>

3421 cgactctaaa taggagtgtt tgcagcaag cagcgaaagt ttatgactgg gtccgaatct
>.....SAG1-I.....>

3481 ctgaacggat gtgtggcgga cctggctgat gttgatcgcc gtcgacacac gcgccacatg
>.....SAG1-I.....>

3541 ggtcaatata caagacagct atcagttgtt ttagtcgaac cggttaacac aattcttgcc
>.....SAG1-I.....>
      HindIII
      -----
3601 cccccgaaag cttcgaatct ctgaacggat gtgtggcgga cctggctgat gttgatcgcc
>....>>
      >>.....SAG1-II.....>

3661 gtcgacacac gcgccacatg ggtcaatata caagacagct atcagttgtt ttagtcgaac
>.....SAG1-II.....>

3721 cggttaacac aattcttgcc ccccgaggg ggatccacta gttctagagc ggccctcagg
>.....SAG1-II.....>>
      >>.....vector.....>

3781 tcgacgggat cgataagcta gagcttcagc atcatctctg gaagcatccc ctgaactgcc
>.....vector.....>>
      >>.....pTUB1.....>

3841 tgagtctacc aagagcactg gcgaaggctg tgagtagtcg gacaggcacg gtgactcatg
>.....pTUB1.....>

3901 ttgttgga gtagcgagct ctgggttaac cgcataattca ctaactggct ccgtcctgtt
>.....pTUB1.....>

3961 gtcattagat actgaatcag gtaacgatac atgagcagca tcctcgtgtt ccaggcgcat
>.....pTUB1.....>

4021 gtccctgctcc ggcttgcaac caaggaccgc tggttcatct ttgggtctct ccgtactggg
>.....pTUB1.....>

4081 tggtagaggt gaaactgtcg acgtggatgc agctgccctg cttagagtac ggacgaagtg
>.....pTUB1.....>

4141 aatagctgcg tctgcatgaa caaggggctc tgaggccgcg tgtgatacga aaggtttgct
>.....pTUB1.....>

4201 ggctactgaa cataggtctc gcagtgcggg ggcatactcc agtcggcctt caccgaaactt
>.....pTUB1.....>

4261 cgtgaccagg cgatcaacaa gcggcgtcac gcgagtaaa tccgtccaga gaaagccacc
>.....pTUB1.....>

4321 atagtgcacc atatactgcc ggcacatctt gctgaaagtc gtggcggtgc gatcagctaa
>.....pTUB1.....>
```

4381 gtcctgtgtg tacttggaag tgagcagctc ttccagttcg gtccagtcac attgccatgc  
>.....pTUB1.....>

4441 cctagccacc gtttcaccag cgtccgatgc aggctcacct agccgatacc tgtgaccgtt  
>.....pTUB1.....>

4501 catatatgag ttgctattac atctgtcggc cgtaacgaca caaggagatg cggctggagg  
>.....pTUB1.....>

4561 caacggggtt gtgagcacca ttttgactct gtcacggagt aaaaaacatt tattcgaact  
>.....pTUB1.....>

4621 ttgtacgagc gcagtcagta gaagtcaacc acgctatca actacgctgc aattacagag  
>.....pTUB1.....>

4681 gacagggaca gggaaaaaaa gccgaagagg ttgccggtgt taggagatga cgagacgttt  
>.....pTUB1.....>

4741 aaacggccgc tggctattgt tcgggcatcg ttgagccaca agcgatcaag gtgaaaacaa  
>.....pTUB1.....>

4801 agttaaatag ttatgctgga gcgattgcc tgctgaatct tcagatcgga cgaacagtca  
>.....pTUB1.....>

4861 ttctggccca ctgtacttga tgtgttcgat gtaaatacaca ctcagtcgag tgctcggtag  
>.....pTUB1.....>

4921 caatcaagtt gctcttttct ctcttttcta gacacggtaa gaacgcttat gaacacgcat  
>.....pTUB1.....>

4981 acacgcatag tttttgctag aatgcagcga ccagatgtcg caaggctcgc tcccccatcg  
>.....pTUB1.....>

5041 actggagaat caagaaaaac ctgcgttgat cccaaacgta ctctgtgggt ggtgcaacca  
>.....pTUB1.....>

5101 gaagtttcat actgatcaaa agccagtga cagctggggg acattgcagg tctggtgctt  
>.....pTUB1.....>

5161 caagaagcgc tgaagaagaa agtggcgaaa ccctcggcag ttgccttgga agaggcgccg  
>.....pTUB1.....>

5221 tgcagtgaac ttctgaagtg cgtagtacct ggctcctaata gctgttttgt gtttcgctgt  
>.....pTUB1.....>

5281 ctgggcagca gtagaatgct gtgccagaat tagccactat tttagacatt tatttacaca  
>.....pTUB1.....>

5341 ttttttttct gatgaacttg gcttattcat tttttcaagt cttgccactg ggtggtggca

>.....pTUB1.....>

5401 tgagactcgc ttagatgtat gtgggtgttg caatcacgct ggatgctcgg cttatttctg  
>.....pTUB1.....>

5461 agttttttgt ggttttgaca atgggaacga tttcagagct actatttcac gtggtacggg  
>.....pTUB1.....>

5521 tatgagccac taaaaaacg aagaaaaacg ctgtttgag aaacaatagc aaactgtttt  
>.....pTUB1.....>

5581 tcgtcatagt aactcagcgc ccccttgccc ccccccagca gtgagatgca agacaatcct  
>.....pTUB1.....>

5641 ctctaccaca gcttttgttg cgtctgtttc aaattttcag cgctcgcgaa aggcatacag  
>.....pTUB1.....>

5701 aacaacatta tgagagggag caggtttgtg gggctggcgg gtgcaggaat gtgttcctgc  
>.....pTUB1.....>

5761 gaaaaaggcc tctgctgaga aggtcgtggc gtttgaaaaa tatccgaggt agcaaagact  
>.....pTUB1.....>

5821 tgtttcagtg ctcccctttt gaagacctgc ggcggcagtg cactgaagag taactccaaa  
>.....pTUB1.....>

5881 tcaccgcggt gagacttggg tttttccggt atccttcaga agagtgtgtt ttcgtttaat  
>.....pTUB1.....>

5941 tcgtcacaga ccacgaaaaa cgaaccatcg aagacgatca ctgcgtccgc gtgcatctgg  
>.....pTUB1.....>

6001 atggatgacc cacatctgtt gcagccgtcg cagacatgca tgtcccgcgt tcgtgaaatt  
>.....pTUB1.....>

6061 ctctgcatca gcggagtgat caggaatcat cgtctcagcg ggatgacgtt gcggagcagg  
>.....pTUB1.....>

6121 ccggctcgcg ggggcagtca gatgccgaag gcgtaactca ggacgcttgc gctcatcgca  
>.....pTUB1.....>

6181 gaacaggggt ggtgcctgca ttgggtgccc ttggtgatcc tggttggacc ggtggagatg  
>.....pTUB1.....>

6241 cgcgcgcacg aaggggatgt gtcagaaaca ttttgtttgt tctctgtgaa ctttttagatg  
>.....pTUB1.....>

6301 tgtaaaggc ggcgaatatt agcagagagt cctccttggt ggattctctc ttgaatttcg  
>.....pTUB1.....>

```

6361 ccctttcctt ctctttgcga gtctcgtaga gaacaagcac tcgttcgccg tccctgacga
>.....pTUB1.....>

6421 cgcaaccgc gcagaagaca tccaccaaac ggtgttacac aatcaccttg tgtgaagttc
>.....pTUB1.....>

                                     BglII
                                     -----
6481 ttgcggaaaa ctactcgttg gcattttttc ttgaattccc agatctaaaa tggtagagcaa
>.....pTUB1.....>>
                                     >>.YFP CDS.>
                                     >>>>

6541 gggcgaggag ctgttcaccg ggggtgtgcc catcctggtc gagctggacg gcgacgtaaa
>.....YFP CDS.....>

6601 cggccacaag ttcagcgtgt ccggcgaggg cgagggcgat gccacctacg gcaagctgac
>.....YFP CDS.....>

6661 cctgaagttc atctgcacca ccggcaagct gcccgtgcc tggcccaccc tcgtgaccac
>.....YFP CDS.....>
                                     PstI
                                     -----
6721 cttcggctac ggcctgcagt gcttcgccc ctaccccgac cacatgaagc agcacgactt
>.....YFP CDS.....>

6781 cttcaagtcc gccatgccc aaggctacgt ccaggagcgc accatcttct tcaaggacga
>.....YFP CDS.....>

6841 cggcaactac aagaccgcg ccgaggtgaa gttcgagggc gacaccctgg ygaaccgcat
>.....YFP CDS.....>

6901 cgagctgaag ggcctcgact tcaaggagga cggcaacatc ctggggcaca agctggagta
>.....YFP CDS.....>

6961 caactacaac agccacaacg tctatatcat ggccgacaag cagaagaacg gcatcaaggt
>.....YFP CDS.....>

7021 gaacttcaag atccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca
>.....YFP CDS.....>

7081 gcagaacacc cccatcggcg acggcccctg gctgctgccc gacaaccact acctgagcta
>.....YFP CDS.....>

7141 ccagtccgcc ctgagcaaag accccaacga gaagcgcgat cacatggtcc tgctggagtt
>.....YFP CDS.....>
                                     AvrII
                                     -----
7201 cgtgaccgcc gccgggatca ctctcgccat ggacgagctg tacaagccta ggatgtctag
>.....YFP CDS.....>>
                                     >>.....>
                                     >>>

