



(86) Date de dépôt PCT/PCT Filing Date: 1999/07/19
(87) Date publication PCT/PCT Publication Date: 2000/02/03
(45) Date de délivrance/Issue Date: 2010/10/12
(85) Entrée phase nationale/National Entry: 2001/01/19
(86) N° demande PCT/PCT Application No.: NL 1999/000460
(87) N° publication PCT/PCT Publication No.: 2000/005378
(30) Priorités/Priorities: 1998/07/22 (EP98202467.1);
1998/07/22 (EP98202465.5)

(51) Cl.Int./Int.Cl. *C12N 15/31* (2006.01),
A61K 39/09 (2006.01), *C07K 14/315* (2006.01),
C12N 1/21 (2006.01), *C12N 15/54* (2006.01),
C12N 9/10 (2006.01), *C12P 19/04* (2006.01),
C12Q 1/14 (2006.01), *C12Q 1/68* (2006.01),
G01N 33/569 (2006.01), *A61K 39/00* (2006.01)

(72) Inventeur/Inventor:
SMITH, HILDA ELIZABETH, NL

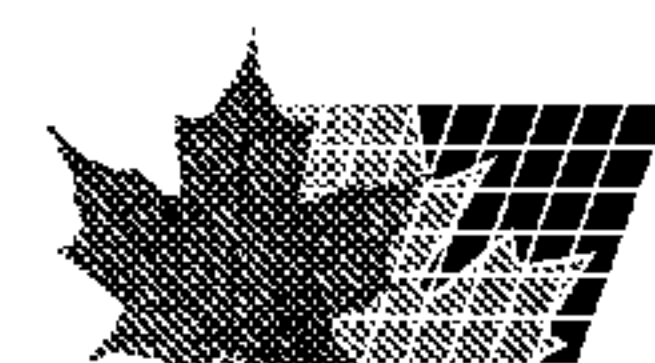
(73) Propriétaire/Owner:
STICHTING DIENST LANDBOUWKUNDIG
ONDERZOEK, NL

(74) Agent: BORDEN LADNER GERVAIS LLP

(54) Titre : VACCINS CONTRE LE STREPTOCOCCUS SUIS ET TESTS DIAGNOSTIQUES
(54) Title: STREPTOCOCCUS SUIS VACCINES AND DIAGNOSTIC TESTS

(57) **Abrégé/Abstract:**

The invention relates to Streptococcus suis infections of pigs, to vaccines directed against those infections and to tests for diagnosing Streptococcus suis infections. The invention provides an isolated or recombinant nucleic acid encoding a capsular gene cluster of Streptococcus suis or a gene or gene fragment derivated thereof. The invention furthermore provides a nucleic acids probe or primer allowing species or serotype specific detection of Streptococcus suis. The invention also provides a Streptococcus suis antigen and vaccine derived thereof.



**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/31, 15/54, 9/10, 1/21, C07K 14/315, C12P 19/04, A61K 39/09, C12Q 1/14, 1/68, G01N 33/569	A2	(11) International Publication Number: WO 00/05378 (43) International Publication Date: 3 February 2000 (03.02.00)
(21) International Application Number: PCT/NL99/00460 (22) International Filing Date: 19 July 1999 (19.07.99) (30) Priority Data: 98202465.5 22 July 1998 (22.07.98) EP 98202467.1 22 July 1998 (22.07.98) EP (71) Applicant (for all designated States except US): STICHTING DIENST LANDBOUWKUNDIG ONDERZOEK [NL/NL]; Bornsesteeg 53, NL-6708 PD Wageningen (NL). (72) Inventor; and (75) Inventor/Applicant (for US only): SMITH, Hilda, Elizabeth [NL/NL]; Golfpark 98, NL-8241 AG Lelystad (NL). (74) Agent: OTTEVANGERS, S., U.; Vereenigde Octrooibureaux, Nieuwe Parklaan 97, NL-2587 BN The Hague (NL).	(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, 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, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(54) Title: <i>STREPTOCOCCUS SUIIS</i> VACCINES AND DIAGNOSTIC TESTS (57) Abstract The invention relates to <i>Streptococcus suis</i> infections of pigs, to vaccines directed against those infections and to tests for diagnosing <i>Streptococcus suis</i> infections. The invention provides an isolated or recombinant nucleic acid encoding a capsular gene cluster of <i>Streptococcus suis</i> or a gene or gene fragment derivated thereof. The invention furthermore provides a nucleic acids probe or primer allowing species or serotype specific detection of <i>Streptococcus suis</i>. The invention also provides a <i>Streptococcus suis</i> antigen and vaccine derived thereof.		

WO 00/05378

PCT/NL99/00460

Title: *Streptococcus suis* vaccines and diagnostic tests.

The invention relates to *Streptococcus* infections of pigs, to vaccines directed against those infections, to tests for diagnosing *Streptococcus* infections and to the field of bacterial vaccines, more in particular to vaccines directed
5 against *Streptococcus* infections.

Streptococcus species, of which there are a large variety causing infections in domestic animals and man, are often grouped according to Lancefield's groups. Typing according to Lancefield occurs on the basis of serological determinants or
10 antigens that are among others present in the capsule of the bacterium and allows for only an approximate determination, often bacteria from a different group show cross-reactivity with each other, while other *Streptococci* can not be assigned a group-determinant at all. Within groups, further
15 differentiation is often possible on the basis of serotyping; these serotypes further contribute to the large antigenic variability of *Streptococci*, a fact that creates an array of difficulties within diagnosis of and vaccination against *Streptococcal* infections.

20 Lancefield group A *Streptococcus* species (GAS, *Streptococcus pyogenes*), are common with children, causing nasopharyngeal infections and complications thereof. Among animals, especially cattle are susceptible to GAS, whereby often mastitis is found.

25 Group A streptococci are the etiologic agents of streptococcal pharyngitis and impetigo, two of the commonest bacterial infections in children, as well as a variety of less common but potentially life-threatening infections, including soft tissue infections, bacteraemia, and pneumonia. In
30 addition, GAS are uniquely associated with the postinfectious autoimmune syndromes of acute rheumatic fever and poststreptococcal glomerulonephritis.

Several recent reports suggest that the incidence both of serious infections due to GAS and of acute rheumatic fever has

WO 00/05378

PCT/NL99/00460

increased during the past decade, focusing renewed interest on defining the attributes or virulence factors of the organism that may play a role in the pathogenesis of these diseases.

5 GAS produce several surface components and extracellular products that may be important in virulence. The major surface protein, M protein, has been studied in the most detail and has been shown convincingly to play a role in both virulence and immunity. Isolates rich in M protein are able to grow in human blood, a property thought to reflect the capacity of M
10 protein to interfere with phagocytosis, and these isolates tend to be virulent in experimental animals.

Lancefield group B *Streptococcus* (GBS) are most often seen with cattle, causing mastitis, however, human infants are susceptible as well, often with fatal consequences. Group B
15 streptococci (GBS) constitute a major cause of bacterial sepsis and meningitis among human neonates born in the United States and Western Europe and are emerging as significant neonatal pathogens in developing countries as well.

It is estimated that GBS strains are responsible for
20 10,000 to 15,000 cases of invasive infection in neonates in the United States alone. Despite advances in early diagnosis and treatment, neonatal sepsis due to GBS continues to carry a mortality rate of 15 to 20%. In addition, survivors of GBS meningitis have 30 to 50% incidence of long-term neurologic
25 sequelae. The increasing recognition over the past two decades of GBS as an important pathogen for human infants has generated renewed interest in defining the bacterial and host factors important in virulence of GBS and in the immune response to GBS infection.

30 Particular attention has focused on the capsular polysaccharide as the predominant surface antigen of the organisms. In a modification of the system originally developed by Rebecca Lancefield, GBS strains are serotyped on the basis of antigen differences in their capsular
35 polysaccharides and the presence or absence of serologically defined C proteins. While GBS isolated from non-human sources

WO 00/05378

PCT/NL99/00460

often lack a serologically detectable capsule, a large majority of strains associated with neonatal infection belong to one of four major capsular serotypes, 1a, 1b, II or III. The capsular polysaccharide forms the outermost layer around the exterior of the bacterial cell, superficial to the cell wall. The capsule is distinct from the cell wall-associated group B carbohydrate. It has been suggested that the presence of sialic acid in the capsule of bacteria that cause meningitis is important for these bacteria to breach the blood-brain barrier. Indeed, in *S. agalactiae* sialic acid has shown to be critical for the virulence function of the type III capsule. The capsule of *S. suis* serotype is composed of glucose, galactose, N-acetylglucosamine, rhamnose and sialic acid.

The group B polysaccharide, in contrast to the type-specific capsule, is present on all GBS strains and is the basis for serogrouping of the organisms into Lancefield's group B. Early studies by Lancefield and co-workers showed that antibodies raised in rabbits against whole GBS organisms protected mice against challenge with strains of homologous capsular type, demonstrating the central role of the capsular polysaccharide as a protective antigen. Studies in the 1970s by Baker and Kasper demonstrated that cord blood of human infants with type III GBS sepsis uniformly had low or undetectable levels of antibodies directed against the type III capsule, suggesting that a deficiency of anticapsular antibody was a key factor in susceptibility of human neonates to GBS disease.

Lancefield group C infections, such as those with *S. equi*, *S. zooepidemicus*, *S. dysgalactiae*, and others are mainly seen with horse, cattle and pigs, but can also cross the species barrier to humans. Lancefield group D (*S. bovis*) infections are found with all mammals and some birds, sometimes resulting in endocarditis or septicaemia.

Lancefield groups E, G, L, P, U and V (*S. porcinus*, *S. canis*, *S. dysgalactiae*) are found with various hosts, causing

WO 00/05378

PCT/NL99/00460

neonatal infections, nasopharyngeal infections or mastitis.

Within Lancefield groups R, S, and T, (and with ungrouped types) *S. suis* is found, an important cause of meningitis, septicemia, arthritis and sudden death in young pigs.

5 Incidentally, it can also cause meningitis in man.

Streptococcus suis is an important cause of meningitis, septicemia, arthritis and sudden death in young pigs (4, 46). Incidentally, it can also cause meningitis in man (1). *S. suis* strains are usually identified and classified by their
10 morphological, biochemical and serological characteristics (58, 59, 46). Serological classification is based on the presence of specific antigenic polysaccharides. So far, 35 different serotypes have been described (9, 56, 14). In several European countries, *S. suis* serotype 2 is the most prevalent type
15 isolated from diseased pigs, followed by serotypes 9 and 1. Serological typing of *S. suis* is carried out using different types of agglutination tests. In these tests, isolated and biochemically characterised *S. suis* cells are agglutinated with a panel of 35 specific sera. These methods are very laborious
20 and time-consuming.

Little is known about the pathogenesis of the disease caused by *S. suis*, let alone about its various serotypes such as type 2. Various bacterial components, such as extracellular and cell-membrane associated proteins, fimbriae, haemagglutinins,
25 and haemolysin have been suggested as virulence factors (9, 10, 11, 15, 16, 47, 49). However, the precise role of these protein components in the pathogenesis of the disease remains unclear (37). It is well known that the polysaccharidic capsule of various *Streptococci* and other gram-positive bacteria plays an
30 important role in pathogenesis (3, 6, 35, 51, 52). The capsule enables these micro-organisms to resist phagocytosis and is therefore regarded as an important virulence factor. Recently, a role of the capsule of *S. suis* in the pathogenesis was suggested as well (5). However, the structure, organisation and
35 functioning of the genes responsible for capsule polysaccharide synthesis (*cps*) in *S. suis* is unknown. Within *S. suis* serotypes

WO 00/05378

PCT/NL99/00460

1 and 2 strains can differ in virulence for pigs (41, 45, 49). Some type 1 and 2 strains are virulent, other strains are not. Because both virulent and non-virulent strains of serotype 1 and 2 strains are fully encapsulated, it may even be that
5 capsule is not a relevant factor required for virulence.

Attempts to control *S. suis* infections or disease are still hampered by the lack of knowledge about the epidemiology of the disease and the lack of effective vaccines and sensitive diagnostics. It is well known and generally accepted
10 that the polysaccharidic capsule of various Streptococci and other gram-positive bacteria plays an important role in pathogenesis. The capsule enables these micro-organisms to resist phagocytosis and is therefore regarded as an important virulence factor.

15 Compared to encapsulated *S. suis* strains, non-encapsulated *S. suis* strains are phagocytosed by murine polymorphonuclear leucocytes to a greater degree. Moreover, an increase in thickness of capsule was noted for *in vivo* grown virulent strains while no increase was observed for avirulent
20 strains. Therefore, these data again demonstrate the role of the capsule in the pathogenesis for *S. suis* as well.

Ungrouped *Streptococcus* species, such as *S. mutans*, causing caries with humans, *S. uberis*, causing mastitis with cattle, and *S. pneumoniae*, causing major infections in humans, and
25 *Enterococcus faecilalis* and *E. faecium*, further contributed to the large group of Streptococci.

Streptococcus pneumoniae (the pneumococcus) is a human pathogen causing invasive diseases, such as pneumonia, bacteraemia, and meningitis. Despite the availability of
30 antibiotics, pneumococcal infections remain common and can still be fatal, especially in high-risk groups, such as young children and elderly people. Particularly in developing countries, many children under the age of five years die each year from pneumococcal pneumonia. *S. pneumoniae* is also the
35 leading cause of otitis media and sinusitis. These infections are less serious, but nevertheless incur substantial medical

WO 00/05378

PCT/NL99/00460

costs, especially when leading to complications, such as permanent deafness. The normal ecological niche of the pneumococcus is the nasopharynx of man. The entire human population is colonised by the pneumococcus at one time or another, and at a given time, up to 60% of individuals may be carriers. Nasopharyngeal carriage of pneumococci by man is often accompanied by the development of protection to infection by the same serotype. Most infections do not occur after prolonged carriage but follow the acquisition of recently acquired strains. Many bacteria contain surface polysaccharides which act as a protective layer against the environment. Surface polysaccharides of pathogenic bacteria usually make the bacteria resistant to the defense mechanisms of the host, e.g., the lytic action of serum or phagocytosis. In this respect, the serotype-specific capsular polysaccharide (CP) of *Streptococcus pneumoniae*, is an important virulence factor. Unencapsulated strains are avirulent, and antibodies directed against the CP are protective. Protection is serotype specific; each serotype has its own, specific CP structure. Ninety different capsular serotypes have been identified. Currently, CPs of 23 serotypes are included in a vaccine.

Vaccines directed against *Streptococcus* infections in general aim at utilising an immune response directed against the polysaccharide capsule of the various *Streptococcus* species, especially since the capsule is considered a main virulence factor for these bacteria. The capsule, during infection, provides resistance to phagocytosis and thus promotes the escape of the bacteria from the immune system of the host, protecting the bacteria by elimination by macrophages and neutrophils.

The capsule particularly confers the bacterium resistance to complement-mediated opsonophagocytosis. In addition, some bacteria express capsular polysaccharides (CPs) that mimic host molecules, thereby avoiding the immune system of the host. Also, even when the bacteria have been phagocytosed, intracellular killing is hampered by the presence of a

WO 00/05378

PCT/NL99/00460

capsule.