7261 attagataaa agtaaagtga ttaacagcgc attagagctg cttaatgagg tcggaatcga
>.....TR CDS.....>

```

7321 aggtttaaca acccgtaaac tcgcccagaa gctaggtgta gagcagccta cattgtattg  
>.....TR CDS.....>

7381 gcatgtaaaa aataagcggg ctttgctcga cgccttagcc attgagatgt tagataggca  
>.....TR CDS.....>

7441 ccatactcac ttttgccctt tagaagggga aagctggcaa gattttttac gtaataacgc  
>.....TR CDS.....>

7501 taaaagtttt agatgtgctt tactaagtca tcgcatgga gcaaaagtac atttaggtac  
>.....TR CDS.....>

7561 acggcctaca gaaaaacagt atgaaactct cgaaaatcaa ttagcctttt tatgccaaaca  
>.....TR CDS.....>

7621 aggtttttca ctagagaatg cattatatgc actcagcgct gtggggcatt ttactttagg  
>.....TR CDS.....>

7681 ttgcgtattg gaagatcaag agcatcaagt cgctaaagaa gaaagggaac cacctactac  
>.....TR CDS.....>

7741 tgatagtatg ccgccattat tacgacaagc tatcgaatta ttgatcacc aagggtgcaga  
>.....TR CDS.....>

7801 gccagccttc ttattcggcc ttgaattgat catatgcgga ttagaaaaac aacttaaattg  
>.....TR CDS.....>  
PstI  
-----

7861 tgaaagtggg tcttaactgc agcccacaga agctgccggt ctctcgtttt cctctctttt  
>....TR CDS....>  
>>.....DHFR.....>

7921 cggagggatc agggagagtg cctcgggtcg gagagagctg acgagggggg gccagagacc  
>.....DHFR.....>

7981 cctgtgtcct ttatcgaaga aaagggatga ctcttcatgt ggcatttcac acagtctcac  
>.....DHFR.....>

8041 ctgcgccttg tttctttttg tcaatcagaa cgaaagcgag ttgcgggtga cgcagatgtg  
>.....DHFR.....>

8101 cgtgtatcca ctcgtgaatg cggtatcggt ctgtatgccg ctagagtgtg ggactgttgc  
>.....DHFR.....>

8161 tgtctgccca cgacagcaga caactttcct tctatgcact tgcaggatgg tgcagcgcaa  
>.....DHFR.....>

8221 acgacggaga gaaaggagca ccctctcagt ttccctacga tgtgctgtca gtttcgactc  
>.....DHFR.....>

8281 ttcaccgcga acgattggcg atacgtctct gttgacttgt taggctccga ccacgaagct  
>.....DHFR.....>



Figure 1 B

**TubYFP/TR-sagCAT**Molecule Features:

Type	Start	End	Name	Description
GENE	2172	2192	T3	T3 primer for sequencing
REGION	2271	2580	pSAG1	SAG1 5' region including promoter
REGION	2581	2583	SAG1 ATG-I	first ATG
GENE	2581	2634	SAG1 CDS	SAG1 coding sequence
REGION	2632	2634	SAG1 ATG-II	second ATG
GENE	2638	3294	CAT CDS	chloramphenicol acetyl transferase coding sequence
REGION	3295	3607	SAG1-I	SAG1 3' untranslated region
REGION	3614	3747	SAG1-II	repeated part of 3' untranslated region used to start tub promoter
REGION	3748	3799	vector	part of unknown vector
REGION	3800	6520	pTUB1	TUB1 5' region including promoter
GENE	6530	7246	YFP CDS	Yellow fluorescent protein coding sequence
REGION	6530	6532	YFP ATG-I	first ATG
GENE	7253	7876	TR CDS	Tet repressor coding sequence
REGION	7253	7255	TR ATG-I	first ATG
REGION	7886	8656	DHFR	DHFR 3' untranslated region
GENE	8710	8690	C T7	T7 primer for sequencing

Figure 1 C

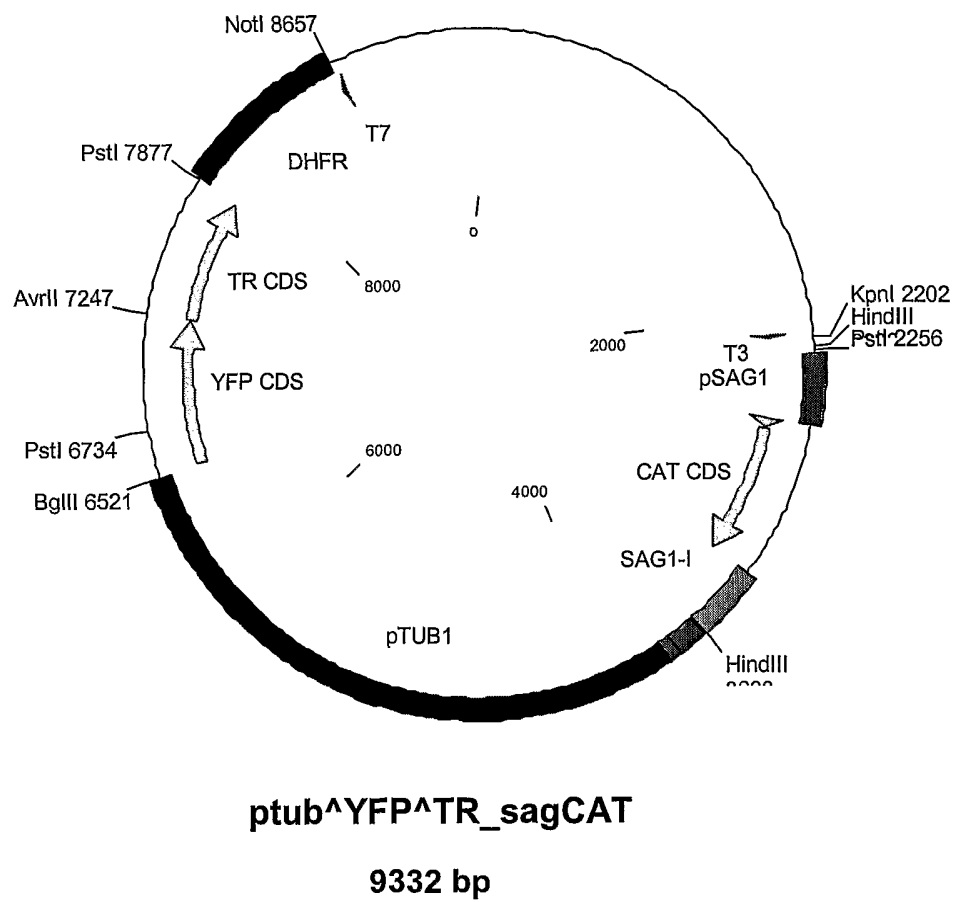




Figure 2 A

```

1  agcttcctcg cagagattgt cagtgcattg cacaaccgcg aaaagccggc agccgcggta
   >.....'promoter.....>

61  atacggggac gaggaaaacg actgagcgtc acaacagaag cagccgagta aacggcgaag
   >.....'promoter.....>

121 gaaatggaaa ggacccaagt aaaatttctt gaagaatttc agcgcaacaa ctctgcgggt
   >.....'promoter.....>

181 tcttgccaat agaggaattt cacttcctca tcgtctgatt tatgctttca tcatctgcg
   >.....'promoter.....>

241 ctcaacagcc gaataaacgg ttctcggtcg cttccttaaa ctctacttca gtagttgaaa
   >.....'promoter.....>

301 ctcttttgct tcacgagcct tcgtctcagc cctcaccgct ctgagttctg tctttgttga
   >.....'promoter.....>

361 ggaaagctcc cgctgaaaaa acaggacttt gtttgagat tttcatgtgt actggaaagt
   >.....'promoter.....>

421 gagatgtgac ttggggaagt ccgctttaa atttccattg tttctcaaa atgaaaagt
   >.....'promoter.....>

                                     PstI
                                     -----
481 taaaaaatcg aagtgcgtgc ccgcgagga attccctct gcagatttgt tttgcattta
   >.....'promoter.....>

541 tatgtcgttt ttacggagaa aagtcccaag ctgctgctcc ttctctaact agatgttgaa
   >.....'promoter.....>

601 cgctagcaca tatgcaccag atgcttctga agtataccta aacgcacctt gggaacaact
   >.....'promoter.....>

661 gtgctcccat tcataaaact catacaagtc accaagcatg ccataccggt gagacataac
   >.....'promoter.....>

721 aacggaagct agactactcc cccctgttat tgcacactat cgaaaaggat tcctaggttt
   >.....'promoter.....>

                                     PstI
                                     -----
781 ctatcctctg ctttttctg gggcacactg cagagaaact accgtgcgcg ctacctccg
   >.....'promoter.....>

841 acgtgcgagg cgatagcaaa acgcttttga aggaaaaagt cgagaaatcg acgactgcgt
   >.....'promoter.....>

```

```
901 ctcttgaatc cgagagaggg atccaaccca ccgagttctc tgcattgtgca gcatctgcaa
>.....'promoter.....>

961 gaacgtgata atgcatgaac tcgatcatcg ctttatctgt gtgcatgcat ttctgaaaaa
>.....'promoter.....>

1021 gaaaggcggt ttctgcgcgg agactcgcgc ggaggcaaga cgagactttc tcctcttcca
>.....'promoter.....>

1081 aactgcgagc cacgggggcg catgcaattt gaacatcacg caaaatccca aaacgggtgg
>.....'promoter.....>

1141 ggtggagccg caaacttttt tggcatgcag cgttgagcct gagctgcggg gggggctttt
>.....'promoter.....>

1201 gtcgcgagcg tggggtgccg cgagagagca acgcggcgct acgcggccga cgggtcttct
>.....'promoter.....>

1261 gggaagcctc gcatttcctc gacgggttct cccctcaatt ctcttccttt ctctgcgtct
>.....'promoter.....>>
>>.....e.....>

1321 tcctcaggtg gcttcgtcac cggtttttct cctcgcgttc gtgctccgct gtgtgtccgg
>.....e.....>

1381 agtgccgca cagatcgagg gcgttctccg ctcccacctt gcggttccca atttcgattt
>.....e.....>

1441 ttctcgtca ccatggggcg catgtacggt cctggaaagg gcatgtctgc ctctaggatg
>.....e.....>>
>>>
>>.....cds.....>>
>>...>

1501 accatgatta cggattcact ggccgtcgtt ttacaacgtc gtgactggga aaaccctggc
>.....'LacZ.....>

1561 gttaccaaac ttaatgcct tgcagcacat ccccttttcg ccagctggcg taatagcgaa
>.....'LacZ.....>

1621 gaggccgca ccgatcgccc ttccaacag ttgcgcagcc tgaatggcga atggcgcttt
>.....'LacZ.....>

1681 gcctggtttc cggcaccaga agcggtgccg gaaagctggc tggagtgcga tcttcctgag
>.....'LacZ.....>

1741 gccgatactg tcgtcgtccc ctcaaactgg cagatgcacg gttacgatgc gccatctac
>.....'LacZ.....>

1801 accaacgtga cctatcccat tacgggtcaat ccgccgtttg tcccacgga gaatccgacg
>.....'LacZ.....>
```

1861 ggttgttact cgctcacatt taatgttgat gaaagctggc tacaggaagg ccagacgcga  
>.....'LacZ.....>

1921 attatatttg atggcgtaa ctcggcggtt catctgtggt gcaacggcg ctgggtcggt  
>.....'LacZ.....>

1981 tacggccagg acagtcgttt gccgtctgaa ttgacctga gcgcattttt acgcgccgga  
>.....'LacZ.....>

2041 gaaaaccgcc tcgcggtgat ggtgctgcgc tggagtgcg gcagttatct ggaagatcag  
>.....'LacZ.....>

2101 gatatgtggc ggatgagcgg cattttccgt gacgtctcgt tgctgcataa accgactaca  
>.....'LacZ.....>

2161 caaatcagcg atttccatgt tgccactcgc tttaatgatg atttcagccg cgctgtactg  
>.....'LacZ.....>

2221 gaggctgaag ttcagatgtg cggcgagttg cgtgactacc tacgggtaac agtttcttta  
>.....'LacZ.....>

2281 tggcaggggtg aaacgcaggt cgccagcggc accgcgcctt tcggcgggtga aattatcgat  
>.....'LacZ.....>

2341 gagcgtgggtg gttatgccga tcgcgtcaca ctacgtctga acgtcgaaaa cccgaaactg  
>.....'LacZ.....>

2401 tggagcgccg aaatcccgaa tctctatcgt gcggtgggtg aactgcacac cgccgacggc  
>.....'LacZ.....>

2461 acgctgattg aagcagaagc ctgcgatgtc ggtttccgcg aggtgcggat tgaaaatggt  
>.....'LacZ.....>

2521 ctgctgctgc tgaacggcaa gccgttgctg attcgaggcg ttaaccgtca cgagcatcat  
>.....'LacZ.....>

2581 cctctgcatg gtcaggatcat ggatgagcag acgatggtgc aggatattcct gctgatgaag  
>.....'LacZ.....>

2641 cagaacaact ttaacgccgt gcgctgttcg cattatccga accatccgct gtggtacacg  
>.....'LacZ.....>

2701 ctgtgcgacc gctacggcct gtatgtgggtg gatgaagcca atattgaaac ccacggcatg  
>.....'LacZ.....>

2761 gtgccaatga atcgtctgac cgatgatccg cgctggctac cggcgatgag cgaacgcgta  
>.....'LacZ.....>

2821 acgcgaatgg tgcagcgcga tcgtaatcac ccgagtgtga tcctctggtc gctggggaat  
>.....'LacZ.....>

2881 gaatcaggcc acggcgctaa tcacgacgcg ctgtatcgct ggatcaaacc tgctgatcct  
>.....'LacZ.....>

2941 tcccgcgccg tgcatgatga aggcggcgga gccgacacca cggccaccga tattatttgc  
>.....'LacZ.....>

3001 ccgatgtacg cgcgcgtgga tgaagaccag cccttcccgg ctgtgccgaa atgggtccatc  
>.....'LacZ.....>

3061 aaaaaatggc ttctgctacc tggagagacg cggccgctga tcctttgcga atacgcccac  
>.....'LacZ.....>

3121 gcgatgggta acagtcttgg cggtttcgct aaatactggc aggcgtttcg tcagtatccc  
>.....'LacZ.....>

3181 cgtttacagg gcggcttcgt ctgggactgg gtggatcagt cgctgattaa atatgatgaa  
>.....'LacZ.....>

3241 aacggcaacc cgtggtcggc ttacggcggg gatthttggcg atacgccgaa cgatcgccag  
>.....'LacZ.....>

3301 ttctgtatga acggtctggg ctttgccgac cgcacgccgc atccagcgct gacggaagca  
>.....'LacZ.....>

3361 aaacaccagc agcagttttt ccagttccgt ttatccgggc aaaccatcga agtgaccagc  
>.....'LacZ.....>

3421 gaatacctgt tccgtcatag cgataacgag ctccctgcact ggatgggtggc gctggatggg  
>.....'LacZ.....>

3481 aagccgctgg caagcgggta agtgccctctg gatgtcgctc cacaaggtaa acagttgatt  
>.....'LacZ.....>

3541 gaactgcctg aactaccgca gccggagagc gccgggcaac tctggctcac agtacgcgta  
>.....'LacZ.....>

3601 gtgcaaccga acgcgaccgc atggtcagaa gccgggcaca tcagcgctg gcagcagtgg  
>.....'LacZ.....>

3661 cgtctggcgg aaaacctcag tgtgacgctc cccgcgcgt cccacgccat cccgcctctg  
>.....'LacZ.....>

3721 accaccagcg aaatggattt ttgcatcgag ctgggtaata agcgttgga atttaaccgc  
>.....'LacZ.....>

3781 cagtcaggct ttctttcaca gatgtggatt ggcgataaaa aacaactgct gacgccgctg  
>.....'LacZ.....>

3841 cgcgatcagt tcaccgctgc accgctggat aacgacattg gcgtaagtga agcgaccgc  
>.....'LacZ.....>

```

3901 attgacccta acgcctgggt cgaacgctgg aaggcggcgg gccattacca ggccgaagca
>.....'LacZ.....>

3961 gcgttggtgc agtgcacggc agatacactt gctgatgcgg tgctgattac gaccgctcac
>.....'LacZ.....>

4021 gcgtggcagc atcaggggaa aaccttattt atcagccgga aaacctaccg gattgatggg
>.....'LacZ.....>

4081 agtgggtcaaa tggcgattac cgttgatggt gaagtggcga gcgatacacc gcacccggcg
>.....'LacZ.....>

4141 cggattggcc tgaactgcca gctggcgcag gtagcagagc gggtaaactg gctcggatta
>.....'LacZ.....>

4201 ggcccgcaag aaaactatcc cgaccgcctt actgccgcct gttttgaccg ctgggatctg
>.....'LacZ.....>

4261 ccattgtcag acatgtatac cccgtacgtc ttcccagagc aaaacgggtc gcgctgcggg
>.....'LacZ.....>

4321 acgcgcgaat tgaattatgg cccacaccag tggcgcggcg acttcagtt caacatcagc
>.....'LacZ.....>

4381 cgctacagtc aacagcaact gatggaacc agccatcgcc atctgctgca cgcggaagaa
>.....'LacZ.....>

4441 ggacatggc tgaatatcga cggtttccat atggggattg gtggcgacga ctctggagc
>.....'LacZ.....>

4501 ccgtcagtat cggcgaatt ccagctgagc gccggtcgct accattacca gttggtctgg
>.....'LacZ.....>

          PstI
          -----
4561 tgtcaaaaat aactgcagcc cacagaagct gcccgctctc cgttttcctc tcttttcgga
>.....'LacZ.....>
          >>.....DHFR .....>

4621 gggatcaggg agagtgcctc gggtcggaga gagctgacga gggggtgcca gagaccctg
>.....DHFR .....>

4681 tgtcctttat cgaagaaaag ggatgactct tcatgtggca tttcacacag tctcacctcg
>.....DHFR .....>

4741 ccttggtttc tttttgtcaa tcagaacgaa agcgagttgc gggtgacgca gatgtgcgtg
>.....DHFR .....>

4801 tatccactcg tgaatgcgtt atcgttctgt atgccgctag agtgctggac tgttgctgtc
>.....DHFR .....>

```



5821 tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag tggactcttg  
>.....pKS+.....>

5881 ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt ataagggatt  
>.....pKS+.....>

5941 ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacgcgaat  
>.....pKS+.....>

6001 ttttaacaaaa tattaacgct tacaatttag gtggcacttt tcggggaaat gtgcgcggaa  
>.....pKS+.....>

6061 cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac  
>.....pKS+.....>

6121 cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg  
>.....pKS+.....>

6181 tcgcccttat tccctttttt ggggcatttt gccttcctgt ttttgcacac ccagaaacgc  
>.....pKS+.....>

6241 tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg  
>.....pKS+.....>

6301 atctcaacag cggttaagatc cttgagagtt ttcgccccga agaacgtttt ccaatgatga  
>.....pKS+.....>

6361 gcacttttaa agttctgcta tgtggcgcggt tattatcccg tattgacgcc gggcaagagc  
>.....pKS+.....>

6421 aactcggtcg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag  
>.....pKS+.....>

6481 aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgcctgcc ataaccatga  
>.....pKS+.....>

6541 gtgataaacac tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg  
>.....pKS+.....>

6601 cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga  
>.....pKS+.....>

6661 atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt  
>.....pKS+.....>

6721 tgcgcaaact attaaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact  
>.....pKS+.....>

6781 ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggg  
>.....pKS+.....>

6841 ttattgctga taaatctgga gccggtgagc gtgggtctcg cggatcatt gcagcactgg  
>.....pKS+.....>

6901 ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta  
>.....pKS+.....>

6961 tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac  
>.....pKS+.....>

7021 tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta  
>.....pKS+.....>

7081 aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt  
>.....pKS+.....>

7141 tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt  
>.....pKS+.....>

7201 tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt  
>.....pKS+.....>

7261 gtttgccgga tcaagagcta ccaactcttt ttccgaaggc aactggcttc agcagagcgc  
>.....pKS+.....>

7321 agataccaaa tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg  
>.....pKS+.....>

7381 tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct gccagtggcg  
>.....pKS+.....>

7441 ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggc  
>.....pKS+.....>

7501 cgggctgaac ggggggttcg tgcacacagc ccagcttgga gcgaacgacc tacaccgaac  
>.....pKS+.....>

7561 tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg  
>.....pKS+.....>

7621 acaggatatcc ggtaagcggc agggctcgaa caggagagcg cacgaggag cttccagggg  
>.....pKS+.....>

7681 gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat  
>.....pKS+.....>

7741 ttttgtgatg ctgctcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt  
>.....pKS+.....>

7801 tacggttcct gcccttttgc tggccttttg ctcacatgtt ctttcctgcg ttatcccctg  
>.....pKS+.....>



```
7861 attctgtgga taaccgtatt accgcctttg agtgagctga taccgctcgc cgcagccgaa
>.....pKS+.....>

7921 cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga gcgccaata cgcaaaccgc
>.....pKS+.....>

7981 ctctccccgc gcgttgccg attcattaat gcagctggca cgacaggttt cccgactgga
>.....pKS+.....>

8041 aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag gcacccagg
>.....pKS+.....>

8101 ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc
>.....pKS+.....>

8161 acacaggaaa cagctatgac catgattacg ccaagcgcg cattaaccct cactaaaggg
>.....pKS+.....>>
>>.....T3 .....>