It is in general thought that only when the host has antibodies or other serum-factors directed against capsule antigens, the bacterium will get recognised by the immune system through the anticapsular-antibodies or serum-factors bound to its capsule, and will, through opsonisation, get phagocytosed and killed.

However, these antibodies are serotype-specific, and will often only confer protection against only one of the many serotypes known within a group of *Streptococci*.

For example, current commercially available *S. suis* vaccines, which are in general based on whole-cell-bacterial preparations, or on capsule-enriched fractions of *S. suis*, confer only limited protection against heterologous strains. Also, the current pneumococcal vaccine, licensed in the United States in 1983, consists of purified CPs of 23 pneumococcal serotypes whereas at least 90 CP types exist.

The composition of this pneumococcal vaccine was based on the frequency of the occurrence of disease isolates in the US and cross-reactivity between various serotypes. Although this vaccine protects healthy adults against infections caused by serotypes included in the vaccine, it fails to raise a protective immune response in infants younger than 18 months and it is less effective in elderly people. In addition, the vaccine confers only limited protection in patients with immunodeficiencies and haematology malignancies. In the light of above, improved vaccines are needed against *Streptococcus* infections. Much attention is being paid at producing CP vaccines by producing the relevant polysaccharides via chemical or recombinant means. However, chemical synthesis of polysaccharides is costly, and capsular polysaccharide synthesis by recombinant means necessitates knowledge about the relevant genes, which are not always available and need to be determined for each and every relevant serotype.

WO 00/05378

PCT/NL99/00460

The invention provides an isolated or recombinant nucleic acid encoding a capsular (*cps*) gene cluster of *Streptococcus suis*. Biosynthesis of capsule polysaccharides in general has been studied in a number of Gram-positive and Gram-negative bacteria (32). In Gram-negative bacteria, but also in a number of gram-positive bacteria, genes which are involved in the biosynthesis of polysaccharides are clustered at a single locus. *Streptococcus suis* capsular genes as provided by the invention show a common genetic organisation involving three distinct regions. The central region is serotype specific and encodes enzymes responsible for the synthesis and polymerisation of the polysaccharides. This region is flanked by two regions conserved in *Streptococcus suis* which encode proteins for common functions such as transport of the polysaccharide across the cellular membrane. However, in between species, only low homologies exist, hampering easy comparison and detection of seemingly similar genes. Knowing the nucleic acid encoding the flanking regions allows type-specific determination of nucleic acid of the central region of *Streptococcus suis* serotypes, as for example described in the experimental part of the description of the invention.

The invention provides an isolated or recombinant nucleic acid encoding a capsular gene cluster of *Streptococcus suis* or a gene or gene fragment derived thereof. Such a nucleic acid is for example provided by hybridising chromosomal DNA derived from any one of the *Streptococcus suis* serotypes to a nucleic acid encoding a gene derived from a *Streptococcus suis* serotype 1, 2 or 9 capsular gene cluster, as provided by the invention (see for example Tables 4 and 5) and cloning of (type-specific) genes as for example described in the experimental part of the description. At least 14 open reading frames are identified. Most of the genes belong to a single transcriptional unit, identifying a co-ordinate control of these genes, they, and the enzymes and proteins they encode, act in concert to provide the capsule with the relevant polysaccharides. The invention provides *cps* genes and proteins

WO 00/05378

PCT/NL99/00460

encoded thereof involved in regulation (CpsA), chain length determination (CpsB, C), export (CpsC) and biosynthesis (CpsE, F, G, H, J, K). Although the overall organisation seemed at first glance to be similar to that of the *cps* and *eps* gene clusters of a number of Gram-positive bacteria (19, 32, 42), overall homologies are low (see table 3). The region involved in biosynthesis is located at the centre of the gene cluster and is flanked by two regions containing genes with more common functions.

10 The invention provides an isolated or recombinant nucleic acid encoding a capsular gene cluster of *Streptococcus suis* serotype 2 or a gene or gene fragment derived thereof, preferably as identified in Figure 3. Genes in this gene cluster are involved in polysaccharide biosynthesis of
15 capsular components and antigens. For a further description of such genes see for example Table 2 of the description, for example a *cpsA* gene is provided functionally encoding regulation of capsular polysaccharide synthesis, whereas *cpsB* and *cpsC* are functionally involved in chain in chain length
20 determination. Other genes, such as *cpsD*, E, F, G, H, I, J, K and related genes, are involved in polysaccharide syntheses, functioning for example as glucosyl- or glycosyltransferase. The *cpsF*, G, H, I, J genes encode more type-specific proteins than the flanking genes which are found more-or-less conserved
25 throughout the species and can serve as base for selection of primers or probes in PCR-amplification or cross-hybridisation experiments for subsequent cloning.

For example, the invention further provides an isolated or
30 recombinant nucleic acid encoding a capsular gene cluster of *Streptococcus suis* serotype 1 or a gene or gene fragment derived thereof, preferably as identified in Figure 4.

In addition, the invention provides an isolated or recombinant nucleic acid encoding a capsular gene cluster of
35 *Streptococcus suis* serotype 9 or a gene or gene fragment derived thereof, preferably as identified in Figure 5.

WO 00/05378

PCT/NL99/00460

Furthermore, the invention provides for example a fragment or parts thereof of the *cps* locus, involved in the capsular polysaccharide biosynthesis, of *S. suis*, exemplified in the experimental part for serotype 1, 2 or 9, and allows easy
5 identification or detection of related fragments derived of other serotype of *S. suis*.

The invention provides a nucleic acid probe or primer derived from a nucleic acid according to the invention allowing species or serotype specific detection of
10 *Streptococcus suis*. Such a probe or primer (herein used interchangeably) is for example a DNA, RNA or PNA (peptide nucleic acid) probe hybridising with capsular nucleic acid as provided by the invention. Species specific detection is provided preferably by selecting a probe or primer sequence
15 from a species-specific region (e.g. flanking region) whereas serotype specific detection is provided preferably by selecting a probe or primer sequence from a type-specific region (e.g. central region) of a capsular gene cluster as provided by the invention. Such a probe or primer can be used
20 in a further unmodified form, for example in cross-hybridisation or polymerase-chain reaction (PCR) experiments as for example described in the experimental part of the description of the invention. Herein the invention provides the isolation and molecular characterisation of additional
25 type specific *cps* genes of *S. suis* types 1 and 9. In addition, we describe the genetic diversity of the *cps* loci of serotypes 1, 2 and 9 among the 35 *S. suis* serotypes yet known. Type-specific probes are identified. Also, a type-specific PCR for for example serotype 9 is provided, being a rapid, reliable
30 and sensitive assay, which is used directly on nasal or tonsillar swabs or other samples of infected or carrier animals.

The invention also provides a probe or primer according to the invention further provided with at least one reporter
35 molecule. Examples of reporter molecules are manifold and known in the art, for example a reporter molecule can comprise

WO 00/05378

PCT/NL99/00460

additional nucleic acid provided with a specific sequence (e.g. oligo-dT) hybridising to a corresponding sequence to which hybridisation can easily be detected for example because it has been immobilised to a solid support.

5 Yet other reporter molecules comprise chromophores, e.g. fluorochromes for visual detection, for example by light microscopy or fluorescent in situ hybridisation (FISH) techniques, or comprise an enzyme such as horseradish peroxidase for enzymatic detection, e.g. in enzyme-linked
10 assays (EIA). Yet other reporter molecules comprise radioactive compounds for detection in radiation-based-assays.

In a preferred embodiment of the invention, at least one probe or primer according to the invention is provided (labelled) with a reporter molecule and a quencher molecule,
15 providing together with unlabeled probe or primer a PCR-based test allowing rapid detection of specific hybridisation.

The invention further provides a diagnostic test or test kit comprising a probe or primer as provided by the invention. Such a test or test kit, for example a cross-hybridisation
20 test or PCR-based test, is advantageously used in rapid detection and/or serotyping of *Streptococcus suis*.

The invention furthermore provides a protein or fragment thereof encoded by a nucleic acid according to the invention. Examples of such a protein or fragment are for example
25 proteins described in for example Table 2 of the description, for example a cpsA protein is provided functionally encoding regulation of capsular polysaccharide synthesis, whereas cpsB and cpsC are functionally involved in chain in chain length determination. Other proteins or functional fragments thereof
30 as provided by the invention, such as cpsD, E, F, G, H, I, J, K and related proteins, are involved in polysaccharide biosynthesis, functioning for example as glucosyl- or glycosyltransferase in polysaccharide biosynthesis of *Streptococcus suis* capsular antigen.

35 The invention furthermore provides a method to produce a *Streptococcus suis* capsular antigen comprising using a protein

WO 00/05378

PCT/NL99/00460

or functional fragment thereof as provided by the invention, and provides therewith a *Streptococcus suis* capsular antigen obtainable by such a method. A comparison of the predicted amino acid sequences of the *cps2* genes with sequences found in the databases allowed the assignment of functions to the open reading frames. The central region contains the type specific glycosyltransferases and the putative polysaccharide polymerase. This region is flanked by two regions encoding for proteins with common functions, such as regulation and transport of polysaccharide across the membrane.

Biosynthesis of *Streptococcus* capsular polysaccharide antigen using a protein or functional fragment thereof is advantageously used in chemo-enzymatic synthesis and the development of vaccines which offer protection against serotype-specific Streptococcal disease, and is also advantageously used in the synthesis and development of multivalent vaccines against Streptococcal infections. Such vaccines elicit anticapsular antibodies which confer protection.

Furthermore, the invention provides an acapsular *Streptococcus* mutant for use in a vaccine, a vaccine strain derived thereof and a vaccine derived thereof. Surprisingly, and against the grain of common doctrine, the invention provides use of a *Streptococcus* mutant deficient in capsular expression in a vaccine.

Acapsular *Streptococcus* mutants have long been known in the art and can be found in nature. Griffith (J. Hyg. 27:113-159, 1928) demonstrated that pneumococci could be transformed from one type to another. If he injected live rough (acapsular or unencapsulated) type 2 pneumococci into mice, the mice would survive. If, however, he injected the same dose of live rough type 2 mixed with heat-killed smooth (encapsulated) type 1 into a mouse, the mouse would die, and from the blood he could isolate live smooth type 1 pneumococci. At that time, the significance of this transforming principle was not understood. However, understanding came when it was shown that

WO 00/05378

PCT/NL99/00460

DNA constituted the genetic material responsible for phenotypic changes during transformation.

Streptococcus mutants deficient in capsular expression are found in several forms. Some are fully deficient and have
5 no capsule at all, others form a deficient capsule, characterised by a mutation in a capsular gene cluster. Deficiency can for instance include capsular formation wherein the organization of the capsular material has been re-arranged, as for example demonstrable by electron microscopy.
10 Yet others have a nearly fully developed capsule which is only deficient in a particular sugar component.

Now, after much advance of biotechnology and despite the fact that little is still known about the exact localisation and sequence of genes involved in capsular synthesis in
15 *Streptococci*, it is possible to create mutants of *Streptococci*, for example by homologous recombination or transposon mutagenesis, which has for example been done for GAS (Wessels et al., PNAS 88:8317-8321, 1991), for GBS (Wessels et al., PNAS 86: 8983-8987, 1989), for *S. suis* (Smith, ID-DLO
20 Annual report 1996, page 18-19; Charland et al., Microbiol. 144:325-332, 1998) and for *S. pneumonia* (Kolkman et al., J. Bact. 178:3736-3741, 1996). Such recombinant derived mutants, or isogenic mutants, can easily be compared with the wild-type strains from which they have been derived.

25 In a preferred embodiment, the invention provides use of a recombinant-derived *Streptococcus* mutant deficient in capsular expression in a vaccine. Recombinant techniques useful in producing such mutants are for example homologous recombination, transposon mutagenesis, and others, whereby
30 deletions, insertions or (point)-mutations are introduced in the genome. Advantages of using recombinant techniques are the stability of the obtained mutants (especially with homologous recombination and double cross-over techniques), and the knowledge about the exact site of the deletion, mutation or
35 insertion.

In a much preferred embodiment, the invention provides a

WO 00/05378

PCT/NL99/00460

stable mutant deficient in capsular expression obtainable for example through homologous recombination or cross over integration events. Examples of such a mutant can be found in the experimental part of this description, for example mutant
5 10cpsB or 10cpsEF is such a stable mutant as provided by the invention.

The invention also provides a *Streptococcus* vaccine strain and vaccine that has been derived from a *Streptococcus* mutant deficient in capsular expression. In general, said
10 strain or vaccine is applicable within the whole range of Streptococcal infections, be it for those with animals or man or with zoonotic infections. It is of course now possible to first select a common vaccine strain and derive a
15 *Streptococcus* mutant deficient in capsular expression thereof for the selection of a vaccine strain and use in a vaccine according to the invention.

In a preferred embodiment, the invention provides use of a *Streptococcus* mutant deficient in capsular expression in a vaccine wherein said *Streptococcus* mutant is selected from
20 the group composed of *Streptococcus* group A, *Streptococcus* group B, *Streptococcus suis* and *Streptococcus pneumoniae*. Herewith the invention provides vaccine strains and vaccines for use with these notoriously heterologous Streptococci, of which a multitude of serotypes exist. With a vaccine as
25 provided by the invention that is derived from a specific *Streptococcus* mutant that deficient in capsular expression, the difficulties relating to lack of heterologous protection can be circumvented since these mutants do not rely on capsular antigens per se to induce protection.

30 In a preferred embodiment, said vaccine strain is selected for its ability to survive or even replicate in an immune-competent host or host cells and thus can persist for a certain period, varying from 1-2 days to more than one or two weeks, in a host, despite its deficient character.

35 Although an immunodeficient host will support replication of a wide range of bacteria that are deficient in one or more

virulence factors, in general it is considered a characteristic of pathogenicity of Streptococci that they can survive for certain periods or replicate in a normal host or host cells such as macrophages. For example, Williams and
5 Blakemore (Neuropath. Appl. Neurobiol.: 16, 345-356, 1990; Neuropath. Appl. Neurobiol.: 16, 377-392, 1990; J. Infect. Dis.: 162, 474-481, 1990) show that both polymorphonuclear cells and macrophage cells are capable of phagocytosing pathogenic *S. suis* in pigs lacking anti-*S. suis* antibodies,
10 only pathogenic bacteria could survive and multiply inside macrophages and the pig.

In a preferred embodiment, the invention, however, provides a deficient or avirulent mutant or vaccine strain which is capable of surviving at least 4-5 days, preferably at
15 least 8-10 days in said host, thereby allowing the development of a solid immune response to subsequent *Streptococcus* infection,

Due to its persistent but avirulent character, a *Streptococcus* mutant or vaccine strain as provided by the
20 invention is well suited to generate specific and/or long-lasting immune responses against Streptococcal antigens, moreover because possible specific immune responses of the host directed against a capsule are relatively irrelevant because a vaccine strain as provided by the invention is in
25 general not recognised by such antibodies.