                KpnI                      HindIII
                -----                      -----
8221 aacaaaagct gggtaccggg cccccctcg aggtcgacgg tatcgata
>>
```

Figure 2 B

Molecule: S13/lacZ, 8268 bps DNA Circular

## Molecule Features:

Type	Start	End	Name	Description
REGION	1	1294	'promoter	promoterregion
REGION	1295	1492	e	exon 1
REGION	1453	1455	atg	ATG start
GENE	1453	1492	cds	gen
GENE	1495	4578	'LacZ	LacZ gene from E.coli BL21
REGION	4582	5354	DHFR	DHFR 3' untranslated region
GENE	5408	5388 C	T7	T7 primer for sequencing
REGION	5408	8202	pKS+	pKS+ vector
GENE	8202	8222	T3	T3 primer for sequencing

Figure 2 C

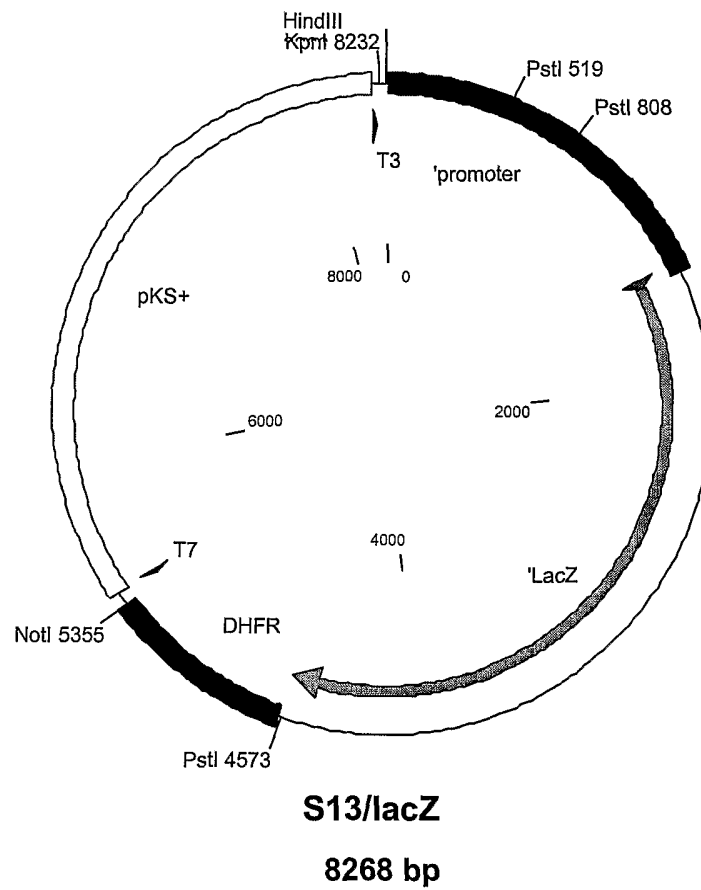


Figure 3 A

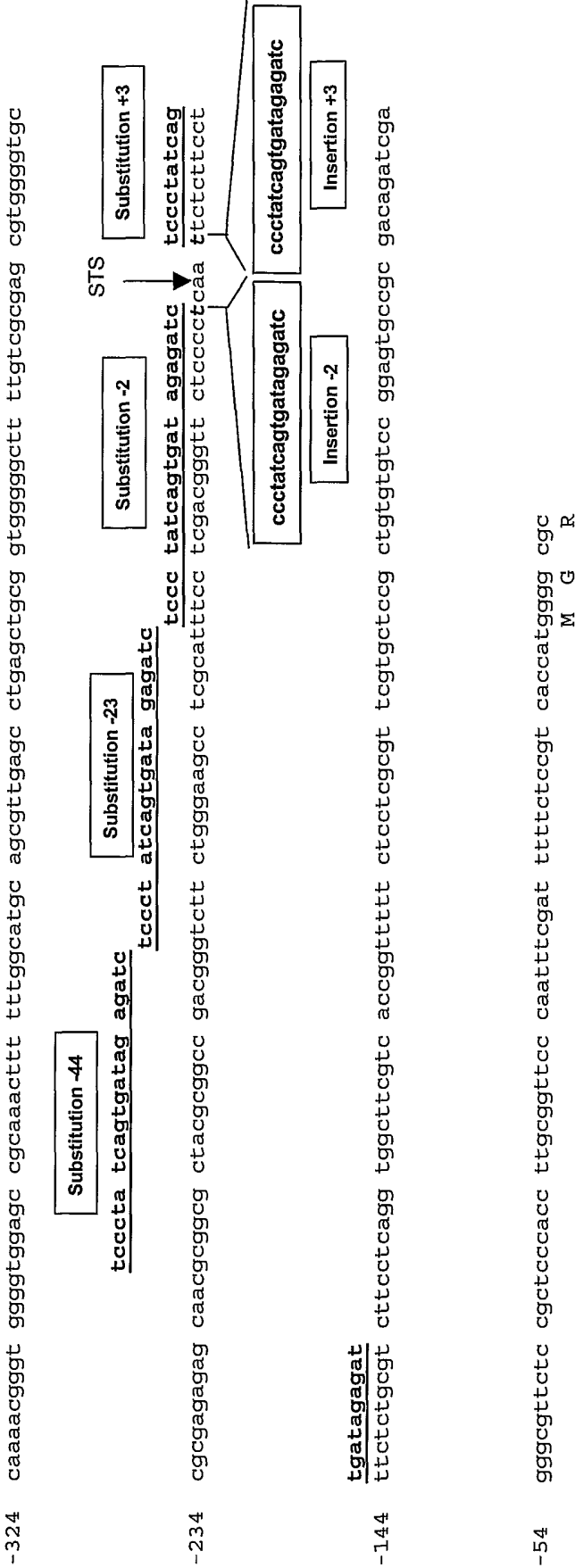


Figure 3 B

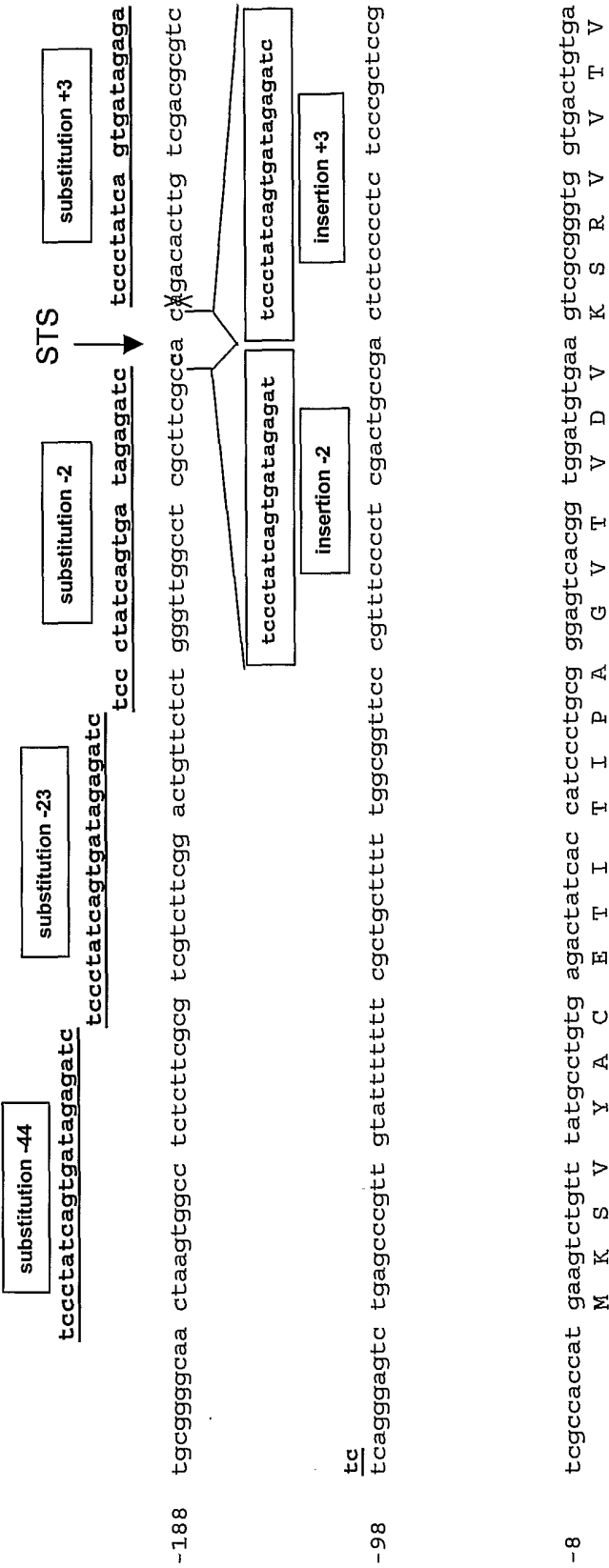


Figure 4

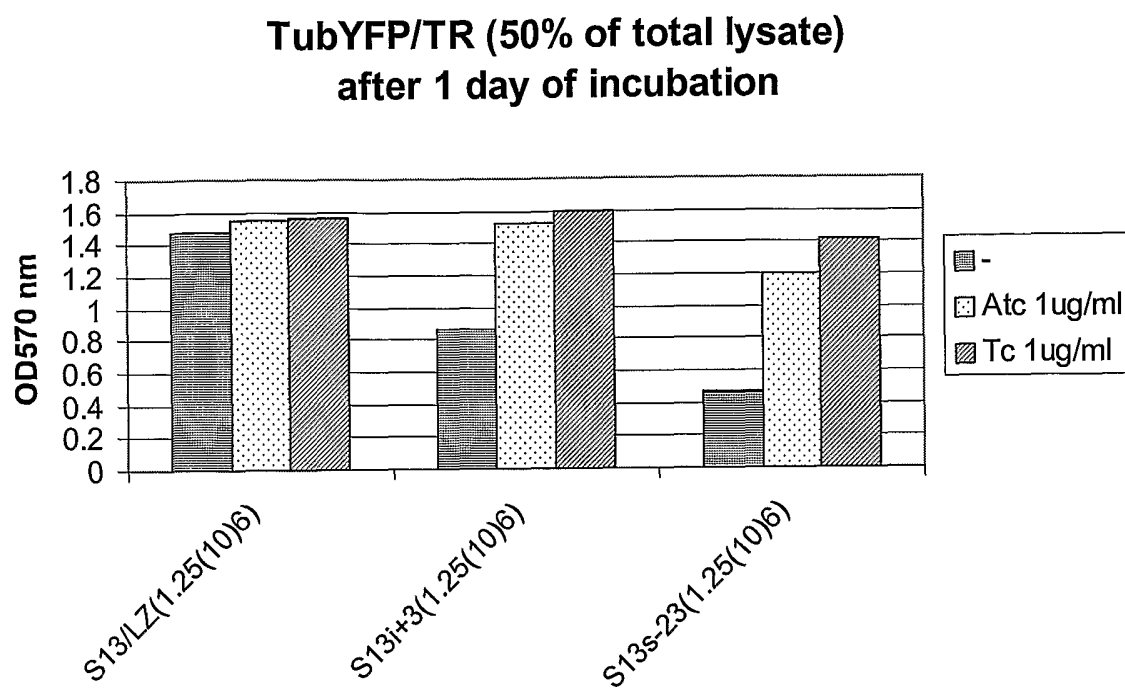
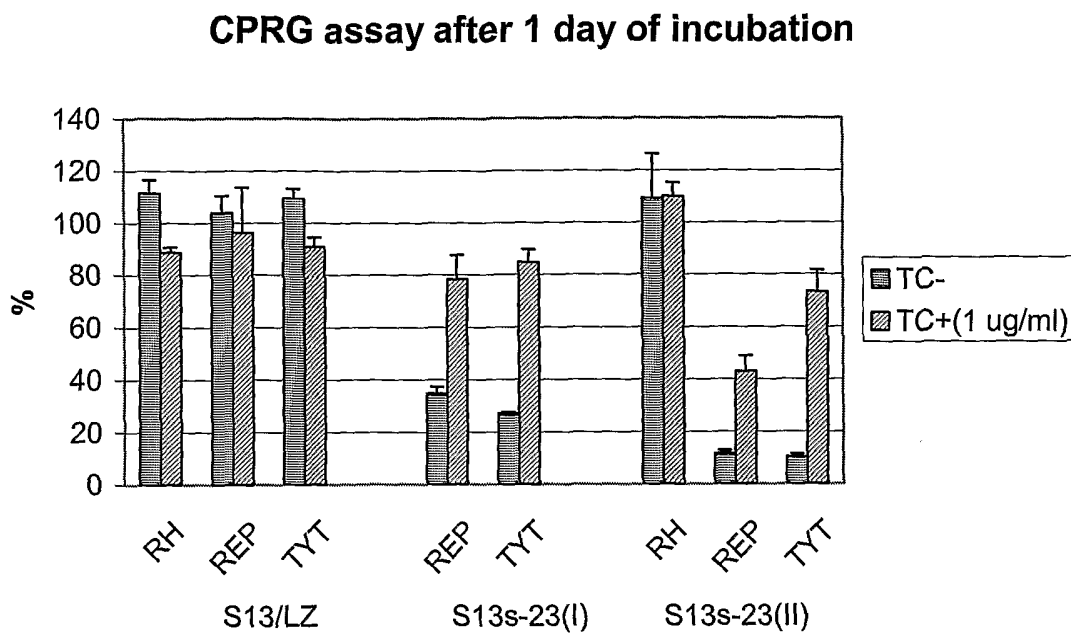
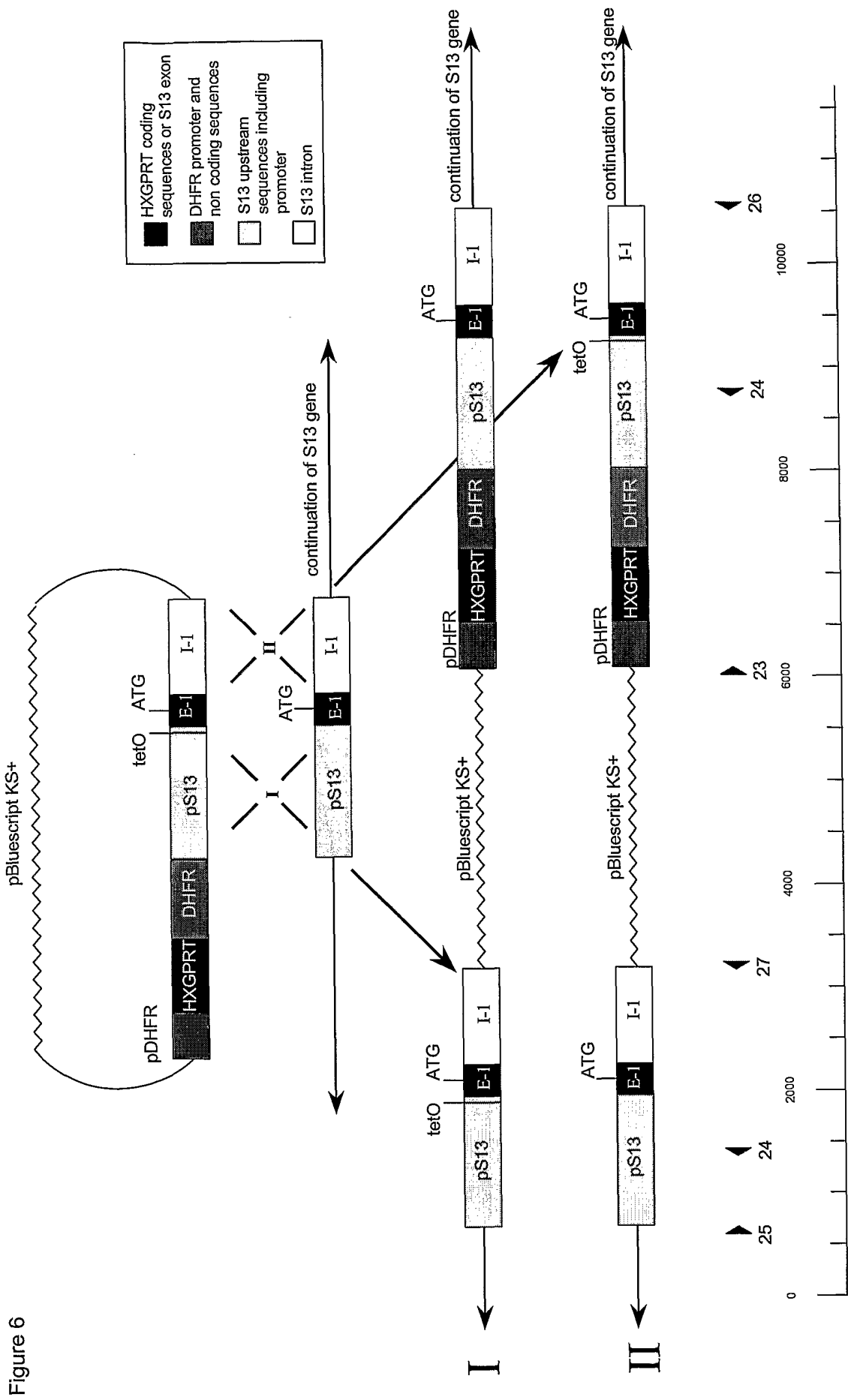
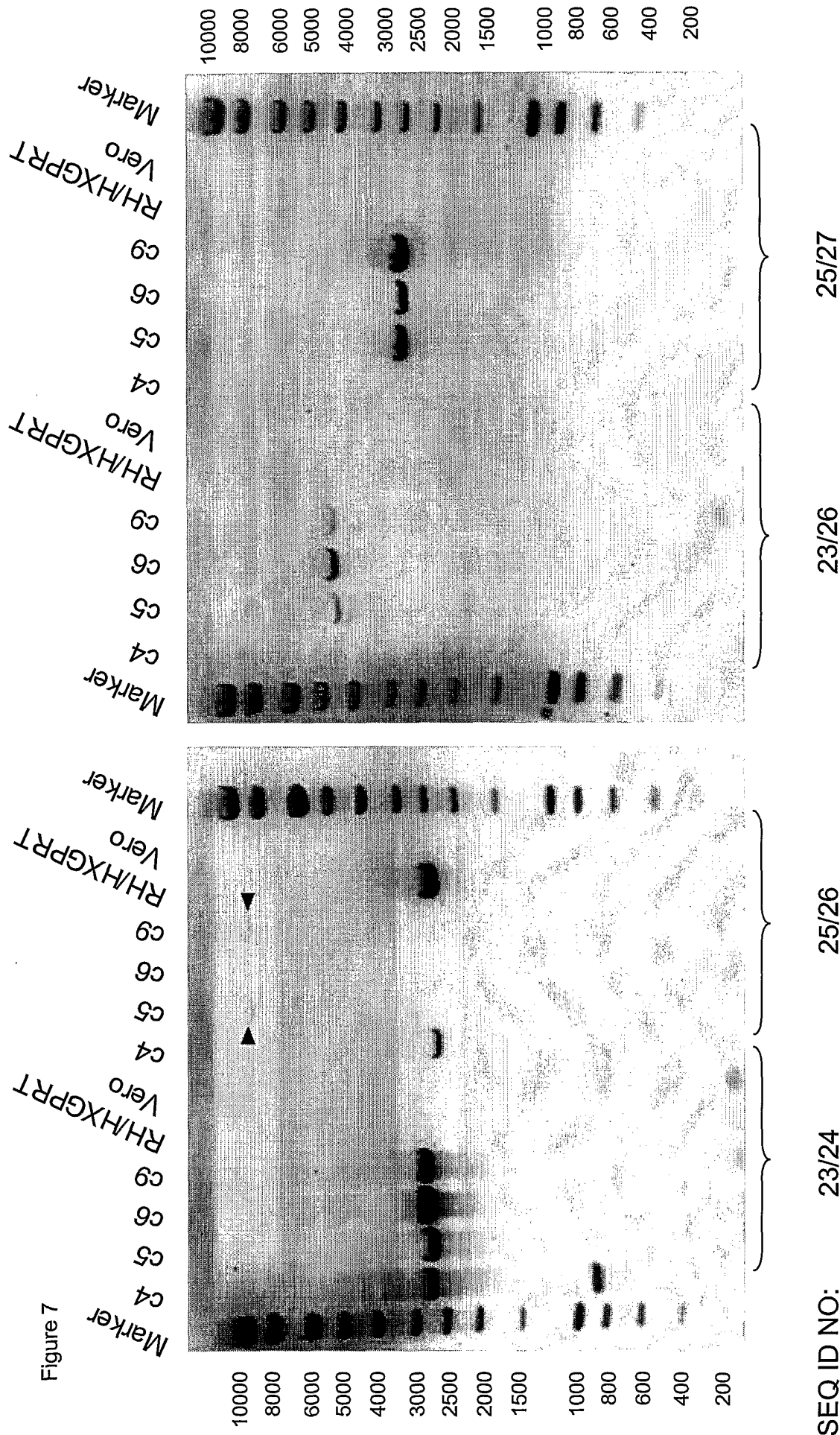


Figure 5









## SEQUENCE LISTING

<110> AKZO Nobel N.V.

<120> live attenuated parasite vaccine

<130> 2002-017-EP

<150> EP 02078953

<151> 2002-09-20

<160> 29

<170> PatentIn version 3.2

<210> 1

<211> 4834

<212> DNA

<213> *Toxoplasma gondii*

<400> 1  
cctagttgtg ttcgcaacag tacaccgtcc tgagtgagtc gagaacatca gagatgagca 60  
cacgcaatag cgggtccgcc aggggtgcatt tgtccacatc gggatgcaca gaggggcacg 120  
agtcgcacaa aagcagatac tagagacaag gagagagtgcc ggcctaacca gaattcgact 180  
cagtttcttg acccattcgt tagggtcggg ctcagcctcc ttcaggattt cgtcaagac 240  
atctttgcta gcttcccgtc gcagacatga aaggcagtggt cacgcataaa gagccgattg 300  
aaacgcagtc acagagatac gaagaaatca aagcccgtgg aaagcgaacg gctgggatgt 360  
agctgagaaa gcaaattcac tggcgggtgca aagagccaat gaaatcaggg tcgcgtagag 420  
gaactataaa acgtgaaaaa cgtgccttcc gagtctcgca aaggtgcgca tcgatccac 480  
atttgagaga aggttgcgag gcagtaataa gggcagggga gaggataaaa tccgatagac 540  
ccagttcttg gtctcccaga acggggacag gaccggacgc ctgcaagggt ggatcacaac 600  
tccagaggca aagccgccac ggaggaacgg aatccatgac cgagtggaat tataacgaag 660  
aggtgtttgt cgtcggaatg gtgccaagac acaaaaaaag aaatgtttag acgctcgact 720  
gtgcactagc ggggggaggg gtgcaaaagg gacgagtggt ctcagtgggt cggaggtaac 780  
tgaaaaaacg gtgcaaaata tggagcctta cgtggagccg cagggggcag aacagatgtc 840  
tcagaagaaa gtccgagaga acagaagaaa aacgagaaaa gtgatgggag actcatgcag 900  
agtggcgcca cgagtctgtc tctcagacga gcttaccagt gctgggcgga ggtaaaggaa 960  
agaagtcaag acgcggacct tgaggggggt ggacagcatg atgaatcgct gatgtatgta 1020  
ctttagaagc gcaggagtta agagtcgagc ggcatggcag gacgaccagt tgtcctttat 1080  
gcttcgcaga taggcaatat atctgctgct gagggcctca tttctggaga gttgcgttgt 1140  
ccctgtcgtc gcctcatcct ttatctccgt gtttgtctct tccagggcag ccttctgact 1200

accgccaac gggcttcctt cttcggattc catttgagat agccgtagaa gcagaggaag 1260  
agccgtcaga acgcttgccg cggcagaaaa acacttaaag ggcgtcacia gattgatata 1320  
ggcaagagga atggacgtca acaggctgat tcataagtga cgctcccagt aagtggcgga 1380  
cagccatgaa aatgagcggc cgagtttgca gaaacagaga aagaggctctg catcctggcg 1440  
aagagccgcc ggacaccctg cttctctttc acagttcgta ggtgccaaga ccaggaccaa 1500  
attatcgccc ttcttagcaa accttgagcc gagttaccgg agaggttagc cgaaaaagaa 1560  
tcgaaacgaa gacgccatth tttgtctcca ttgcacacgg acggaccgta gcttgtctct 1620  
cagcatatct tacgacgttt tgcggctggt atcgctaaca caccacaaag agaaatggtt 1680  
tatcgaaaaa cttgttagcc ggatggtaaa gagatgcaga aggcagtccg cagtaattcg 1740  
gttttcgtca gttgtggcgt gctggcacac tcacgttttt ccagcgtcac atgctgcctg 1800  
attcacgcag aaactgcatg tgcgctgcgt gtctcgctg cctcaggatg cccttgtcgg 1860  
ccgatagtga ggaaggaaaa acggctccag caaaatggtt gttctattcg gcgagtgcg 1920  
gtattccttc cacaaggctc agacaccgtc gagtgttttt cttccggact gaaccccgga 1980  
aaagtcactt tgcaccgtag attccacgtg ctccagcgcg gctgtcaatt ttcgacactg 2040  
cgcgaaacggc ttgccaacaa gaccaggctc gcgcgcccgg cttttcacat tcccgcggc 2100  
ttatatacgg aaggctttgc caggcgtatt ctggccgcgt ggggtcgaaa gaaagtcgaa 2160  
aaagagcatg cttgtcaagt gcatgcggcc atgtagggtt ctaggacccc tgttaaattt 2220  
ccagggtgcg gggcaactaa gtggcctctc ttgcgctcgt cttcggactg ttctctgggt 2280  
tggcctcgct tcgccacaga cacttgtcga cgcgtctcag ggagtctgag cccgttgtat 2340  
ttttttcgct gtcttttttg cggttccgt ttccctcga ctgccgactc tcccctctcc 2400  
cgctccgtcg ccaccatgaa gtctgtttat gcctgtgaga ctatcaccat ccctgcggga 2460  
ggtaagtthc tcgacctacg agagggtgaa ctgcggagaa gacgaatgaa acattgcccc 2520  
gcttgatctt tgaggagag ttgccagatt ctgcggctcc acagccctcg ttttttttcc 2580  
tcccgcattg gttagatgtg tcccgacccc gaggggaagcg atcgacacgc tgggaaggaa 2640  
cgcccgatga gcggaaggt tttggaattc aggccccgat gcgcaaagtg gcaagtgtct 2700  
tggacccac tgaggagcc gaacagcagc attttacaga tcttcgccac tgaggagggg 2760  
ggcggatctg ggaggtgaag aggcgcggaa ccgtgttcca cttggctttt ctccgcattc 2820  
gctgtgtctg ctctgcgtca aaaatccgca tgctttgttg tcattcaaag aggtcatctg 2880  
ggcgccttgt tctttgttct gccgcattca caacagtctg acccgccaga gaatacggtc 2940

tggtctgtcc	ggtgactggc	gatggggaaa	tgggggaaac	tgtgtcgtca	gcgagtgaag	3000
gcgtttttta	gtggaatttc	tacattgtgc	aagcacacag	aaggtgtccc	gtgctaatat	3060
ctggaacagt	agattatgat	taggtagtgg	aacagggaga	gcgtctgttg	tacatcactg	3120
tctgcactcg	tttgtactac	aacgaagtgg	ttgatgcgct	gacttggttg	tcgattgcat	3180
agacatagcg	tggaaaagta	gaagacaggg	ttgtatgcga	ggctctgtgt	gcacctgttt	3240
catgtggaca	agaccaccgg	gcatatgctg	gctgttgctt	caacacgctg	ccgaaacatg	3300
tcacggcggt	gcgggggaaa	ggagtcgttg	tagaaaccat	agagagagtt	gaggtagctc	3360
ttgatgtctc	gcaaaaatgg	gactggcacc	tgttgtctgt	gtcttcgatt	aacacgagcg	3420
ccgccactgc	gtttgatgct	cgctaactgg	gcagcgtcgt	gtacgtacag	ctcgaatagc	3480
gtaattgtgt	gtttttgtac	tctttctgtt	tgagtttcat	caaagagggc	cagccacaaa	3540
atgggagcag	ggggatattg	gagggcatat	gataagtgcc	gcctgtgtgc	atacttctag	3600
aatagacagg	aattcgagag	cgaagctgtc	tgaacagaga	tctgcagggt	ctggtttgac	3660
tgtgtaggca	ggtttctgta	gcgacgggag	tcgcaatgca	gagtgcgct	tggtattggt	3720
tgtttcaaga	tgtttgcata	cctctgagag	caatcgcttt	ttgtcctgtt	ttgcgtgtct	3780
ctcggctgtg	tgccttctga	aagaaaatgt	tgcattccgt	tgcggttttc	tgctgcagtc	3840
acgggtggatg	tgaagtcgcg	ggtggtgact	gtgaagggca	agtaggcgaa	atcacgcgtg	3900
cattccgcca	cctccctgtc	gacatccaga	agaccaagtc	tggaaaccga	ctgaaggctc	3960
agatgtggta	tggaaacctgc	acagacctca	gctgcatccg	cacgctgtgc	tctcacatca	4020
agaacatggt	cactgggtgtg	atgaagaagt	tccagtacaa	gatgcgcttc	gtgtatgcac	4080
attttcccat	caacgtgaac	atcagcggca	acggaactgt	cgtcgaaatc	cgcaacttct	4140
tgggcgagga	gcgtgtgctg	atcgtcaaga	tgcttccggg	agttaagtgc	gagaaggcca	4200
caaacgtcaa	ggatgaaatc	gcgctcactg	gaactgacgt	cgagctcgtc	tctcgatcag	4260
gtagaggctc	cgaggaactg	aaaagggggc	tggtgtgccg	gtatgcgcgc	atctaatatg	4320
agttttggag	gtgctggaggc	agcaaggaag	cgtatagatg	tgggcattta	tgaatgtgca	4380
tctatgttgg	tgtaattctg	tgtatgcctg	attgcgacgt	gcccacaacc	acctccaggt	4440
tggagaaga	gagaaactga	taacgggtgga	ccccgagagc	gggattaccg	ggaactctcg	4500
gacggtcgtt	gtataactcat	ctcacgtggg	cgagggggag	gtggtttgtc	cttcgatgtt	4560
gccacagatt	tggagggtgag	gtgtcttcat	atcctgcatg	tgtgtctgca	ccagcagata	4620
tggttaactgc	caggtgagac	acgttgtcga	gccacaggta	tttttgtgta	tctgcattgc	4680
attaacatgg	tttgtattct	tctgttctgt	atgcttctct	tcttcagcgg	ctctcatcca	4740

tcaatcgact ttggtcagga gaaaggatat ccgtaaattt ttggatggca tctacgtgtc 4800  
agagaccagc actgtcgaac aggacgcgta aatg 4834

<210> 2  
<211> 4338  
<212> DNA  
<213> *Toxoplasma gondii*

<220>  
<221> misc\_feature  
<222> (565)..(566)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (625)..(626)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (639)..(640)  
<223> n is a, c, g, or t

<400> 2  
atcgcatgac ctgatcacgc acggaaagaa acgaatagtc gccatttgaa gtgagatccg 60  
tgtgaacagg tcagatcacc gaatcggacg atatatgcac tgaaggcagc ggagccagct 120  
gtaaacaaag aactggacag cagctaccgt agctgtagac ggacgcgact tcgagagcgt 180  
ccacgtcaaa cctcaccgat ctgcacctca taaagaggca tgtgggctgg gagatacagg 240  
ggtgaagaag gagagacaat ttgcgtaagg aggcgaagct ttcgatttcc aggtgcgatt 300  
ggagtcgccg ccacaggaga cgcgaactcc tcaaaaacgg acacggagaa gccctgtgca 360  
gagacaacgg aaagaatgtc ctgacgagag agttgcaaaa gaatgttgaa caattaaagc 420  
aatgatgcag actcgaagat ctaacgcctc gcaggtctca acggttgctg tgatcgccca 480  
tttacacagt ccttaagttg agtgcattgag aggccttgca gctcaaggca acgctgtaaa 540  
cagcagtgtt atgaatcggg tgccnntatt gaggcgtctg cgtctggctg gtccatcaag 600  
ccaaaagacg cttgtaaaca ggatnntcca ttogaatggn gacagacagt ttggcaactg 660  
tcatcacacg tgacgttaaa aggcaccgtt aagcgcattg caaggaaagg tcacccgcga 720  
tttacacaca ccaggtgccg tagctgtcga tgaatgogaa ttccagagtt tttctctccg 780  
acactacata agctgtaaat gctcattctg tcattcgctg accgtgttta ctacggggaa 840  
tcgagaaacg gaatatcaag aacacaggct gtcaaaagac accgcgaaac ctgcttgccg 900  
aatctaacgg ttgcctctgg ccatttatgt gtttctcgcc tgtgccttgt tcgctgcaga 960

cacagcctga gtccgcagcg aggtaaatac gaagaaaaac ctgacgagct ctgtcagatc 1020  
tgtacaagcg acagaagcgg attgacagag gagagtgcgc gacggtgacg agagtgagag 1080  
tcgactacga agttagagga caccaggggtg gcgaatgtgc caatacgcag cttgaaaggg 1140  
tcgagatcga caatcgaaac tcacttcact cgttaaacaac tcgagcggtt gctgcaggtt 1200  
ttgtttgggg caccgccct ttgccttccc acccatcggg ttcagccccg agtactccac 1260  
cagcaaaaca gcatcgaggc cgtatgcctc gaagaagtct ccaacctgca aaagaaaggc 1320  
accacgtcag agcaaggaaa tcaagactca accaggtgta cagacaccgc ataccgtcgc 1380  
caggaacccc ggtctaggac aactttgcta gtgctgctca aaaggtggaa cggagaaggc 1440  
gagacagcag actggcgggtg ccagttcaaa tcaccactgc ctgaagcgcg gggaccgaga 1500  
cagtttgca tgttggaatt cctgtgacgc acacactttg gaacttgcct gaataatcag 1560  
aactttgtcg ggccgaagtc gtttttgttc tcgtacgaag acgaggagag gaggcataatg 1620  
cagggagtca ggaaccatcg acggatggtc gaaggaaaga aaagagagct gccgccggga 1680  
agcgggacgg gaaagaagcg gcagccttgc caagaacgtc ggggtgtactg ccaccgaggg 1740  
aggcgggcag ctttctgcag acgaacgcag aacggagcag ttttcttccc ctcttcgact 1800  