In addition, the invention provides a *Streptococcus* vaccine strain according to the invention which strain comprises a mutant capable of expressing a *Streptococcus* virulence factor or antigenic determinant.

30 In a preferred embodiment, the invention provides a *Streptococcus* vaccine strain according to the invention which strain comprises a mutant capable of expressing a *Streptococcus* virulence factor wherein said virulence factor or antigenic determinant is selected from a group of cellular
35 components, such as muramidase-released protein (MRP) extracellular factor (EF) and cell-membrane associated

WO 00/05378

PCT/NL99/00460

proteins, 60kDA heat shock protein, pneumococcal surface protein A (Psp A), pneumolysin, C protein, protein M, fimbriae, haemagglutinins and haemolysin or components functionally related thereto.

5 In a much preferred embodiment, the invention provides a *Streptococcus* vaccine strain according to the invention which strain comprises a mutant capable of over-expressing said virulence factor. In this way, the invention provides a vaccine strain for incorporation in a vaccine which
10 specifically causes a host to provide a immune response directed against antigenically important determinants of virulence (listed above), thereby providing specific protection directed against said determinants. Over-expression can for example be achieved by cloning the gene involved
15 behind a strong promoter, which is for example constitutionally expressed in a multicopy system, either in a plasmid or via intergration in a genome.

In yet another embodiment, the invention provides a *Streptococcus* vaccine strain according to the invention which
20 comprises a mutant capable of expressing a non-*Streptococcus* protein. Such a vector-*Streptococcus* vaccine strain allows, when used in a vaccine, protection against other pathogens than *Streptococcus*.

Due to its persistent but avirulent character, a
25 *Streptococcus* vaccine strain or mutant as provided by the invention is well suited to generate specific and long-lasting immune responses, not only against Streptococcal antigens, but also against other antigens when these are expressed by said strain. Especially antigens derived from another pathogen are
30 now expressed without the detrimental effects of said antigen or pathogen which would otherwise have harmed the host.

An example of such a vector is a *Streptococcus* vaccine strain or mutant wherein said antigen is derived from a pathogen, such as *Actinobacillus pleuropneumonia*,
35 *Mycoplasmatae*, *Bordetella*, *Pasteurella*, *E. coli*, *Salmonella*, *Campylobacter*, *Serpulina* and others.

WO 00/05378

PCT/NL99/00460

The invention also provides a vaccine comprising a *Streptococcus* vaccine strain or mutant according to the invention and further comprising a pharmaceutically acceptable carrier or adjuvant. Carriers or adjuvants are well known in the art, examples are phosphate buffered saline, physiological salt solutions, (double-)oil-in-water-emulsions, aluminumhydroxide, Specol, block- or co-polymers, and others.

A vaccine according to the invention can comprise a vaccine strain either in a killed or live form. For example, a killed vaccine comprising a strain having (over)expressed a Streptococcal or heterologous antigen or virulence factor is very well suited for eliciting an immune response. In a preferred embodiment, the invention provides a vaccine wherein said strain is live, due to its persistent but avirulent character, a *Streptococcus* vaccine strain as provided by the invention is well suited to generate specific and long-lasting immune responses.

Now that a Streptococcal vaccine is provided by the invention, the invention also provides a method for controlling or eradicating a Streptococcal disease in a population comprising vaccinating subjects in said population with a vaccine according to the invention.

In a preferred embodiment, a method for controlling or eradicating a Streptococcal disease is provided comprising testing a sample, such as a blood sample, or nasal or throat swab, faeces, urine, or other samples such as can be sampled at or after slaughter, collected from at least one subject, such as an infant or a pig, in a population partly or wholly vaccinated with a vaccine according to the invention for the presence of encapsulated Streptococcal strains or mutants. Since a vaccine strain or mutant according to the invention is not pathogenic, and can be distinguished from wild-type strains by capsular expression, the detection of (fully) encapsulated Streptococcal strains indicates that wild-type infections are still present. Such wild-type infected subjects can than be isolated from the remainder of the population

WO 00/05378

PCT/NL99/00460

until the infection has passed away. With domestic animals, such as pigs, it is even possible to remove the infected subject from the population as a whole by culling. Detection of wild-type strains can be achieved via traditional culturing techniques, or by rapid detection techniques such as PCR detection.

In yet another embodiment, the invention provides a method for controlling or eradicating a Streptococcal disease comprising testing a sample collected from at least one subject in a population partly or wholly vaccinated with a vaccine according to the invention for the presence of capsule-specific antibodies directed against Streptococcal strains. Capsule specific antibodies can be detected with classical techniques known in the art, such as used for Lancefield's group typing or serotyping.

A much preferred embodiment of a method provided by the invention for controlling or eradicating a Streptococcal disease in a population comprises vaccinating subjects in said population with a vaccine according to the invention and testing a sample collected from at least one subject in said population for the presence of encapsulated Streptococcal strains and/or for the presence of capsule-specific antibodies directed against Streptococcal strains.

For example, a method is provided according to the invention wherein said Streptococcal disease is caused by *Streptococcus suis*.

The invention also provides a diagnostic assay for testing a sample for use in a method according to the invention comprising at least one means for the detection of encapsulated Streptococcal strains and/or for the detection of capsule-specific antibodies directed against Streptococcal strains.

The invention furthermore provides a vaccine comprising an antigen according to the invention and further comprising a suitable carrier or adjuvant. The immunogenicity of a capsular antigen provided by the invention is for example increased by

linking to a carrier (such as a carrier protein), allowing the recruitment of T-cell help in developing an immune response.

The invention further provides a recombinant micro-organism provided with at least a part of a capsular gene cluster derived from *Streptococcus suis*. The invention provides for example a lactic acid bacterium provided with at least a part of a capsular gene cluster derived from *Streptococcus suis*. Various food-grade lactic acid bacteria (*Lactococcus lactis*, *Lactobacillus casei*, *Lactobacillus plantarium* and *Streptococcus gordonii*) have been used as delivery systems for mucosal immunization. It has now been shown that oral (or mucosal) administration of recombinant *L. lactis*, *Lactobacillus*, and *Streptococcus gordonii* can elicit local IgA and /or IgG antibody responses to an expressed antigen. The use of oral routes for immunization against infective diseases is desirable because oral vaccines are easier to administer, have higher compliance rates, and because mucosal surfaces are the portals of entry for many pathogenic microbial agents. It is within the skill of the artisan to provide such micro-organisms with (additional) genes.

The invention further provides a recombinant *Streptococcus suis* mutant provided with a modified capsular gene cluster. It is within the skill of the artisan to swap genes within a species. In a preferred embodiment, an avirulent *Streptococcus suis* mutant is selected to be provided with at least a part of a modified capsular gene cluster according to the invention.

The invention further provides a vaccine comprising a micro-organism or a mutant provided by the invention. An advantage of such a vaccine over currently used vaccines is that they comprise accurately defined micro-organisms and well-characterised antigens, allowing accurate determination of immune responses against various antigens of choice.

The invention is further explained in the experimental part of this description without limiting the invention thereto.

Experimental part

MATERIAL AND METHODS

5

Bacterial strains and growth conditions.

The bacterial strains and plasmids used in this study are listed in Table 1. *S. suis* strains were grown in Todd-Hewitt broth (code CM189, Oxoid), and plated on Columbia agar blood base (code CM331, Oxoid) containing 6% (v/v) horse blood. *E. coli* strains were grown in Luria broth (28) and plated on Luria broth containing 1.5% (w/v) agar. If required, antibiotics were added to the plates at the following concentrations: spectinomycin: 100 ug/ml for *S. suis* and 50 ug/ml for *E. coli* and ampicillin, 50 ug/ml.

Serotyping. The *S. suis* strains were serotyped by the slide agglutination test with serotype-specific antibodies (44).

DNA techniques. Routine DNA manipulations were performed as described by Sambrook et al. (36).

Alkaline phosphatase activity. To screen for PhoA fusions in *E. coli*, plasmid libraries were constructed. Therefore, chromosomal DNA of *S. suis* type 2 was digested with *AluI*. The 300-500-bp fragments were ligated to *SmaI*-digested pPHOS2. Ligation mixtures were transformed to the PhoA⁻ *E. coli* strain CC118. Transformants were plated on LB media supplemented with 5-Bromo-4-chloro-3-indolylfosfaat (BCIP, 50 ug/ml, Boehringer, Mannheim, Germany). Blue colonies were purified on fresh LB/BCIP plates to verify the blue phenotype.

DNA sequence analysis. DNA sequences were determined on a 373A DNA Sequencing System (Applied Biosystems, Warrington, GB). Samples were prepared by use of a ABI/PRISM dye terminator cycle sequencing ready reaction kit (Applied Biosystems). Sequencing data were assembled and analyzed using the MacMollyTetra program. Custom-made sequencing primers were purchased from Life Technologies. Hydrophobic stretches within

WO 00/05378

PCT/NL99/00460

proteins were predicted by the method of Klein et al. (17). The BLAST program available on Netscape NavigatorTM was used to search for protein sequences related to the deduced amino acid sequences.

5 **Construction of gene-specific knock-out mutants of *S. suis*.** To construct the mutant strains 10cpsB and 10cpsEF we electrotransformed the pathogenic serotype 2 strain 10 (45, 49) of *S. suis* with pCPS11 and pCPS28 respectively. In these plasmids the *cpsB* and *cpsEF* genes were disturbed by the
10 insertion of a spectinomycin-resistance gene. To create pCPS11 the internal 400 bp *Pst*I-*Bam*HI fragment of the *cpsB* gene in pCPS7 was replaced by the *Sp*c^R gene. For this purpose pCPS7 was digested with *Pst*I and *Bam*HI and ligated to the 1,200-bp *Pst*I-*Bam*HI fragment, containing the *Sp*c^R gen, from pIC-spc. To
15 construct pCPS28 we have used pIC20R. In this plasmid we inserted the *Kpn*I-*Sal*I fragment from pCPS17 (resulting in pCPS25) and the *Xba*I-*Cla*I fragment from pCPS20 (resulting in pCPS27). pCPS27 was digested with *Pst*I and *Xho*I and ligated to the 1,200-bp *Pst*I-*Xho*I fragment, containing the *Sp*c^R gene of
20 pIC-spc. The electrotransformation to *S. suis* was carried out as described before (38).

Southern blotting and hybridization. Chromosomal DNA was isolated as described by Sambrook et al. (36). DNA fragments were separated on 0.8% agarose gels and transferred to Zeta-
25 Probe GT membranes (Bio-Rad) as described by Sambrook et al. (36). DNA probes were labelled with [(-³²P]dCTP (3000 Ci mmol⁻¹; Amersham) by use of a random primed labelling kit (Boehringer). The DNA on the blots was hybridized at 65°C with appropriate DNA probes as recommended by the supplier of the
30 Zeta-Probe membranes. After hybridization, the membranes were washed twice with a solution of 40 mM sodium phosphate, pH 7.2, 1 mM EDTA , 5% SDS for 30 min at 65°C and twice with a solution of 40 mM sodium phosphate, pH 7.2, 1 mM EDTA, 1% SDS for 30 min at 65°C.

35 **PCR.** The primers used in the *cps2J* PCR correspond to the positions 13791-13813 and 14465-14443 in the *S. suis cps2*

WO 00/05378

PCT/NL99/00460

locus. The sequences were: 5'-CAAACGCAAGGAATTACGGTATC-3' and 5'-GAGTATCTAAAGAATGCCTATTG-3'. The primers used for the *cps11* PCR correspond to the positions 4398-4417 and 4839-4821 in the *S. suis* *cps1* sequence. The sequences were: 5'-

5 GGCGGTCTAGCAGATGCTCG-3' and 5'-GCGAACTGTTAGCAATGAC-3'. The primers used in the *cps9H* PCR correspond to the positions 4406-4126 and 4494-4475 in the *S. suis* *cps9* sequence. The sequences were: 5'-GGCTACATATAATGGAAGCCC3' and 5'-CGGAAGTATCTGGGCTACTG-3'.

10 **Construction of gene-specific knock-out mutants of *S. suis*.** To construct the mutant strains 10cpsB and 10cpsEF we electrotransformed the pathogenic serotype 2 strain 10 of *S. suis* with pCPS11 and pCPS28 respectively. In these plasmids the *cpsB* and *cpsEF* genes were disturbed by the
15 insertion of a spectinomycin-resistance gene. To create pCPS11 the internal 400 bp *Pst*I-*Bam*HI fragment of the *cpsB* gene in pCPS7 was replaced by the *Spc*^R gene. For this purpose pCPS7 was digested with *Pst*I and *Bam*HI and ligated to the 1,200-bp *Pst*I-*Bam*HI fragment, containing the *Spc*^R gen, from pIC-spc. To
20 construct pCPS28 we have used pIC20R. In this plasmid we inserted the *Kpn*I-*Sal*I fragment from pCPS17 (resulting in pCPS25) and the *Xba*I-*Cla*I fragment from pCPS20 (resulting in pCPS27). pCPS27 was digested with *Pst*I and *Xho*I and ligated to the 1,200-bp *Pst*I-*Xho*I fragment, containing the *Spc*^R gene of
25 pIC-spc. The electrotransformation to *S. suis* was carried out as described before (38).

Phagocytosis assay. Phagocytosis assays were performed as described by Leij *et al.* (23). Briefly, to opsonize the cells, 10⁷ *S. suis* cells were incubated with 6% SPF-pig serum for 30
30 min at 37°C in a head-over-head rotor at 6 rpm. 10⁷ AM and 10⁷ opsonized *S. suis* cells were combined and incubated at 37°C under continuous rotation at 6 rpm. At 0, 30, 60 and 90 min, 1-ml samples were collected and mixed with 4 ml of ice-cold EMEM to stop phagocytosis. Phagocytes were removed by centrifugation
35 for 4 min at 110 x g and 4°C. The number of colony forming units (CFU) in the supernatants was determined. Control

WO 00/05378

PCT/NL99/00460

experiments were carried out simultaneously by combining 10^7 opsonized *S. suis* cells with EMEM (without AM).

Killing assays. AM (10^7 /ml) and opsonized *S. suis* cells (10^7 /ml) were mixed 1 : 1 and incubated for 10 min at 37°C under continuous rotation at 6 rpm. Ice-cold EMEM was added to stop further phagocytosis and killing. To remove extracellular *S. suis* cells, phagocytes were washed twice (4 min, 110 x g, 4°C) and resuspended in 5 ml EMEM containing 6% SPF serum. The tubes were incubated at 37°C under rotation at 6 rpm. After 0, 15, 30, 60 and 90 min, samples were collected and mixed with ice-cold EMEM to stop further killing. The samples were centrifuged for 4 min at 110 x g at 4°C and the phagocytic cells were lysed in EMEM containing 1% saponine for 20 min at room temperature. The number of CFU in the suspensions was determined.

Pigs. Germfree pigs, cross-breeds of Great Yorkshire and Dutch landrace, were obtained from sows by caesarian sections. The surgery was performed in sterile flexible film isolators. Pigs were allotted to groups, each consisting of 4 pigs, and were housed in sterile stainless steel incubators.

Experimental infections. Pigs were inoculated intranasally with *S. suis* type 2 as described before. To predispose the pigs for infection with *S. suis*, five-day old pigs were inoculated intranasally with about 10^7 CFU of *Bordetella bronchiseptica* strain 92932. Two days later the pigs were inoculated intranasally with *S. suis* type 2 (10^6 CFU). Pigs were monitored twice daily for clinical signs of disease, such as fever, nervous signs and lameness. Blood samples were collected three times a week from each pig. White blood cells were counted with a cell counter. To monitor infection with *S. suis* and *B. bronchiseptica* and to check for absence of contaminants, we collected swabs of nasopharynx and feces daily. The swabs were plated directly onto Columbia agar containing 6% horse blood. After three weeks the pigs were killed and examined for pathological changes. Tissue specimens from the central nervous system, serosae, and joints were examined bacteriologically and

WO 00/05378

PCT/NL99/00460

histologically as described before (45, 49). Colonization of the serosae was scored positively when *S. suis* was isolated from the pericardium, thoracal pleura or the peritoneum.

Colonization of the joints was scored positively when *S. suis* was isolated from one or more joints (12 joints per animal were scored).

Vaccination and challenge

One week old pigs were vaccinated intravenously with a dosage of 10⁶ cfu of the *S. suis* strains 10cpsEF or 10cpsB. Three weeks later the pigs were challenged intravenously with the pathogenic serotype 2 strain 10 (10⁷ cfu). Disease monitoring, haematological, serological and bacteriological examinations as well as post-mortum examinations were as described before under experimental infections.

Electron Microscopy. Bacteria were prepared for electron microscopy as described by Wagenaar et al. (50). Shortly, bacteria were mixed with agarose MP (Boehringer) of 37° C to a concentration of 0.7%. The mixture was immediately cooled on ice. Upon gelifying, samples were cut into 1 to 1.5 mm slices and incubated in a fixative containing 0.8% glutaraldehyde and 0.8% osmiumtetroxide. Subsequently, the samples were fixed and stained with uranyl acetate by microwave stimulation, dehydrated and imbedded in eponaraldite resin. Ultra-thin sections were counterstained with lead citrate and examined with a Philips CM 10 electron microscope at 80 kV.

Isolation of porcine alveolar macrophages (AM). Porcine AM were obtained from the lungs of specific pathogen free (SPF) pigs. Lung lavage samples were collected as described by van Leengoed et al. (43). Cells were suspended in EMEM containing 6% (v/v) SPF-pig serum and adjusted to 10⁷ cells per ml.

RESULTS**Identification of the *cps* locus.**

The *cps* locus of *S. suis* type 2 was identified by making use of a strategy developed for the genetic identification of exported proteins (13, 31). In this system we made use of a plasmid (pPHOS2) containing a truncated alkaline phosphatase gene (13). The gene lacked the promoter sequence, the translational start site and the signal sequence. The truncated gene is preceded by a unique *Sma*I restriction site. Chromosomal DNA of *S. suis* type 2, digested with *Alu*I, was randomly cloned in this restriction site. Because translocation of *PhoA* across the cytoplasmic membrane of *E. coli* is required for enzymatic activity, the system can be used to select for *S. suis* fragments containing a promoter sequence, a translational start site and a functional signal sequence. Among 560 individual *E. coli* clones tested, 16 displayed a dark blue phenotype when plated on media containing BCIP. DNA sequence analysis of the inserts from several of these plasmids were performed (results not shown) and the deduced amino acid sequences were analyzed. The hydrophobicity profile of one of the clones (pPHOS7, results not shown) showed that the N-terminal part of the sequence resembled the characteristics of a typical signal peptide: a short hydrophilic N-terminal region is followed by a hydrophobic region of 38 amino acids. These data indicate that the *phoA* system was successfully used for the selection of *S. suis* genes encoding exported proteins. Moreover, the sequences were analyzed for similarities present in the databases. The sequence of pPHOS7 showed a high similarity (37% identity) with the protein encoded by the *cps14C* gene of *Streptococcus pneumoniae* (19). This strongly suggests that pPHOS7 contains a part of the *cps* operon of *S. suis* type 2.

Cloning of the flanking *cps* genes. In order to clone the flanking *cps* genes of *S. suis* type 2 the insert of pPHOS7 was used as a probe to identify chromosomal DNA fragments which contain flanking *cps* genes. A 6-kb *Hind*III fragment was

WO 00/05378

PCT/NL99/00460

identified and cloned in pKUN19. This yielded clone pCPS6 (Fig. 1C). Sequence analysis of the insert of pCPS6 revealed that pCPS6 most probably contained the 5'-end of the *cps* locus, but still lacked the 3'-end. Therefore, sequences of the 3' -end of pCPS6 were in turn used as a probe to identify chromosomal fragments containing *cps* sequences located further downstream. These fragments were also cloned in pKUN19, resulting in pCPS17. Using the same system of chromosomal walking we subsequently generated the plasmid pCPS18, pCPS20, pCPS23 and pCPS26, containing downstream *cps* sequences.

Analysis of the *cps* operon. The complete nucleotide sequence of the cloned fragments was determined (figure 4). Examination of the compiled sequence revealed the presence of at least 13 potential open reading frame (Orfs), which were designated as Orf 2Y, Orf2X and Cps2A-Cps2K (Fig. 1A). Moreover, a 14th, incomplete, Orf (Orf 2Z) was located at the 5'-end of the sequence. Two potential promoter sequences were identified. One was located 313 bp (locations 1885-1865 and 1884-1889) upstream of Orf2X. The other potential promoter sequence was located 68 bp upstream of Orf2Y (locations 2241-2236 and 2216-2211). Orf2Y is expressed in opposite orientation. Between Orfs 2Y and 2Z the sequence contained a potential stem-loop structure, which could act as a transcription terminator. Each Orf is preceded by a ribosome-binding site and the majority of the Orfs are very closely linked. The only significant intergenic gap was found between Cps2G and Cps2H (389 nucleotides). However, no obvious promoter sequences or potential stem-loop structures were found in this region. These data suggest that Orf2X and Cps2A-Cps2K are arranged as an operon.

An overview of all Orfs with their properties is shown in Table 2. The majority of the predicted gene products is related to proteins involved in polysaccharide biosynthesis. Orf2Z showed some similarity with the YitS protein of *Bacillus subtilis*. YitS was identified during the sequence analysis of the complete genome of *B. subtilis*. The function of the protein

is unknown.

Orf2Y showed similarity with YcxD protein of *B. subtilis* (53). Based on the similarity between YcxD and MocR of *Rhizobium meliloti* (33), YcxD was suggested to be a regulatory protein.

Orf2X showed similarity with the hypothetical YAAA proteins of *Haemophilus influenzae* and *E. coli*. The function of these proteins is unknown.

The gene products encoded by the *cps2A*, *cps2B*, *cps2C* and *cps2D* genes showed approximate similarity with the CpsA, CpsC, CpsD and CpsB proteins of several serotypes of *Streptococcus pneumoniae* (19), respectively. This suggest similar functions for these proteins. Hence, Cps2A may have a role in the regulation of the capsular polysaccharide synthesis. Cps2B and Cps2C could be involved in the chain length determination of the type 2 capsule and Cps2C can play an additional role in the export of the polysaccharide. The Cps2D protein of *S. suis* is related to the CpsB protein of *S. pneumoniae* and to proteins encoded by genes of several other Gram-positive bacteria involved in polysaccharide or exopolysaccharide synthesis, but their function is unknown (19).

The protein encoded by *cps2E* gene showed similarity to several bacterial proteins with glycosyl transferase activities: Cps14E and Cps19fE of *S. pneumoniae* serotypes 14 and 19F (18, 19, 29), CpsE of *Streptococcus salvarius* (X94980) and CpsD of *Streptococcus agalactiae* (34). Recently, Kolkman et al. (18) showed that Cps14E is a glucosyl-1-phosphate transferase that links glucose to a lipid carrier, the first step in the biosynthesis of the *S. pneumoniae* type 14 repeating unit. Based on these data a similar function may be fulfilled by Cps2E of *S. suis*.

The protein encoded by the *cps2F* gene showed similarity to the protein encoded by the *rfbU* gene of *Salmonella enteritica*. (25). This similarity is most pronounced in the C-terminal regions of these proteins. The *rfbU* gene was shown to encoded mannosyltransferase activity (25).

WO 00/05378

PCT/NL99/00460

The *cps2G* gene encoded a protein that showed moderate similarity with the *rfbF* gene product of *Campylobacter hyoilei* (22), the *epsF* gene product of *S. thermophilus* (40) and the *capM* gene product of *S. aureus* (24). On the basis of
5 similarity the *rfbF*, *epsF* and *capM* genes are suggested to encoded galactosyltransferase activities. Hence, a similar glycosyl transferase activity could be fulfilled by the *cps2G* gene product.

The *cps2H* gene encodes a protein that is similar to the N-
10 terminal region of the *lgtD* gene product of *Haemophilus influenzae* (U32768). Moreover, the hydrophobicity plots of Cps2H and LgtD looked very similar in these regions (data not shown). Based on sequence similarity the *lgtD* gene product was suggested to have glycosyl transferase activity (U32768).

15 The gene product encoded by the *cps2I* gene showed some similarity with a protein of *Actinobacillus actinomycetemcomitans* (AB002668). This protein is part of the gene cluster responsible for the serotype-b-specific antigen of *A. actinomycetemcomitans*. The function of the protein is unknown.

20 The gene products encoded by the *cps2J* and *cps2K* genes showed significant similarities to the Cps14J protein of *S. pneumoniae*. The *cps14J* gene of *S. pneumoniae* was shown to encode a β -1,4-galactosyltransferase activity. In *S. pneumoniae* CpsJ is responsible for the addition of the fourth
25 (i.e. last) sugar in the synthesis of the *S. pneumoniae* serotype 14 polysaccharide (20). Even some similarity was found between Cps2J and Cps2K (Fig. 2, 25.5% similarity). This similarity was most pronounced in the N-terminal regions of the proteins. Recently, two small conserved regions were identified
30 in the N-terminus of Cps14J and Cps14I and their homologues (20). These regions were predicted to be important for catalytic activity. Both regions, DXS and DXDD (Fig. 2), were also found in Cps2J and Cps2K.

Distribution of the *cps2* genes in other *S. suis* serotypes. To examine the relationship between the *cps2* genes and *cps* genes in the other *S. suis* serotypes, we performed cross-hybridization experiments. DNA fragments of the individual *cps2* genes were amplified by PCR, labelled with ^{32}P , and used to probe Southern blots of chromosomal DNA of the reference strains of the 35 different *S. suis* serotypes. Large variation in the hybridization patterns were observed (Table 4). As a positive control we used a probe specific for 16S rRNA. The 16S rRNA probe hybridized with all serotypes tested. However, none of the other genes tested were common in all serotypes. Based on the genetic organization of the genes we previously suggested that *orfX* and *cpsA-cpsK* genes are part of one operon and that the protein encoded by these genes are all involved in polysaccharide biosynthesis. *OrfY* and *OrfZ* are not a part of this operon, and their role in the polysaccharide biosynthesis is unclear. Based on sequence similarity data, *OrfY* may be involved in regulation of the *cps2* genes. *OrfZ* is proposed to be unrelated to polysaccharide biosynthesis.

Probes specific for the *orfZ*, *orfY*, *orfX*, *cpsA*, *cpsB*, *cpsC* and *cpsD* genes hybridized with most other serotypes. This suggests that the protein encoded by these genes are not type-specific, but may perform more common functions in biosynthesis of the capsular polysaccharide. This confirms previous data which showed that the *cps2A-cps2D* genes showed strong similarity to *cps* genes of several serotype of *Streptococcus pneumoniae*. Based on this similarity *Cps2A* is possibly a regulatory protein, whereas *Cps2B* and *Cps2C* may play a role in length determination and export of polysaccharide. The *cps2E* gene hybridized with DNA of serotypes 1, 2, 14 and 1/2. The *cps2E* gene showed a strong similarity to the *cps14E* gene of *S. pneumoniae* (18). This enzyme was shown to have a glucosyl-1-phosphate activity and catalyzed the transfer of glucose to a lipid carrier (18). These data indicate that a glycosyltransferase closely related to *Cps14E* may be responsible for the first step in the biosynthesis of

WO 00/05378

PCT/NL99/00460

polysaccharide in the *S. suis* serotypes 1, 2, 14 and 1/2. The *cps2F*, *cps2G*, *cps2H*, *cps2I* and *cps2J* genes hybridized with chromosomal DNA of serotypes 2 and 1/2 only. The *cps2G* gene showed an additional weak hybridization signal with DNA of serotype 34. In agglutination tests serotype 1/2 showed agglutination with sera specific for serotype 2 as well as with sera specific for serotype 1. This suggests that serotype 1/2 shares antigenic determinants with both types 1 and 2. The hybridization data confirmed these data. All putative glycosyltransferases present in serotype 2 are also present in serotype 1/2. The *cps2K* gene showed a similar hybridization pattern as the *cps2E* gene. Hybridization was observed with DNA of serotypes 1, 2, 14 and 1/2. Taken together these hybridization data show that the *cps2* gene cluster can be divided in three regions: a central region containing the type-specific genes is flanked by two regions containing common genes for various serotypes.

Cloning of the type-specific *cps* genes of serotypes 1 and 9.

To clone the type-specific *cps* genes of *S. suis* serotype 1 we used the *cps2E* gene as a probe to identify chromosomal DNA fragments of type 1 which contain flanking *cps* genes. A 5 kb *EcoRV* fragment was identified and cloned in pKUN19. This yielded pCPS1-1 (Fig. 1B). This fragment was in turn used as a probe to identify an overlapping 2.2 kb *HindIII* fragment. pKUN19 containing this *HindIII* fragment was designated pCPS1-2. The same strategy was followed to identify and clone the type-specific *cps* genes of serotype 9. In this case, we used the *cps2D* gene as a probe. A 0.8 kb *HindIII*-*XbaI* fragment was identified and cloned, yielding pCPS9-1 (Fig. 1C). This fragment was in turn used as a probe to identify a 4 kb *XbaI* fragment. pKUN19 containing this 4 kb *XbaI* fragment was designated pCPS9-2.

Analysis of the cloned *cps1* genes. The complete nucleotide sequence of the inserts of pCPS1-1 and pCPS1-2 was determined (figure 5). Examination of the sequence revealed the presence of five complete and two incomplete Orfs (Fig.1B). Each Orf is preceded by a ribosome-binding site. In accord with data obtained for the *cps2* genes of serotype 2, the majority of the Orfs is very closely linked. The only significant gap (718 bp) was found between Cps1G and Cps1H. No obvious promoter sequences or potential stem-loop structures could be found in this region. This suggests that, as in serotype 2, the *cps* genes in serotype 1 are arranged in an operon.

An overview of the Orfs and their properties is shown in Table 2. As expected on the basis of the hybridization data (Table 4), the protein encoded by the *cps1E* gene was related to Cps2E of *S. suis* type 2 (identity of 86%). The fragment cloned in pCPS1-1 lacked the coding region for the first 7 amino acids of the *cps1E* gene.