cggcttcggt ccagcaggt tgtcgcgct cggcgcttcc ggacgcttct gcgtgtggaa 1860  
gaagcggccg gacaggcggg atgcgtttga agggaaatga gtctgcctgt ctgaaatcgc 1920  
gtgaggcaac tgaatggtcg gacgtgcgag gcgtcggtcc ggatggacaa caaaagcgac 1980  
gcgtcgagaa gaggcaagaa agggcaagct ggcgcgacag ccacgcaact ggcctgctgc 2040  
gtcccttgca gctgccgaaa agaagtccag cagctgatgc tgccaagag tggcccaccc 2100  
tccgcagctg ctgtcgcggt tcgcgacctt ctgtaaaaaa ctcttgagcg tgcggccgtg 2160  
catcttggtc gcggaaagca aggcatgat cgaccagtac gcttgcgga cggagcaggc 2220  
taaaccctgc cggcaaccgc gctgtcgcg aaacatttcc atcagcagtc tcgcgttgac 2280  
ctccactgag atgcacacaa gtgaaaagag aatcgtctgc agtttccagc gatagcgcgc 2340  
gaaagaagcg gagccgcgga gaggcgcccc cggacccgaa gcgggcccga gagcgcgaga 2400  
ggcaatggag caaaaaaatt cctgcctaga ggagaatcgc cattcctggt cacggtcact 2460  
gcatacagaa tcccctcca tcgtgggtct gatgaaaaag agaaaaggac accattgtgg 2520  
tgtggagccg cggtgccgtc tcggctttcc cgttgggtcaa tgcagaagcg cgtccccagt 2580  
gactagagcc gaaccaccgg cgactccaat aaaggggcct tctactccat tcaggggtgt 2640  
cgcattgtaa aactgagttc tctaatacat ggcacacctc ggaaaaacac tactcacagt 2700  
cgtggaatat cttaaatagt cgagcgctcc agagtaccaa gctgcaatgc ccagctgcct 2760

ctcggacaac accggtggac taggcagcgc atcaaggaga cccccccag aggcggcctt 2820  
actccagaac atgcgaagga tgtcactgtt gatgcaaaat cagggtagtg accatcaagt 2880  
gaacagtttc tagagtatcc acgtggctcc aaaaagtaaa gtatggcccc acatcaaacc 2940  
gagacyagcc tgtcttcgtt agaaacgctc gtgccgatga cgtcagtgca tgcacgctgc 3000  
agacacgacc caaagtagct gatacgggtga ggaatgagca tgctgccgat tcaaaatcgt 3060  
cagtcgcggc tacagcagtg ttttatgtac atcctcgttt tctttttatt caagaaaccc 3120  
aggtaacatg tgcttgaagg ggcgacttgg gtagaatggc cctcgaacgt tggtgccgag 3180  
gggttcggtg cagccttcca tcttgggctg taatgtgtct tgcgtatttt ctctgcgggc 3240  
ctgcagtcgg aactgccttc atagtccctc aacgcaacct gcatagcttt accaggaatt 3300  
tccgtgtcct tcccttgggt gggaagcgat ggtttgcagg agatgcagct gtgcaacact 3360  
gtaaaggctg gtttcagcaa acgggactct cccctcggtc actcccaggg agtgtcggcc 3420  
aggtgacggc ctggtctgcg gtcacagaag gatgtccttt cgaagatgca gtccatgaaa 3480  
cttcgcttcc tagtgactgg gtggagagga gtgaaccac agctctaggg acggatgcga 3540  
cactctcaag cattccatcc gagctacctg agcctggaac gccgcggaac agcttctcat 3600  
ggggcacccg gaaaaggtag tgacgagaga atgctttctc agaattgatc cttcgccttc 3660  
tgtgagaaac actgtcacga ggactgcaat cttttcaggt gtatcttttc tgtacagctg 3720  
tgcagcgttc acgcttttga ccggagagta cccgtgatgt gacttgccgg ttgggtagag 3780  
cctagtccag ctgtctgtca tcttctgaac gggagttagg aagccgtaac aactaatatg 3840  
tgcacgcgg tttctccggt gtgcgcgtgg gcaggcttta tgcaagggtt cggcttcccc 3900  
ctggttctgg aaagctgacg atcaacaaca gagacgcagc cgactatctt caagacaatc 3960  
catggtggat tcataattgc atcgcgccac tcatggaact gcagctggag aatgaatttg 4020  
atatcattgc ggaggtaagc cacgttagct agcttcgggt gacagtcgat atgtttcgca 4080  
atgatttctt acagcgtctt tctacgttt ttgcttgctt tcgacatcag gccacgggg 4140  
gaggcctcgg cgggcagctc ggagcaatca tgcttgctgt cgctcgggag attgtgcgac 4200  
agcgaccgga actgcgaccc cctcttcggc gagcaggttt tctgaccgta gacgcaagga 4260  
aggtcgagag gaagaaattt ggtcttcgga aggcgagaaa gaaggaacag tacagcaaac 4320  
ggtaggggtgc gtaggtca 4338

&lt;210&gt; 3

&lt;211&gt; 3639

&lt;212&gt; DNA

<213> *Toxoplasma gondii*

<400> 3  
cctcgcagag attgtcagtg catgacacaa ccgcgaaaag ccggcagccg cggtaatacgc 60  
gggacgagga aaacgactga gcgtcacaaac agaagcagcc gagtaaaccg cgaaggaaat 120  
ggaaaggacc caagtaaaat ttcttgaaga atttcagcgc aacaactctg cgggttcttg 180  
cgaatagagg aatttcactt cctcatcgtc tgatttatgc tttcatcatc tgccgctcaa 240  
cagccgaata aacggttctc ggctcgcttcc ttaaactcta cttcagtagt tgaaactctt 300  
ttgcttcacg agccttcgtc tcagccctca ccgtcctgag ttctgtcttt gttgaggaaa 360  
gctcccgcgtg aaaaaacagg actttgtttg cagattttca tgtgtactgg aaagtgagat 420  
gtgacttggg gaagtccgct ttaaaatttc cattgttttc tcaaatgaa aagtctaaaa 480  
aatcgaagtg cgtgccccgc gaggaattcc cctctgcaga tttgttttgc atttatatgt 540  
cgtttttacg gagaaaagtc ccaagctgct gctccttctc taactagatg ttgaacgcta 600  
gcacatatgc accagatgct tctgaagtat acctaaacgc accttgggaa caactgtgct 660  
cccattcata aaactcatac aagtcaccaa gcatgccata ccgctgagac ataacaacgg 720  
aagctagact actccccct gttattgcac actatcgaaa aggattccta ggtttctatc 780  
ctctgccttt tcctggggca cactgcagag aaactaccgt gcgcgctacc tcccgcgtg 840  
cgaggcgata gcaaaacgct tttgaaggaa aaagtcgaga aatcgacgac tgcgtctctt 900  
gaatccgaga gagggatcca acccaccgag ttctctgcat gtgcagcatc tgcaagaacg 960  
tgataatgca tgaactcgat catcgctta tctgtgtgca tgcattttcg aaaaagaaag 1020  
gcgttttctg cgcggagact cgcgcggagg caagacgaga ctttctcctc ttccaaactg 1080  
cgagccacgg gggcgcatgc aatttgaaca tcacgcaaaa tcccaaacg ggtggggtg 1140  
agccgcaaac ttttttgga tgcagcgttg agcctgagct gcggtggggg cttttgtcgc 1200  
gagcgtgggg tgccgcgaga gagcaacgcg gcgctacgcg gccgacgggt cttctgggaa 1260  
gcctcgcatc tcctcgacgg gttctcccct caattctctt cctttctctg cgtcttcctc 1320  
aggtggcttc gtcaccggtt tttctcctcg cgttcgtgct ccgctgtgtg tccggagtgc 1380  
cgcgacagat cgagggcggt ctccgctccc accttgccgt tcccaatttc gatttttctc 1440  
cgtcaccatg gggcgcatgt acggtcctgg aaagggcatg tctgccagcg ctctcccctg 1500  
gcgcagaaaag ccccgacat ggctgaaaat caagccgtcg gacgtcgaag agcacattgc 1560  
caagcttgca aagaagggcc agacccctc ccaggtactt tcggcgggaa gaaggagaa 1620  
aaacgacgga gttgccgcgg ctgcggtctg gggaacaacg gggaagtgc caggaataat 1680  
acgcgttctc caggtcatcg ggggaaaacg ctgcgcaatc ccagcccctc gactctccgc 1740



agttgttttt cgcagtcttt tcgccctcgg accgttcagc ggagatgggg acgagaatcc 1800  
tctccctccc tctgctgagt tttcccgct ctctcgtgtc tcaaaaaagg ctgcagaaat 1860  
gctcgcttgc ctcacagccg gacctctctg ttgaccaaag cggccagtcg ggatttcttg 1920  
cgcggggatg ggctgaagca acgaaacggg atggacgttt gtcggtttcc ccgctgtcgt 1980  
gtgacttgct cgaggaacca aagaacggga acgagcggga ggaagccggg ggaaaatttg 2040  
cgttcctccc cgcaaacct atccgcaaaa atgcctcgtt tcgcgaactg tggacggggg 2100  
ggaacaacgc gttgtgttgt ttgtgcatac ctgactgaca cggctctgcg cgtgggtggc 2160  
tgggatccga gcggtccgaa gacagtctct cggaaattcg cgagcggacc tactgttggt 2220  
cactgaacgg ttcgcttctc ctgtgaaatc aacatggttt cttgtgcagt ttaccgaaag 2280  
tgaggacgac atgttttttg tgaccgtggc gggccggtt cgcgtcggcg gcaaaccgga 2340  
cgatccaatg cggcgaaacc gggggagagt cattgcacac gatgaagtcg cgaacgacca 2400  
aggcaatttt ttcggaggtc aaatcacatc tttcgagaag tatgaatgca tgtggaaggt 2460  
cgtctgtttt tccattctcc acgttcttct tgcttttggc tgcttgctgt tctctggcga 2520  
cttacagatc ggcgtcacac tgagagactc cttcggagtg ccgcagggtta aatccgtcac 2580  
tggaacaag atcctccgta tcctcaagct ccaaggtagg ttacgcgtt aagggaaaca 2640  
ctagtttcaa tcttctcgag aacactggag gggcggagat cggggcgag ctccttccag 2700  
gctcaagaag gtctgggagt gaaaaacgaa cgcaaatgca tggacgttgt atatatgtat 2760  
gcctggatac ggtgtgaggg tagcgccttt ggcaggagca agtgtgaagt ttgcgtctgt 2820  
ttggagaagg aatgacgccg cgtcgtcgcg agggcggttct ctgccttccc ggttctctgt 2880  
tctttgagaa agaacgtttt tcgcgtttct ccgcagttag aggttccctt cgaagaggca 2940  
cctagatcag tcgactcgtt cttgaggagg ctggccttcg tcagtgtgtc tgctgcttct 3000  
ctcactgcaa cactgtctcc cttgaagaga tttagcgcag atgctgattt tctggcgttc 3060  
agctctctgc cgtcgccctc tocaacatgt caagaagcac gttgcttgct tccctctttg 3120  
tccagcaaag tggagttttt gtatgcgtgc aatctatgca atcgagagct tgctgaagcg 3180  
acgttgctct cctctctccc aagtgtatgc tctccgcgtt tcttcgtctg gttaaaaaaga 3240  
aacggcgctg ctgtcctttc cttcgtggcg aactcgggat tgtttctcaa atccgatcta 3300  
ctgtcagccg tccaagtgcc tgtctgacct ctttctctga ctcccgcag cacatttaga 3360  
gcgcgtggaa gcgactgttc aagtctctct ctcatctgtt tctctagggt ctgaagagcg 3420  
ccaagtgcgt tttcgaggtc ctccagacct ggcgccacca gtgttctccc gactgttctt 3480  
ttttttcagg tcttgcccc gagctgcctg aggacttgta ctacttgatc aagaaagccg 3540

tgagcgtgcg aaagcacttg gagagaaaca ggaaggacaa ggacgccaag ttccgtctga 3600  
ttcttggtga gtctcgaatt caccgtcttg ctcgctaca 3639

<210> 4  
<211> 2748  
<212> DNA  
<213> *Toxoplasma gondii*

<400> 4  
ctgccgcttc cttaacctct ggaaggggtt gaagcttttt cgactgaaa cgcgagagac 60  
acgaatgagc tgaccacttt tcctgcattc gctggccctg taccggccgc attctctact 120  
tcgtacacct tcactgtact cacacccgaa aacttcagaa gtcgggcttt gctgcaggcg 180  
actcagggca gaggagaagg tgaacatgcc ttccccaatt ttgccgcag tttgcctcgg 240  
gttgctctct tccccaacct agatgaaatg aaaccactcg atcagccatg tatttcccg 300  
aaaaggttgg ctaccaagcc cacatttggc aagcactctg gaagatgcgg cgccacggaa 360  
gcaacctcgc accgcaccgc tgtcgcgggg tctcgacgtc gtccgcgtta acatgatgtc 420  
ccgccagagc ttccactgtt ctgacgagta catggatgca ataagcaagt cgtctccact 480  
cgtacacgaa tgtccagcac cagcaagctg catcacttca atcgcttggc agcccaacgc 540  
ttcttttgtc tgctttctgt ttttgctgcg agactgaccg agcacgaagc ttacacgaaa 600  
cgcgaaagctc taatcgcta catccactc cttcttcaga aatggcaggg aagctacact 660  
cttggttaacg tctctaggtt taaaggggtt gtgcaccgag ccgatcacgt tacggacacg 720  
ggctcggggc tcactacgtt acaaggaata aacaactaac gcacaatcct ggttacattc 780  
gggcccacag catataacct ctccggaggc ttgtattcca gctatcgaaa aaaaagcaat 840  
tcgatgtaat tccctccaat agccccgagc gyatgtcatc tacaagtggc agccttggtc 900  
gggaccctc ttgggtatgc ccggacagat gcgccagtga gatctttaac ctccgcgtaa 960  
agataggtgt ctgctgtagc ggtgcgcttt tttgtgttgc atgcatgcat caagctggcc 1020  
gggacaaggc gggtgccccg atggatggat ggcaaaacca ctgtgctgca ggcagcagcc 1080  
ctccctcggc ggctcctctt gtggggcacg gcgcattgcc cagcacaacg tagcggcctt 1140  
gtccagcatg gacgaaaagc agtgggggag actcccagag gaaagcgttg ccatgcaaag 1200  
gggaaacagg ggatttttgt caccgacgagc tgtttgtcca cttttgtgag gtggtaattt 1260  
gacacagttc tcatccctgt tttgtccaag atggcgctga ccggcaacgt aaggggattc 1320  
tccgcactgc attgctctct tgaggagaaa aaggccgccc tgaaaacatg ctgttctcca 1380  
ccagttggca gattgaagca gtctacggga agatgcacgg tttaatcatt gttgaatttc 1440  
gtctgcagtc tgactcttcc gtattggagc acgcgatgtc tcgttgctg tgaaatcgta 1500