The protein encoded by the *cps1F* and *cps1G* genes showed strong similarity to the Cps14F and Cps14G proteins of *Streptococcus pneumoniae* serotype 14, respectively (20). The function of the Cps14F is not completely clear, but it has been suggested that Cps14F can enhance role in glycosyltransferase activity. The *cps14G* gene of *S. pneumoniae* was shown to encode β -1,4-galactosyltransferase activity. In *S. pneumoniae* type 14 this activity is required for the second step in the biosynthesis of the oligosaccharide subunit (20). Based on the similarity data found similar glycosyltransferase and enhancing activities are suggested for the *cps 1G* and *cps1F* genes of *S. suis* type 1.

The protein encoded by the *cps1H* gene showed similarity to the Cps14H protein of *S. pneumoniae* (20). Based on sequence similarity Cps14H was proposed to be the polysaccharide polymerase (20).

The protein encoded by the *cps1I* gene showed some similarity with the Cps14J protein of *S. pneumoniae* (19). The *cps14J* gene was shown to encode a β -1,4-galactosyltransferase

WO 00/05378

PCT/NL99/00460

activity, responsible for the addition of the fourth (i.e. last) sugar in the synthesis of the *S. pneumoniae* serotype 14 polysaccharide.

Between Cps1G and Cps1H a gap of 718 bp was found. This region revealed three small Orfs. The three Orfs were expressed in three different reading frames and were not preceded by potential ribosome binding sites, nor contained potential start sites. However, the three potential gene products encoded by this region showed some similarity with three successive regions of the C-terminal part of the EpsK protein of *Streptococcus thermophilus* (27% identity, 40). The region related to the first 82 amino acids is lacking.

Analysis of the cloned *cps9* genes. We also determined the complete nucleotide sequence of the inserts of pCPS9-1 and pCPS9-2 (figure 6). Examination of the sequence revealed the presence of three complete and two incomplete Orfs (Fig.1C). As in serotypes 1 and 2, all Orfs are preceded by a ribosome-binding site and are very closely coupled. As suggested by the hybridization data (Table 4) the Cps2D and Cps9D proteins were highly related (Table 2). Based on sequence comparisons pCPS9-1 lacked the first 27 amino acids of the Cps9D protein.

The protein encoded by the *cps9E* gene showed some similarity with the CapD protein of *Staphylococcus aureus* serotype 1 (24). Based on sequence similarity data the Cap1D protein was suggested to be an epimerase or a dehydratase involved in the synthesis of N-acetylfructosamine or N-acetylgalactosamine (63).

Cps9F showed some similarity to the CapM proteins of *S. aureus* serotypes 5 and 8 (61, 64, 65). Based on sequence similarity data Cap5M and Cap8M are proposed to be glycosyltransferases (63).

The protein encoded by the *cps9G* gene showed some similarity with a protein of *Actinobacillus actinomycetemcomitans* (AB002668_4). This protein is part of a gene cluster responsible for the serotype-b specific antigens

WO 00/05378

PCT/NL99/00460

of *Actinobacillus actinomycetemcomitans*. The function of the protein is unknown.

The protein encoded by the *cps9H* gene showed some similarity with the *rfbB* gene of *Yersinia enterocolitica* (68).
5 The RfbB protein was shown to be essential for O-antigen synthesis, but the function of the protein in the synthesis of the O:3 lipopolysaccharide is unknown.

Serotype 1 and serotype 9 specific cps genes. To determine
10 whether the cloned fragments in pCPS1-1, pCPS1-2, pCPS9-1 and pCPS9-2 contained the type-specific genes for serotype 1 and 9, respectively, cross hybridization experiments were performed. DNA fragments of the individual *cps1* and *cps9* genes were amplified by PCR, labelled with ³²P, and used to probe
15 Southern blots of chromosomal DNA of the reference strains of the 35 different *S. suis* serotypes. The results are shown in Table 5. Based on the data obtained with the *cps2E* probe (Table 4), the *cps1E* probe was expected to hybridize with chromosomal DNA of *S. suis* serotypes 1,2, 14, 27 and 1/2. The
20 *cps1H*, *cps9E* and *cps9F* probes hybridized with most other serotypes. However, the *cps1F* and *cps1G* and *cps1I* probes hybridized with chromosomal DNA of serotypes 1 and 14 only. The *cps9G* and *cps9H* probe hybridized with serotype 9 only. These data suggest that the *cps9G* and *cps9H* probes are
25 specific for serotype 9 and therefore could be useful tools for the development of rapid and sensitive diagnostic tests for *S. suis* type 9 infections.

Type specific PCR. So far, the probes were tested on the 35
30 different reference strains only. To test the diagnostic value of the type-specific *cps* probes further, several other *S. suis* serotype 1, 2, 1/2, 9 and 14 strains were used. Moreover, since a PCR based method would be even more rapid and sensitive than a hybridization test, we tested whether we
35 could use a PCR for the serotyping of the *S. suis* strains. The

WO 00/05378

PCT/NL99/00460

oligonucleotide primer sets were chosen within the *cps2J*,
cps1I and *cps9H* genes. Amplified fragments of 675 bp, 380 bp
and 390 bp were expected respectively. The results show that
5 using *cps2J* primers; 380 bp fragments were amplified on type 1
and 14 strains using *cps1I* primers and 390 bp fragments were
amplified on type 9 strains using *cps9H* primers.

Construction of mutants impaired in capsule production. To
10 evaluate the role of the capsule of *S. suis* type 2 in the
pathogenesis, we constructed two isogenic mutants in which
capsule production was disturbed. To construct mutant 10cpsB,
pCPS11 was used. In this plasmid a part of the *cps2B* gene was
replaced by the spectinomycin-resistance gene. To construct
15 mutant strain 10cpsEF the plasmid pCPS28 was used. In pCPS28
the 3'-end of *cps2E* gene as well as the 5'-end of *cps2F* gene
were replaced by the spectinomycin-resistance gene. pCPS11 and
pCPS28 were used to electrotransform strain 10 of *S. suis* type
2 and spectinomycin-resistant colonies were selected. Southern
20 blotting and hybridization experiments were used to select
double cross over integration events (results not shown).
To test whether the capsular structure of the strains 10cpsB
and 10cpsEF was disturbed, we used a slide agglutination test
using a suspension of the mutant strains in hyperimmune anti-*S.*
25 *suis* type 2 serum (44). The results showed that even in the
absence of serotype specific antisera, the bacteria
agglutinated. This indicates that in the mutant strains the
capsular structure was disturbed. To confirm this, thin
sections of wild type and mutant strains were compared by
30 electron microscopy. The results showed that compared to the
wild type (Fig. 3A) the amount of capsule produced by the
mutant strains was greatly reduced (Figs. 3B and 3C). Almost no
capsular material could be detected on the surface of the
mutant strains.

Capsular mutants are sensitive to phagocytosis and killing by porcine alveolar macrophages (PAM) .

The capsular mutants were tested for their ability to resist phagocytosis by PAM in the presence of porcine SPF serum. The wild type strain 10 seemed to be resistant to phagocytosis under these conditions (Fig. 4A). In contrast, the mutant strains were efficiently ingested by macrophages (Fig. 4A). After 90 min. more than 99.7% (strain 10cpsB) and 99.8% (strain 10cpsEF) of the mutant cells were ingested by the macrophages. Moreover, as shown in Fig. 4B the ingested strains were efficiently killed by the macrophages. 90-98% of all ingested cells were killed within 90 min. No differences could be observed between wild type and mutant strains. These data indicate that the capsule of *S. suis* type 2 efficiently protects the organism from uptake by macrophages *in vitro*.

Capsular mutants are less virulent for germfree piglets. The virulence properties of the wild-type and mutant strains were tested after experimental infection of newborn germfree pigs (45, 49). Table 1 shows that specific and nonspecific signs of disease could be observed in all pigs inoculated with the wild type strain. Moreover, all pigs inoculated with the wild type strain died during the course of the experiment or were killed because of serious illness or nervous disorders (Table 3). In contrast, the pigs inoculated with strains 10cpsB and 10cpsEF showed no specific signs of disease and all pigs survived until the end of the experiment. The temperature of the pigs inoculated with the wild type strain increased 2 days after inoculation and remained high until day 5 (Table 3). The temperature of the pigs inoculated with the mutant strains sometimes exceeded the 40°C, however, we could observe significant differences in the fever index [i.e % of observations in an experimental group during which pigs showed fever (>40°C)] between pigs inoculated with wild type and mutant strains. All pigs showed increased numbers of polymorphonuclear leucocytes (PMLs) (>10 x 10⁹ PMLs per litre)

WO 00/05378

PCT/NL99/00460

(Table 3). However, in pigs inoculated with the mutant strains the percentage of samples with increased numbers of PMLs was considerably lower. *S. suis* strains and *B. bronchiseptica* could be isolated from the nasopharynx and feces swab samples of all pigs from 1 day post-infection until the end of the experiment (Table 3). Postmortem, the wild type strain could frequently be isolated from the central nervous system (CNS), kidney, heart, liver, spleen, serosae, joints and tonsils. Mutant strains could easily be recovered from the tonsils, but were never recovered from the kidney, liver or spleen. Interestingly, low numbers of the mutant strains were isolated from the CNS, the serosae, the joints, the lungs and the heart. Taken together, these data strongly indicated that mutant *S. suis* strains, impaired in capsule production, are not virulent for young germfree pigs.

We describe the identification and the molecular characterisation of the *cps* locus, involved in the capsular polysaccharide biosynthesis, of *S. suis*. Most of the genes seemed to belong to a single transcriptional unit, suggesting a co-ordinate control of these genes. We assign functions to most of the gene products. We thereby identified regions involved in regulation (Cps2A), chain length determination (Cps2B, C), export (Cps2C) and biosynthesis (Cps2E, F, G, H, J, K). The region involved in biosynthesis is located at the centre of the gene cluster and is flanked by two regions containing genes with more common functions. The incomplete *orf2Z* gene was located at the 5'-end of the cloned fragment. Orf2Z showed some similarity with the YitS protein of *B. subtilis*. However, because the function of the YitS protein is unknown this did not give us any information about the possible function of Orf2Z. Because the *orf2Z* gene is not a part of the *cps* operon, a role of this gene in polysaccharide biosynthesis is not expected. The Orf2Y protein showed some similarity with the YcxD protein of *B. subtilis* (53). The YcxD protein was suggested to be a regulatory protein. Similarly, Orf2Y may be involved in the regulation of polysaccharide biosynthesis. The Orf2X

protein showed similarity with the YAAA proteins of *H. influenzae* and *E. coli*. The function of these proteins is unknown. In *S. suis* type 2 the *orf2X* gene seemed to be the first gene in the *cps2* operon. This suggests a role of Orf2X in the polysaccharide biosynthesis. In *H. influenzae* and *E. coli*, however, these proteins are not associated with capsular gene clusters. The analysis of isogenic mutants impaired in the expression of Orf2X should give more insight in the presumed role of Orf2X in the polysaccharide biosynthesis of *S. suis* type 2.

The gene products encoded by the *cps2E*, *cps2F*, *cps2G*, *cps2H*, *cps2J* and *cps2K* genes showed little similarity with glycosyltransferases of several Gram-positive or Gram-negative bacteria (18, 19, 20, 22, 25). The *cps2E* gene product shows some similarity with the Cps14E protein of *S. pneumoniae* (18, 19). Cps14E is a glucosyl-1-phosphate transferase that links glucose to a lipid carrier (18). In *S. pneumoniae* this is the first step in the biosynthesis of the oligosaccharide repeating unit. The structure of the *S. suis* serotype 2 capsule contains glucose, galactose, rhamnose, N-acetyl glucoseamine and sialic acid in a ratio of 3:1:1:1:1 (7). Based on these data we conclude that Cps2E of *S. suis* has glycosyltransferase activity, and is involved in the linkage of the first sugar to the lipid carrier.

The C-terminal region of the *cps2F* gene product showed some similarity with the RfbU of *Salmonella enteritica*. RfbU was shown to have mannosyltransferase activity (24). Because mannosyl is not a component of the *S. suis* type 2 polysaccharide a mannosyltransferase activity is not expected in this organism. Nevertheless, *cps2F* encodes a glycosyltransferase with another sugar specificity.

Cps2G showed moderate similarity to a family of gene products suggested to encode galactosyltransferase activities (22, 24, 40). Hence a similar activity is shown for Cps2G.

Cps2H showed some similarity with LgtD of *H. influenzae* (U32768). Because LgtD was proposed to have glycosyltransferase

WO 00/05378

PCT/NL99/00460

activity , a similar activity is fulfilled by Cps2H.

Cps2J and Cps2K showed similarity to Cps14J of *S. pneumoniae* (20). Cps2J showed similarity with Cps14I of *S. pneumoniae* as well. Cps14I was shown to have N-acetyl glucosaminyltransferase activity, whereas Cps14J has a β -1,4-galactosyltransferase activity (20). In *S. pneumoniae* Cps14I is responsible for the addition of the third sugar and Cps14J for the addition of the last sugar in the synthesis of the type 14 repeating unit (20). Because the capsule of *S. suis* type 2 contains galactose as well as N-acetyl glucosamine components, galactosyltransferase as well as N-acetyl glucoaminyltransferase activities could be envisaged for the cps2J and cps2K gene products, respectively. As was observed for Cps14I and Cps14J, the N-termini of Cps2J and Cps2K showed a significant degree of sequence similarity. Within the N-terminal domains of Cps14I and Cps14J, two small regions were identified, which were also conserved in several other glycosyltransferases (22). Within these two regions, two Asp residues were proposed to be important for catalytic activity. The two conserved regions, DXS and DXDD, were also found in Cps2J and Cps2K.

The function of Cps2I remains unclear. Cps2I showed some similarity with a protein of *A. actinomycetemcomitans*. Although this protein part is of the gene cluster responsible for the serotype-B-specific antigens, the function of the protein is unknown.

We further describe the identification and characterization of the cps genes specific for *S. suis* serotypes 1, 2 and 9. After the entire cps2 locus of *S. suis* serotype 2 was cloned and characterized, functions for most of the cps2 gene products could be assigned by sequence homologies. Based on these data the glycosyltransferase activities, required for type specificity, could be located in the centre of the operon. Cross-hybridization experiments, using the individual cps2 genes as probes on chromosomal DNAs of the 35 different serotypes, confirmed this idea. The regions containing the

type-specific genes of serotypes 1 and 9 could be cloned and characterized, showing that an identical genetic organization of the *cps* operons of other *S. suis* serotypes exists. The *cps1E*, *cps1F*, *cps1G*, *cps1H*, and *cps1I* genes revealed a striking similarity with *cps14E*, *cps14F*, *cps14G*, *cps14H* and *cps14J* genes of *S. pneumoniae*. Interestingly, *S. pneumoniae* serotype 14 is the serotype most commonly associated with pneumococcal infections in young children (54), whereas *S. suis* serotype 1 strains are most commonly isolated from piglets younger than 8 weeks (46). In *S. pneumoniae* the *cps14E*, *cps14G*, *cps14I* and *cps14J* encode the glycosyltransferases required for the synthesis of the type 14 tetrameric repeating unit, showing that the *cps1E*, *cps1G* and *cps1I* genes encoded glycosyltransferases. The precise functions of these genes as well as the substrate specificities of the enzymes can be established. In *S. pneumoniae* the *cps14E* gene was shown to encode a glucosyl-1-phosphate transferase catalyzing the transfer of glucose to a lipid carrier. Moreover, *cpsE*-like genes were found in *S. pneumoniae* serotypes 9N, 13, 14, 15B, 15C, 18F, 18A and 19F (60). *CpsE* mutants were constructed in the serotypes 9N, 13, 14 and 15B. All mutant strains lacked glucosyltransferase activity (60). Moreover, in all these *S. pneumoniae* serotypes the *cpsE* gene seemed to be responsible for the addition of glucose to the lipid carrier. Based on these data we suggest that in *S. suis* type 1 the *cps1E* gene may fulfil a similar function. The structure of the *S. suis* type 1 capsule is unknown, but it is composed of glucose, galactose, N-acetyl glucosamine, N-acetyl galactosamine and sialic acid in a ratio of 1: 2.4: 1: 1:1.4 (5). Therefore a role of a *cpsE*-like glucosyltransferase activity can easily be envisaged. *CpsE* like sequences were also found in serotypes 2, 1/2 and 14.

For polysaccharide biosynthesis in *S. pneumoniae* type 14, transfer of the second sugar of the repeating unit to the first lipid-linked sugar is performed by the gene products of *cps14F* and *cps14G* (20). Similar to *Cps14F* and *Cps14G*, the *S.*

WO 00/05378

PCT/NL99/00460

suis type 1 proteins Cps1F and Cps1G may act as one glycosyltransferase performing the same reaction. Cps14F and Cps14G of *S. pneumoniae* showed similarity to the N-terminal half and C-terminal half of the SpsK protein of *Sphingomonas* (20, 67), respectively. This suggests a combined function for both proteins. Moreover, *cps14F* and *cps14G* like sequences were found in several serotypes of *S. pneumoniae* and these genes always seemed to exist together (60). The same was observed for *S. suis* type 1. The *cps1F* and *cps1G* probes hybridized with type 1 and type 14 strains.

According to the similarity found between the *cps1H* gene and the *cps14H* gene of *S. pneumoniae* (20), *cps1H* is expected to encode a polysaccharide polymerase.

The protein encoded by the *cps1I* gene showed some similarity with the Cps14J protein of *S. pneumoniae* (19). The *cps14J* gene was shown to encode a β -1,4-galactosyltransferase activity, responsible for the addition of the fourth (i.e. last) sugar in the synthesis of the *S. pneumoniae* serotype 14 polysaccharide. In *S. suis* type 2 the proteins encoded by the *cps2J* and *cps2K* genes showed similarity to the Cps14J protein. However, no significant homologies were found between Cps2J, Cps2K and Cps1I. In the N-terminal regions of Cps14J and Cps14I two small conserved regions, DXS and DXDD, were identified (19). These regions seemed to be important for catalytic activity (13). At the same positions in the sequence Cps2I contained the regions DXS and DXED.

In the region between Cps1G and Cps1H three small Orfs were identified. Since the Orfs were expressed in three different reading frames, and did not contain potential start sites, expression is not expected. However, the three potential gene products encoded by this region showed some similarity with three successive regions of the C-terminal part of the EpsK protein of *Streptococcus thermophilus* (27% identity, 40). The region related to the first 82 amino acids is lacking. The EpsK protein was suggested to play a role in the export of the exopolysaccharide by rendering the polymerized

exopolysaccharide more hydrophobic through a lipid modification. These data could suggest that the sequences in the region between *Cps1G* and *Cps1H* originated from *epsK*-like sequence. Hybridization experiments showed that this *epsK*-like region is also present in other serotype 1 strains as well as in serotype 14 strains (results not shown).

The function of most of the cloned serotype 9 genes can be established. Based on sequence similarity data the *cps9E* and *cps9F* genes could be glycosyltransferases (61, 24, 63, 64, 65). Moreover, the *cps9G* and *cps9H* genes showed similarity to genes located in regions involved in polysaccharide biosynthesis, but the function of these genes is unknown (68).

Cross-hybridization experiments using the individual *cps2*, *cps1* and *cps9* genes as probes showed that the *cps9G* and *cps9H* probes specifically hybridized with serotype 9 strains. Therefore, these are useful as tools for the identification of *S. suis* type 9 strains both for diagnostic purposes as well as in epidemiological and transmission studies. We previously developed a PCR method which can be used to detect *S. suis* strains in nasal and tonsil swabs of pigs (62). The method was for example used to identify pathogenic (EF-positive) strains of *S. suis* serotype 2. During the last years, beside *S. suis* type 2 strains, serotype 9 strains are frequently isolated from organs of diseased pigs. However, until now a rapid and sensitive diagnostic test was not available for type 9 strains. Therefore, the type 9 specific probes or the type 9 specific PCR is of great diagnostic value. The *cps1F*, *cps1G* and *cps1I* probes hybridized with serotype 1 as well as with serotype 14 strains. In coagglutination tests type 1 strains react with the anti-type 1 as well as with the anti-type 14 antisera (56). This suggests the presence of common epitopes between these serotypes. On the other hand type 1 strains agglutinated only with anti-type 1 serum (56,57), indicating that it is possible to detect differences between those serotypes.

The *cps2F*, *cps2G*, *cps2H*, *cps2I* and *cps2J* probes hybridized

WO 00/05378

PCT/NL99/00460

with serotypes 2 and 1/2 only. Serotype 34 showed a weak hybridizing signal with the *cps2G* probe. As shown in agglutination tests type 1/2 strains react with sera directed against type 1 as well as with sera directed against type 2 strains (46). Therefore, type 1/2 shared antigens with both types 1 and 2. Based on the hybridization patterns of serotype 1/2 strains with the *cps1* and *cps2* specific genes, serotype 1/2 seemed to be more closely related to type 2 strains than to type 1 strains. In our current studies we identify type-specific genes, primers or probes which are used for the discrimination of serotypes 1, 14 and 2 and 1/2 and others of the 35 serotypes yet known. Furthermore, type-specific genes, primers or probes can now easily be developed for yet unknown serotypes, once they become isolated.

Cloning and characterization of a further part of the *cps2* locus.

Based on the established sequence 11 genes, designated *cps2L* to *cps2T*, *orf2U* and *orf2V*, were identified. A gene homologous to genes involved in the polymerization of the repeating oligosaccharide unit (*cps2O*) as well as genes involved in the synthesis of sialic acid (*cps2P* to *cps2T*) were identified. Moreover, hybridization experiments showed that the genes involved in the sialic acid synthesis are present in *S. suis* serotype 1, 2, 14, 27 and 1/2. The "*cps2M*" and "*cps2N*" regions showed similarity to proteins involved in the polysaccharide biosynthesis of other gram-positive bacteria. However, these regions seemed to be truncated or were non-functional as the result of frame-shift or point mutations. At its 3'-end the *cps2* locus contained two insertional elements ("*orf2U*" and "*orf2V*") both of which seemed to be non-functional.

To clone the remaining part of the *cps2* locus, sequences of the 3'-end of pCPS26 (Fig. 1C) were used to identify a chromosomal fragment containing *cps2* sequences located further downstream. This fragment was cloned in pKUN19 resulting in pCPS29. Using a similar approach we subsequently isolated the

WO 00/05378

PCT/NL99/00460

plasmids pCPS30 and pCPS34 containing downstream cps2 sequences (Fig. 1C).

Analysis of the cps2 operon.

5 The complete nucleotide sequence of the cloned fragments was determined. Examination of the compiled sequence revealed the presence of : a sequence encoding the C-terminal part of Cps2K, six apparently functional genes (designated cps20-
10 cps2T) and the remnants of 5 different ancestral genes (designated "cps2L", "cps2M", "cps2N" , "orf2U" and "orf2V"). The latter genes seemed to be truncated or incomplete as the result of the presence of stop codons or frame-shift mutations (Fig. 1A). Neither potential promoter sequences nor potential stem-loop structures could be identified within the sequenced
15 region. A ribosome-binding site precedes each ORF and the majority of the ORFs is very closely linked. Three intergenic gaps were found: one between "cps2M" and "cps2N" (176 nucleotides), one between cps20 and cps2P (525 nucleotides), and one between cps2T and "orf2U" (200 nucleotides). These and
20 our above data show that Orf2X and Cps2A-Orf2T are part of a single operon.

 A list of all loci and their properties is shown in Table 4. The "cps2L" region contained three potential ORFs, of 103, 79 and 152 amino acids, respectively, which were only
25 separated from each other by stop codons. Only the first ORF is preceded by a potential ribosomal binding site and contained a methionine start codon. This suggests that "cps2L" originates from an ancestral cps2L gene, which coded for a protein of 339 amino acids. The function of this hypothetical
30 Cps2L protein remains unclear so far: no significant homologies were found between Cps2L and proteins present in the data libraries. It is not clear whether the first ORF of the "cps2L" region is expressed into a protein of 103 amino acids. The "cps2M " region showed homology to the N-terminal
35 134 amino acids of the NeuA proteins of Streptococcus agalactiae and Escherichia coli (AB017355, 32). However,

WO 00/05378

PCT/NL99/00460

although the "cps2 M" region contained a potential ribosome binding site, a methionine start codon was absent. Compared with the *S. agalactiae* sequence, the ATG start codon was replaced by a lysin encoding AAG codon. Moreover, the region homologous to the first 58 amino acids of the *S. agalactiae* NeuA (identity 77%) was separated from the region homologous to amino acids 59-134 of NeuA by a repeated DNA sequence of 100-bp (see later). In addition, the region homologous to amino acids 59 to 95 of NeuA (identity 32%) and the region homologous to the amino acids 96 to 134 of NeuA (identity 50%) were present in different reading frames. Therefore, the partial and truncated NeuA homologue is probably nonfunctional in *S. suis*. The "cps2N" region showed homology to CpsJ of *S. agalactiae* (accession no. AB017355). However, sequences homologous to the first 88 amino acids of CpsJ were lacking in *S. suis*. Moreover, the homologous region was present in two different reading frames. The protein encoded by the cps20 gene showed homology to proteins of several streptococci involved in the transport of the oligosaccharide repeating unit (accession no. AB017355), suggesting a similar function for Cps20. The proteins encoded by the cps2P, cps2S and cps2T genes showed homology to the NeuB, NeuD and NeuA proteins of *S. agalactiae* and *E. coli* (accession no AB017355). Because the "cps2M" region also showed homology to NeuA of *E. coli*, the *S. suis* cps2 locus contains a functional neuA gene (cps2T) as well as a nonfunctional ("cps2M") gene. The mutual homology between these two regions showed an identity of 77% at the amino acid level over amino acids 1-58 and 49% over the amino acids 59-134. Cps2Q and Cps2R showed homology to the N-terminal and C-terminal parts of the NeuC protein of *S. agalactiae* and *E. coli*, respectively. This suggests that the function of the *S. agalactiae* NeuC protein in *S. suis* is likely fulfilled by two different proteins. In *E. coli* the neu genes are known to be involved in the synthesis of sialic acid. NeuNAc is synthesized from N-acetylmannosamine and phosphoenolpyruvate by NeuNAc synthetase. Subsequently, NeuNAc

WO 00/05378

PCT/NL99/00460

is converted to CMP-NeuNAc by the enzyme CMP-NeuNAc synthetase. CMP-NeuNAc is the substrate for the synthesis of polysaccharide. In *E. coli* K1 NeuB is the NeuNAc synthetase, NeuA is the CMP-NeuNAc synthetase. NeuC has been implicated in the NeuNAc synthesis, but its precise role is not known. The precise role of NeuD is not known. A role of the Cps2P-Cps2T proteins in the synthesis of sialic acid can easily be envisaged, since the capsule of *S. suis* serotype 2 is rich in sialic acid. In *S. agalactiae* sialic acid has been shown to be critical to the virulence function of the type III capsule. Moreover, it has been suggested that the presence of sialic acid in capsule of bacteria which can cause meningitis may be important for the capacity of these bacteria to breach the blood-brain barrier. So far, however, the requirement of the sialic acid for virulence of *S. suis* remains unclear.

"Orf2U" and "Orf2V" showed homology to proteins located on two different insertional elements. "Orf2U" is homologous to IS1194 of *Streptococcus thermophilus*, whereas "Orf2V" showed homology to a putative transposase of *Streptococcus pneumoniae*. This putative transposase was recently found to be associated with the type 2 capsular locus of *S. pneumoniae*. Compared with the original insertional elements in *S. thermophilus* and *S. pneumoniae*, both "Orf2U" and "Orf2V" are likely to be non-functional due to frame shift mutations within their coding regions.

A striking observation was the presence of a sequence of 100 bp (Fig. 9) which was repeated three times within the cps2 operon. The sequence is highly conserved (between 94% and 98%) and was found in the intergenic regions between cps2G and cps2H, within "cps2M" and between cps2O and cps2P. No significant homologies were found between this 100-bp direct repeat sequence and sequences present in the data libraries, suggesting that the sequence is unique for *S. suis*.

Distribution of the cps2 sequences among the 35 *S. suis* serotypes. To examine the presence of sialic acid encoding genes in other *S. suis* serotypes, we performed cross-

WO 00/05378

PCT/NL99/00460

hybridization experiments. DNA fragments of the individual cps2 genes were amplified by PCR, radiolabelled with ^{32}P and hybridized to chromosomal DNA of the reference strains of the 35 different *S. suis* serotypes. As a positive control we used
5 a probe specific for *S. suis* 16S rRNA. The 16S rRNA probe hybridized with almost equal intensities to all serotypes tested (Table 4). The "cps2L" sequence hybridized with DNA of serotype 1, 2, 14 and 1/2. The "cps2M", cps2O, cps2P, cps2Q, cps2R, cps2S and cps2T genes hybridized with DNA of serotype
10 1, 2, 14, 27 and 1/2. Because the cps2P-cps2T genes are most probably involved in the synthesis of sialic acid these results suggest that sialic acid is also a part of the capsule in the *S. suis* serotype 1, 2, 14, 27 and 1/2. This is in agreement with the finding that the serotypes 1, 2 and 1/2
15 possess a capsule that is rich in sialic acid. Although the chemical compositions of the capsules of serotype 14 and 27 are unknown, recent agglutination studies using sialic acid-binding lectins suggested the presence of sialic acid in *S. suis* serotype 14, but not in serotype 27. In these studies,
20 sialic acid was also detected in serotypes 15 and 16. Since the latter observation is not in agreement with our hybridization studies, it might be that other genes, not homologous to the cps2P-cps2T genes, are responsible for the sialic acid synthesis in serotypes 15 and 16.