```

ctgtctggat ctctttgagt gaagaacacg cggcaggccg cagttttttt gcagggcctt 1560
ggtaccaaatt gcttggtttac atatttgctt tgtcgagttc ttttgcattg ttttagatat 1620
gcgtggagac tgttaaataca acaaccgctc ggagatattg tgcgcggccc agcaatgctc 1680
tggttccact cccgtcgtga gcagggaacg catgggtgggc ttttgtggct tctgtgtgta 1740
tgccgtctgc agacttgcaa aaagagaaag ttcgtcaagg atgggtgtctt ccaggcggag 1800
ctcaatgagt tcctctctg cactctgtcc gaggatgggt actcgggagt tgaagtccgt 1860
gtgactccca tccgcacaga gatcatcatc cgcgccacca ggactagga agtgctcggc 1920
gacaagggaa ggcgtatccg cgaattgacg tcggtcgttc agaagcgatt cggcttcgcg 1980
cccgactcgg ttgagctctt cgcgagcgt gtggagaacc gtggtctgtg cgccatggca 2040
caggcagaat cgttgcggtg caagcttcta aaggggcttg cagtcagacg cgctgctat 2100
ggtgtcctcc gccacatcat ggagtcggga gctaaagggt agtgctgaca aagtgccatg 2160
tattgtatga ggtaacttga atttagagtg tgaacaaaaa gcattagtgc actgtcacac 2220
gtatcttcgc cggacttttt tcttttcagg ttgcgaggtc gtcgtgtccg gtaaacttcg 2280
cgctcagcgt gccaaagagca tgaagttcaa ggatgggttac ctgatctcta ctggagagcc 2340
ctcgaagatg ttcgtcgacc aagcaatccg ctccgtgcaa ctcgacaag taagtttcaa 2400
attattaagc ctcagttacg tagtaaaggg caatttgtgt aggagctagt atgtacagag 2460
gcagtgtatg tgtgtttttt ttgcagggtg ttcttggtgt tagagtcaag atcatgctgc 2520
cgcatgaccc ggagggcaaa cgtggccccc cgaaccgcgt gccggatact attatcgtga 2580
tggtatccaa gccagagatc cccgttggtc agcctgagga gatggacgag ggagtgtcgc 2640
gtccaatgta atgagtgatt cgtgcgtgac tgttgattta tgggaggagg gtgtccacat 2700
gtgcgtgacc gtggagcagc cgcttaacga aattcgcatg ctccttcg 2748

```

```

<210> 5
<211> 31
<212> DNA
<213> Artificial

```

```

<220>
<223> primer: SAG3-FW

```

```

<400> 5
cgataagctt cgaatctctg aacggatgtg t

```

31

```

<210> 6
<211> 33
<212> DNA
<213> Artificial

```

<220>

<223> primer: TUB5-RV

<400> 6

cgagatctgg gaattcaaga aaaaatgcc aacg

33

<210> 7

<211> 30

<212> DNA

<213> Artificial

<220>

<223> primer: TETAVR5-FW

<400> 7

cgatcctagg atgtctagat tagataaaag

30

<210> 8

<211> 33

<212> DNA

<213> Artificial

<220>

<223> primer: TETPST3-RV

<400> 8

cgtctgcagt taagaccac tttcacattt aag

33

<210> 9

<211> 21

<212> DNA

<213> Artificial

<220>

<223> primer: T3

<400> 9

attaaccctc actaaaggga a

21

<210> 10

<211> 31

<212> DNA

<213> Artificial

<220>

<223> primer: SAG1/1634-RV

<400> 10

cgataagctt tcgggggggc aagaattgtg t

31

<210> 11

<211> 27

<212> DNA

<213> Artificial

<220>  
<223> primer: REV 13A

<400> 11  
gcgccccatg gtgacggaga aaaatcg 27

<210> 12  
<211> 27  
<212> DNA  
<213> Artificial

<220>  
<223> primer: REV 13B (nested primer)

<400> 12  
gggaaccgca aggtgggagc ggagaac 27

<210> 13  
<211> 30  
<212> DNA  
<213> Artificial

<220>  
<223> primer: S13PROMFUS FW

<400> 13  
gcataagctt cctcgagag attgtcagtg 30

<210> 14  
<211> 31  
<212> DNA  
<213> Artificial

<220>  
<223> primer: S13PROMFUS RV

<400> 14  
gcattctaga ggcagacatg ccctttccag g 31

<210> 15  
<211> 33  
<212> DNA  
<213> Artificial

<220>  
<223> primer: LACZ-AVR11 FW

<400> 15  
cgatcctagg atgacctga ttacggattc act 33

<210> 16  
<211> 31  
<212> DNA  
<213> Artificial

<220>

<223> primer: LACZ-PSTI RV

<400> 16

cgatctgcag ttatttttga caccagacca a 31

<210> 17

<211> 50

<212> DNA

<213> Artificial

<220>

<223> primer: S13INSTETO+3FW

<400> 17

ggttctcccc tcaatcccta tcagtgatag agatctctct tcctttctct 50

<210> 18

<211> 50

<212> DNA

<213> Artificial

<220>

<223> primer: S13INSTETO+3RV

<400> 18

agagaaagga agagagatct ctatcactga tagggattga ggggagaacc 50

<210> 19

<211> 51

<212> DNA

<213> Artificial

<220>

<223> primer: S13SUBTETO-23FW

<400> 19

ctacgcggcc gacggtccct atcagtgata gagatcttcc tcgacgggtt c 51

<210> 20

<211> 51

<212> DNA

<213> Artificial

<220>

<223> primer: S13SUBTETO-23RV

<400> 20

gaacccgtcg aggaagatct ctatcactga tagggaccgt cggccgcgta g 51

<210> 21

<211> 32

<212> DNA

<213> Artificial

<220>

<223> primer: S13NOTI-FW

<400> 21

cgatgcggcc gcgtcagtgc atgacacaac cg

32

<210> 22

<211> 32

<212> DNA

<213> Artificial

<220>

<223> primer: S13SACI-RV

<400> 22

gctagagctc ctgtaagtcg ccagagaagc ac

32

<210> 23

<211> 23

<212> DNA

<213> Artificial

<220>

<223> primer: M13-REV

<400> 23

aacagctatg accatgatta cgc

23

<210> 24

<211> 20

<212> DNA

<213> Artificial

<220>

<223> primer: S13CL FW3

<400> 24

cgatagtgtg caataacagg

20

<210> 25

<211> 21

<212> DNA

<213> Artificial

<220>

<223> primer: HRCHECK II 5 S13-FW

<400> 25

gtcgagtcct gtaggttcat c

21

<210> 26

<211> 21

<212> DNA

<213> Artificial

<220>

<223> primer: HRCHECK II S13-RV

<400> 26

ctccgaagga gtctctcagt g

21

<210> 27

<211> 17

<212> DNA

<213> Artificial

<220>

<223> primer: T7

<400> 27

aatacgactc actatag

17

<210> 28

<211> 32

<212> DNA

<213> Artificial

<220>

<223> primer: HXGPRT/BGLII-FW

<400> 28

cgatagatct aaaatggcgt ccaaaccat tg

32

<210> 29

<211> 31

<212> DNA

<213> Artificial

<220>

<223> primer: HXGPRT/PSTI-RV

<400> 29

cgatctgcag ttacttctcg aactttttgc g

31