25 A probe based on "cps2N" sequences hybridized with DNA from serotypes 1, 2, 14 and 1/2. A probe specific for "orf2U" hybridized with serotypes 1, 2, 7, 14, 24, 27, 32, 34, and 1/2, whereas a probe specific for "orf2V" hybridized with many different serotypes. In addition, we prepared a probe specific
30 for the 100-bp direct repeat sequence. This probe hybridized with the serotypes 1, 2, 13, 14, 22, 24, 27, 29, 32, 34 and 1/2 (Table 4). To analyze the number of copies of the direct repeat sequence within the *S. suis* serotype 2 chromosome, a Southern blot hybridization and analysis was performed.
35 Therefore, chromosomal DNA of *S. suis* serotype 2 was digested with NcoI and hybridized with a ^{32}P -labelled direct repeat

WO 00/05378

PCT/NL99/00460

sequence. Only one hybridizing fragment, containing the three direct repeats present on the *cps2* locus, was found (results not shown). This indicates that the 100-bp direct repeat sequence is only associated with the *cps2* locus. In *S.*

5 *pneumoniae* a 115-bp long repeated sequence was found to be associated with the capsular genes of serotypes 1, 3, 14 and 19F. In *S. pneumoniae* this 115-bp sequence was also found in the vicinity of other genes involved in pneumococcal virulence (hyaluronidase and neuraminidase genes). A regulatory role of
10 the 115-bp sequence in co-ordinate control of these virulence-related genes was suggested.

To study the role of the capsule in resistance to phagocytosis and in virulence, we constructed two isogenic mutants in which capsule synthesis was disturbed. In 10*cpsB*,
15 the *cps2B* gene was disturbed by the insertion of an antibiotic-resistance gene, whereas in 10*cpsEF* parts of the *cps2E* and *cps2F* genes were replaced. Both mutant strains seemed to be completely unencapsulated. Because the *cps 2* genes seemed to be part of an operon polar effects cannot be
20 excluded. Therefore these data did not give any information about the role of Cps2B, Cps2E or Cps2F in the polysaccharide biosynthesis. However, the results clearly show that the capsular polysaccharide of *S. suis* type 2 is a surface component with antiphagocytic activity. In vitro wild type
25 encapsulated bacteria are ingested by phagocytes at a very low frequency, whereas the mutant unencapsulated bacteria are efficiently ingested by porcine macrophages. Within 2 hours, over 99.6% of mutant bacteria were ingested and over 92% of the ingested bacteria were killed. Intracellularly, wild type
30 as well as mutant strains seemed to be killed with the same efficiency. This suggests that the loss of capsular material is associated with loss of capacity to resist uptake by macrophages. This loss of resistance to *in vitro* phagocytosis was associated with a substantial attenuation of the virulence
35 in germfree pigs. All pigs inoculated with the mutant strains survived the experiment and did not show any specific clinical

WO 00/05378

PCT/NL99/00460

signs of disease. Only some aspecific clinical signs of disease could be observed. Moreover, mutant bacteria could be reisolated from the pigs. This supports the idea that, as in other pathogenic Streptococci, the capsule of *S. suis* acts as an important virulence factor. Transposon mutants prepared by Charland impaired in the capsule production showed a reduced virulence in pigs and mice. To construct these mutants the type 2 reference strain S735 was used. We previously showed that this strain is only weakly virulent for young pigs. Moreover, the insertion site of the transposon is unsolved so far.

As a further example herein a rapid PCT test for *Streptococcus suis* type 7 is described.

15

Recent epidemiological studies on *Streptococcus suis* infections in pigs indicated that, besides serotypes 1, 2 and 9, serotype 7 is also frequently associated with diseased animals. For the latter serotype, however, no rapid and sensitive diagnostic methods are available. This hampers prevention and control programs. Here we describe the development of a type-specific PCR test for the rapid and sensitive detection of *S. suis* serotype 7. The test is based on DNA sequences of capsular (*cps*) genes specific for serotype 7. These sequences could be identified by cross-hybridization of several individual *cps* genes with the chromosomal DNAs of 35 different *S. suis* serotypes.

Streptococcus suis is an important cause of meningitis, septicemia, arthritis and sudden death in young pigs [69,70]. It can, however, also cause meningitis in man [71]. Attempts to control the disease are still hampered by the lack of sufficient knowledge about the epidemiology of the disease and the lack of effective vaccines and sensitive diagnostics.

S. suis strains can be identified and classified by their morphological, biochemical and serological characteristics [70, 73, 74]. Serological classification is based on the

WO 00/05378

PCT/NL99/00460

presence of specific antigenic determinants. Isolated and biochemically characterized *S. suis* cells are agglutinated with a panel of specific sera. These typing methods are very laborious and time-consuming and can only be performed on isolated colonies. Moreover, it has been reported that nonspecific cross-reactions may occur among different types of *S. suis* [75, 76].

So far, 35 different serotypes have been described [7, 78, 79]. *S. suis* serotype 2 is the most prevalent type isolated from diseased pigs, followed by serotypes 9, and 1. However, recently serotype 7 strains were also frequently isolated from diseased pigs [80, 81, 82]. This suggests that infections with *S. suis* serotype 7 strains seemed to be an increasing problem. Moreover, the virulence of *S. suis* serotype 7 strains was confirmed by experimental infection of young pigs [83].

Recently, rapid and sensitive PCR assays specific for serotypes 2 (and 1/2), 1 (and 14) and 9 were developed [84]. These assays were based the *cps* loci of *S. suis* serotypes 2, 1 and 9 [84, 85]. However, until now no rapid and sensitive diagnostic test is available for *S. suis* serotype 7. Herein we describe the development of a PCR test for the rapid and sensitive detection of *S. suis* serotype 7 strains. The test is based on DNA sequences which form a part of the *cps* locus of *S. suis* serotype 7. Compared with the serological serotyping methods the PCR assay was a rapid, reliable and sensitive assay. Therefore, this test, in combination with the PCR tests which we previously developed for serotype 1, 2 and 9, will undoubtedly contribute to a more rapid and reliable diagnosis of *S. suis* and may facilitate control and eradication programs.

Materials and Methods

Bacterial strains, growth conditions and serotyping.

The bacterial strains and plasmids used in this study are listed in Table 7. The *S. suis* reference strains were obtained from M. Gottschalk, Canada. *S. suis* strains were grown in Todd-Hewitt broth (code CM189, Oxoid), and plated on Columbia agar blood base (code CM331, Oxoid) containing 6% (v/v) horse blood. *E. coli* strains were grown in Luria broth [86] and plated on Luria broth containing 1.5% (w/v) agar. If required, ampicillin was added to the plates. The *S. suis* strains were serotyped by the slide agglutination test with serotype-specific antibodies [70].

15 DNA techniques.

Routine DNA manipulations and PCR reactions were performed as described by Sambrook et al. [88]. Blotting and hybridization was performed as described previously [84,86].

20 DNA sequence analysis.

DNA sequences were determined on a 373A DNA Sequencing System (Applied Biosystems, Warrington, GB). Samples were prepared by use of a ABI/PRISM dye terminator cycle sequencing ready reaction kit (Applied Biosystems). Custom-made sequencing primers were purchased from Life Technologies. Sequencing data were assembled and analyzed using the McMollyTetra program. The BLAST program was used to search for protein sequences homologous to the deduced amino acid sequences.

30

PCR.

The primers used for the *cps7H* PCR correspond to the positions 3334-3354 and 3585-3565 in the *S. suis cps7* locus. The sequences were:

35 5'-AGCTCTAACACGAAATAAGGC-3' and 5'-GTCAAACACCCTGGATAGCCG-3'.

The reaction mixtures contained 10 mM Tris-HCl, pH 8.3; 1.5 mM

WO 00/05378

PCT/NL99/00460

MgCl₂; 50 mM KCl; 0.2 mM of each of the four deoxynucleotide triphosphates; 1 microM of each of the primers and 1U of AmpliTaq Gold DNA polymerase (Perkin Elmer Applied Biosystems, New Jersey). DNA amplification was carried out in a Perkin Elmer 9600 thermal cycler and the program consisted of an incubation for 10 min at 95°C and 30 cycles of 1 min at 95°C, 2 min at 56°C and 2 min at 72°C.

Results and discussion

Cloning of the serotype 7-specific cps genes.

To isolate the type-specific cps genes of *S. suis* serotype 7 we used the cps9E gene of serotype 9 as a probe to identify chromosomal DNA fragments of type 7 containing homologous DNA sequences [84]. A 1.6-kb PstI fragment was identified and cloned in pKUN19. This yielded pCPS7-1 (Fig. 11C). In turn, this fragment was used as a probe to identify an overlapping 2.7 kb ScaI-ClaI fragment. pGEM7 containing the latter fragment was designated pCPS7-2 (Fig. 11C).

Analysis of the cloned cps7 genes.

The complete nucleotide sequences of the inserts of pCPS7-1, pCPS7-2 were determined. Examination of the cps7 sequence revealed the presence of two complete and two incomplete open reading frames (ORFs) (Fig. 11C). All ORFs are preceded by a ribosome-binding site. In accord with the data obtained for the cps1, cps2 and cps9 genes of serotypes 1, 2 and 9, respectively, the type 7 ORFs are very closely linked to each other. The only significant intergenic gap was that found between cps7E and cps7F (443 nucleotides). No obvious promoter sequences or potential stem-loop structures were found in this region. This suggests that, as in serotype 1, 2 and 9, the cps genes in serotype 7 form part of an operon.

An overview of the ORFs and their properties is shown in Table 8. As expected on the basis of the hybridization data [84], the Cps9E and Cps7E proteins showed a high similarity

WO 00/05378

PCT/NL99/00460

(identity 99%, Table 8). Based on sequence comparisons between Cps9E and Cps7E, the PstI fragment of pCPS7-1 lacks the region encoding the first 371 codons of Cps7E. The C-terminal part of the protein encoded by the cps7F gene showed some similarity
5 with the BplG protein of *Bordetella pertussis* [88], as well as with the C-terminal part of *S. suis* Cps2E [85]. Both BplG and Cps2E were suggested to have glycosyltransferase activity and are probably involved in the linkage of the first sugar to the lipid carrier [85,88]. The protein encoded by the cps7G
10 gene showed similarity with the BlpF protein of *Bordetella pertussis* [88]. BlpF is likely to be involved in the biosynthesis of an amino sugar, suggesting a similar function for Cps7G. The protein encoded by the cps7H gene showed similarity with the WbdN protein of *E. coli* [89] as well as
15 with the N-terminal part of the Cps2K protein of *S. suis* [81]. Both WbdN and Cps2K were suggested to have glycosyltransferase activity [85, 89].

Serotype 7 specific cps genes.

20 To determine whether the cloned fragments in pCPS7-1 and pCPS7-2 contained serotype 7-specific DNA sequences, cross hybridization experiments were performed. DNA fragments of the individual cps7 genes were amplified by PCR, labelled with 32P, and used to probe spot blots of chromosomal DNA of the
25 reference strains of 35 different *S. suis* serotypes. The results are summarized in Table 9. As expected, based on the data obtained with the cps9E probe [84], the cps7E probe hybridized with chromosomal DNA of many different *S. suis* serotypes. The cps7F and cps7G probes showed hybridization
30 with chromosomal DNA of *S. suis* serotypes 4, 5, 7, 17, and 23. However, the cps7H probe hybridized with chromosomal DNA of serotype 7 only, indicating that this gene is specific for serotype 7.

WO 00/05378

PCT/NL99/00460

Type specific PCR.

We tested whether we could use PCR instead of hybridization for the typing of the *S. suis* serotype 7 strains. For that purpose we selected an oligonucleotide primer set within the cps7H gene with which an amplified fragment of 251-bp was expected. In addition, we included in our analysis several *S. suis* serotype 7 strains, other than the reference strain. These strains were obtained from different countries and were isolated from different organs (Table 7). The results show that indeed a fragment of about 250-bp was amplified with all type 7 strains used (Fig. 12B), whereas no PCR products were obtained with serotype 1, 2 and 9 strains (Fig. 12A). This suggests that the PCR test, as described here, is a rapid diagnostic tool for the identification of *S. suis* serotype 7 strains. Until now such a diagnostic test was not available for serotype 7 strains. Together with the recently developed PCR assays for serotype 1, 2, 1/2, 14 and 9, this assay may be an important diagnostic tool to detect pigs carrying serotype 2, 1/2, 1, 14, 9 and 7 strains and may facilitate control and eradication programs.

WO 00/05378

PCT/NL99/00460

TABLE 1. Bacterial strains and plasmids

strain/plasmid	relevant characteristics	source/reference
Strain		
<i>E. coli</i>		
CC118	PhoA ⁻	(28)
XL2 blue	Stratagene	
<i>E. coli</i>		
XL2 blue	Stratagene	
<i>S. suis</i>		
10	virulent serotype 2 strain	(49)
3	serotype 2	(63)
17	serotype 2	(63)
735	reference strain serotype 2	(63)
T15	serotype 2	(63)
6555	reference strain serotype 1	(63)
6388	serotype 1	(63)
6290	serotype 1	(63)
5637	serotype 1	(63)
5673	serotype 1/2	(63)
5679	serotype 1/2	(63)
5928	serotype 1/2	(63)
5934	serotype 1/2	(63)
5209	reference strains serotype 1/2	(63)
5218	reference strain serotype 9	(63)
5973	serotype 9	(63)
6437	serotype 9	(63)
6207	serotype 9	(63)
reference strains	serotypes 1-34	(9, 56, 14)
<i>S. suis</i>		
10	virulent serotype 2 strain	(51)
10cpsB	isogenic cpsB mutant of strain 10	this work
10cpsEF	isogenic cpsEF mutant of strain 10	this work
Plasmid		
pKUN19	replication functions pUC, Amp ^R	(23)
pGEM7Zf(+)	replication functions pUC, Amp ^R	Promega Corp.
pIC19R	replication functions pUC, Amp ^R	(29)
pIC20R	replication functions pUC, Amp ^R	(29)
pIC-spc	pIC19R containing spc ^R gene of pDL282	labcollection

WO 00/05378

PCT/NL99/00460

pDL282	replication functions of pBR322 and pVT736-1, Amp ^R , Spc ^R	(43)
pPHOS2	pIC-spc containing the truncated <i>phoA</i> gene of pPHO7 as a <i>Pst</i> I- <i>Bam</i> HI fragment	this work
pPHO7	contains truncated <i>phoA</i> gene	(15)
pPHOS7	pPHOS2 containing chromosomal <i>S. suis</i> DNA	this work
pCPS6	pKUN19 containing 6 kb <i>Hind</i> III fragment of <i>cps</i> operon	this work (Fig.1)
pCPS7	pKUN19 containing 3,5 kb <i>Eco</i> RI- <i>Hind</i> III fragment of <i>cps</i> operon	this work (Fig.1)
pCPS11	pCPS7 in which 0.4 kb <i>Pst</i> I- <i>Bam</i> HI fragment of <i>cpsB</i> gene is replaced by Spc ^R gene of pIC-spc	this work (Fig.1)
pCPS17	pKUN19 containing 3.1 kb <i>Kpn</i> I fragment of <i>cps</i> operon	this work (Fig.1)
pCPS18	pKUN19 containing 1.8 kb <i>Sna</i> BI fragment of <i>cps</i> operon	this work (Fig.1)
pCPS20	pKUN19 containing 3.3 kb <i>Xba</i> I- <i>Hind</i> III fragment of <i>cps</i> operon	this work (Fig.1)
pCPS23	pGEM7Zf(+) containing 1.5 kb <i>Mlu</i> I fragment of <i>cps</i> operon	this work (Fig.1)
pCPS25	pIC20R containing 2.5 kb <i>Kpn</i> I- <i>Sal</i> I fragment of pCPS17	this work (Fig.1)
pCPS26	pKUN19 containing 3.0 kb <i>Hind</i> III fragment of <i>cps</i> operon	this work (Fig.1)
pCPS27	pCPS25 containing 2.3 kb <i>Xba</i> I (blunt)- <i>Cla</i> I fragment of pCPS20	this work (Fig.1)
pCPS28	pCPS27 containing the 1.2 kb <i>Pst</i> I- <i>Xho</i> I Spc ^R gene of pIC-spc	this work (Fig.1)
pCPS29	pKUN19 containing 2.2 kb <i>Sac</i> I- <i>Pst</i> I fragment of <i>cps</i> operon	this work (Fig.1)
pCPS1-1	pKUN19 containing 5 kb <i>Eco</i> RV fragment of <i>cps</i> operon of type 1	this work (Fig.1)
pCPS1-2	pKUN19 containing 2.2 kb <i>Hind</i> III fragment of <i>cps</i> operon of type 1	this work (Fig.1)
pCPS9-1	pKUN19 containing 1 kb <i>Hind</i> III- <i>Xba</i> I fragment of <i>cps</i> operon of serotype 9	this work (Fig.1)
pCPS9-2	pKUN19 containing 4.0 kb <i>Xba</i> I- <i>Xba</i> I fragment of <i>cps</i> operon of serotype 9	this work (Fig.1)

Amp^R: ampicillin resistantSpc^R: spectinomycin resistant*cps*: capsular polysaccharide

**Table 2. Properties of Orfs in the cps locus of *S. suis* serotype 2 and similarities to gene product
other bacteria**

WO 00/05378

PCT/NL99/00460

ORF	nucleotide position in sequence	number of amino acids	GC%	proposed function of gene product ¹	similar gene product (% identity)
Orf2Z	1 -719	240	44	Unknown	<i>B. subtilis</i> YitS (26%)
Orf2Y	2079-822	419	38	Transcription regulation	<i>B. subtilis</i> YcxD (39%)
Orf2X	2202-2934	244	39	Unknown	<i>H. influenzae</i> YAAA (24%)
Cps2A	3041-4484	481	39	Regulation	<i>S. pneumoniae</i> Cps19fA (58%)
Cps2B	4504-5191	229	40	Chain length determination	<i>S. pneumoniae</i> type 3 Orf1 (58%)
Cps2C	5203-5878	225	40	Chain length determination/ Export	<i>S. pneumoniae</i> Cps23fD (63%)
Cps2D	5919-6648	243	38	Unknown	<i>S. pneumoniae</i> CpsB (62%)
Cps2E	6675-8052	459	33	Glycosyltransferase	<i>S. pneumoniae</i> Cps14E (56%)
Cps2F	8089-9256	389	32	Glycosyl transferase	<i>S. pneumoniae</i> Cps23fT
Cps2G	9262-10417	385	36	Glycosyltransferase	<i>S. thermophilus</i> EpsF (25%)
Cps2H	10808-12176	457	31	Glycosyltransferase	<i>S. mutans</i> RGPEC, ^N (29%)
Cps2I	12213- 13443	410	29	CP polymerase	<i>S. pneumoniae</i> Cps23fI (48%)
Cps2J	13583-14579	332	29	Glycosyltransferase	<i>S. pneumoniae</i> Cps14J (31%)
Cps2K	14574-15576	334	37	Glycosyltransferase	<i>S. pneumoniae</i> Cps14J (40%)

WO 00/05378

PCT/NL99/00460

Table 2 continued

"Cps2L"	15618-16635	103	37	Unknown	-
"Cps2M"	16811-17322	-	38	-	<i>S. agalactiae</i> CpsF ^N (77%) <i>E. coli</i> NeuA ^N (47%)
"Cps2N"	17559-18342	-	39	-	<i>S. agalactiae</i> CpsJ (43%)
Cps2O	18401-19802	476	40	Repeat unit transporter	<i>S. agalactiae</i> CpsK (41%)
Cps2P	20327-21341	338	39	Sialic acid synthesis	<i>S. agalactiae</i> NeuB (80%) <i>E. coli</i> NeuB (59%)
Cps2Q	21355-21865	170	42	Sialic acid synthesis	<i>S. agalactiae</i> NeuC ^N (61%) <i>E. coli</i> NeuC ^N (54%)
Cps2R	21933-22483	184	40	Sialic acid synthesis	<i>S. agalactiae</i> NeuC ^C (55%) <i>E. coli</i> NeuC ^C (40%)
Cps2S	22501-23125	208	42	Sialic acid synthesis	<i>E. coli</i> NeuD (32%)
Cps2T	23136-24366	395	40	CMP-NeuNAC synthetase	<i>S. agalactiae</i> CpsF (49%) <i>E. coli</i> NeuA (34%)
"Orf2U"	24566-25488	168	42	Transposase	<i>S. thermophilus</i> IS1194 (51%)
"Orf2V"	25691-26281	116	37	Transposase	<i>S. pneumoniae</i> orf1 (85%)

¹ Predicted by sequence similarity^N Similarity refers to the amino-terminal part of the gene product^C Similarity refers to the carboxy-terminal part of the gene product

ORFs between " " are truncated or non-functional as the result of frame-shift or point mutations

WO 00/05378

PCT/NL99/00460

TABLE 3. Properties of ORFs in the cps genes of *S. suis* serotypes 1 and 9 and similarities to gene products of other bacteria

ORF	nucleotide position in sequence	G + C%	number of amino acids	predicted mol. mass (kDa)	predicted pI	proposed function of gene product ¹	similar gene product (% identity)	reference/ accession nr.
Cps1E ²	1-1363	34%	454	52.2	8.0	Glucosyltransferase	<i>Streptococcus suis</i> Cps2E (86%)	(26)
(48%)							<i>Streptococcus pneumoniae</i> Cps14E (12)	
Cps1F	1374-1821	33%	149	17.3	8.2	Unknown	<i>Streptococcus pneumoniae</i> Cps14F (83%)	(14)
Cps1G	1823-2315	25%	164	19.5	7.5	Glycosyltransferase	<i>Streptococcus pneumoniae</i> Cps14G(50%) (14)	
Cps1H	3035-4202	24%	389	45.5	8.4	CP polymerase	<i>Streptococcus pneumoniae</i> Cps14H (30%)	(14)
Cps1I	4197-					Glycosyltransferase	<i>Streptococcus pneumoniae</i> Cps14J (38%)	(13)
							<i>Lactococcus lactis</i> EpsG (31%)	(29)
							<i>Streptococcus thermophilus</i> EpsI (33%)	(28)
Cps1J						Glycosyltransferase	<i>Streptococcus pneumoniae</i> Cps14J ()	

WO 00/05378

PCT/NL99/00460

Table 3 continued

Cps1K ³	37%	278	32.5	7.8	Glycosyltransferase	(13)
						Streptococcus pneumoniae Cps14J (44%) (13)
Cps9D ²	37%	215	24.9	8.1	Unknown	(26)
						Streptococcus suis Cps2D (89%)
Cps9E					Glycosyltransferase	(18)
						Staphylococcus aureus Cap1D (27%)
Cps9F	36%	200	22.3	8.2	Glycosyltransferase	(17)
						Staphylococcus aureus Cap5M (52%)
Cps9G	35%	269	31.5	8.0	Unknown	(43%)
						Actinobacillus actinomycetemcomitans (AB002668_4)
						Haemophilus influenzae Lsg (005081)
Cps9H ³	30%	143	16.5	7.2	Unknown	(33)
						Yersinia enterocolitica RfbB (28%)

¹ Predicted by sequence similarity² N-terminal part of protein is lacking³ C-terminal part of protein is lacking

PCT/NL99/00460

[illegible]

PCT/NL99/00460

Table 5. Hybridization of serotypes 1 and 9 cps genes with chromosomal DNA of other *S. suis* serotypes

[illegible]

PCT/NL99/00460

23	+	-	-	+	+	-
24	+	-	-	+	+	-
25	+	-	-	-	-	-
26	+	-	+	+	-	-
27	+	-	-	-	-	+
28	+	-	-	+	-	-
29	+	-	-	+	-	-
30	+	-	-	+	-	-
31	+	-	-	+	-	-
32	+	-	-	-	-	-
33	+	-	-	-	-	-
34	+	-	-	-	-	-
1/2	+	-	-	-	+	-

WO 00/05378

PCT/NL99/00460

TABLE 6. Virulence of wild type and capsular mutant *S. suis* strains in germfree pigs

<i>S. suis</i> strains ¹	pigs/ group [n]	mortality ² [%]	morbidity ³ [%]	clinical index of the group	spec symptoms ⁵	non-spec. symptoms ⁶	fever index ⁷	leuco- cyte index ⁸	CNS	isolation of <i>S. suis</i> in pigs [n] per group in	serosae	joints
10	4	100	100		11	88	43	44	2	3		4
10cpsB	4	0	0		0	10	1	3	1	3		2
10cpsEF	4	0	0		0	0	1	0	1	3		2

¹ strain10 in the wild type strain, strains 10cpsB and 10cpsEF are isogenic capsular mutant strains

² piglets which died spontaneously or had to be killed for animal welfare reasons

³ only considering pigs with specific symptoms

⁴ clinical index: % of observations which matched the described criteria

⁵ specific symptoms: ataxia, lemeness on at lest one joint, stiffness

⁶ non-specific symptoms: inappetance, depression

⁷ % of observations in the experimental group with a body temperature > 40° C

⁸ % of blood samples in the group in which number of granulocytes > 10¹⁰/l

Table 7. Bacterial strains and plasmids

strain/plasmid	relevant characteristics
Strain	
<i>E. coli</i>	
XL2 blue	
<i>S. suis</i>	
reference strains	serotypes 1-34
5667	serotype 7, tonsil (1993)
7037	serotype 7, organs (1994)
7044	serotype 7, brains (1994)
7068	serotype 7 (1994)
7646	serotype 7 (1994)
7744	serotype 7, lungs (1996)
7759	serotype 7, joints (1996)
8169	serotype 7 (1997)
15913	serotype 7, meninges (1998)
Plasmid	
pKUN19	replication functions pUC, Amp ^R
pGEM7Zf(+)	replication functions pUC, Amp ^R
pCPS9-1	pKUN19 containing 1 kb <i>HindIII</i> - <i>XbaI</i> fragment of cps operon of serotype 9
pCPS9-2	pKUN19 containing 4.0 kb <i>XbaI</i> - <i>XbaI</i> fragment of cps operon of serotype 9
pCPS7-1	pKUN19 containing 1.6-kb <i>PstI</i> fragment of cps operon of type 7
pCPS7-2	pGEM7 containing 2.7-kb <i>ScaI</i> - <i>Clal</i> fragment of cps operon of type 7

*Amp^R: ampicillin resistant
cps: capsular polysaccharide

WO 00/05378

PCT/NL99/00460

Table 8. Properties of Orfs in the cps genes of *S. suis* serotype 7 and similarities to gene products of other bacteria

Orf	nucleotide position in sequence	proposed function of gene product	similar gene product (% identity)
Cps7E	1-719	Glycosyltransferase	<i>Streptococcus suis</i> Cps9E (99%)
Cps7F	1164-1863	Glycosyltransferase	<i>Bordetella pertussis</i> Bp1G ¹ (43%) <i>Streptococcus suis</i> Cps2E ¹ (33%)
Cps7G	1872-3086	Biosynthesis amino sugar	<i>Bordetella pertussis</i> Bp1F (48%)
Cps7H	3104-3737	Glycosyltransferase	<i>Escherichia coli</i> WbdN (35%) <i>Streptococcus suis</i> Cps2K ² (31%)

¹similarity refers to the C-terminal part of the gene product

²similarity refers to the N-terminal part of the gene product

Table 9.

[illegible]

cps7E
cps7F
cps7G
cps7H
16SrRNA

LEGENDS TO FIGURES

Figure 1.

Organization of the *cps2* gene cluster of *S. suis* type 2.

- 5 (A) Genetic map of the *cps2* gene cluster. The shadowed arrows represent potential ORFs. Interrupted ORFs indicate the presence of stop codons or frame-shift mutations. Gene designations are indicated below the ORFs. The closed arrows indicate the position of the potential promoter sequences. | indicates the position of the potential transcription regulator sequence. ||| indicates the position of the 100-bp repeated sequence.

(B) Physical map of the *cps2* locus.

- Restriction sites are as follows: A: AluI; C: ClaI; E, EcoRI; 15 H, HindIII; K, KpnI; M, MluI; N, NsiI; P, PstI; S, SnaBI; Sa: SacI; X, XbaI.

(C) The DNA fragments cloned in the various plasmids.

Figure 2

- 20 Ethidium bromide stained agarose gel showing PCR products obtained with chromosomal DNA of *S. suis* strains belonging to the serotypes 1, 2, $\frac{1}{2}$, 9 and 14 and *cps2J*, *cps1I* and *cps9H* primer sets as described in Materials and Methods. (A) *cps1I* primers.

- 25 (B) *cps2J* primers and (C) *cps9H* primers. Lanes 1-3: serotype 1 strains; lanes 4-6: serotype 2 strains; lanes 7-9: serotype $\frac{1}{2}$ strains; lanes 10-12: serotype 9 strains and lanes 13-15: serotype 14 strains.

- (B) Ethidium bromide stained agarose gel showing PCR products 30 obtained with tonsillar swabs collected from pigs carrying *S. suis* type 2, type 1 or type 9 strains and *cps2j*, *cps1I* and *cpsH* primer sets as described in Materials and Methods.

- Bacterial DNA suitable for PCR was prepared by using the multiscreen methods as described previously (20). (A) *cps1I* 35 primers. (B) *cps2J* primers and (C) *cps9H* primers. Lanes 1-3: PCR products obtained with tonsillar swabs collected from pigs

WO 00/05378

PCT/NL99/00460

carrying *S.suis* type 1 strains; lanes 4-6: PCR products obtained with tonsillar swabs collected from pigs carrying *S.suis* type 2 strains; lanes 7-9: PCR products obtained with tonsillar swabs collected from pigs carrying *S.suis* type 9 strains; lanes 10-12: PCR products obtained with chromosomal DNA from serotype 9, 2 and 1 strains respectively; lane 13: negative control, no DNA present.

Figure 3

10 CPS2 nucleotide sequences and corresponding amino acid sequences from the open reading frames.

Figure 4

15 CPS1 nucleotide sequences and corresponding amino acid sequences from the open reading frames.

Figure 5

20 CPS9 nucleotide sequences and corresponding amino acid sequences from the open reading frames.

Figure 6

25 CPS7 nucleotide sequences and corresponding amino acid sequences from the open reading frames.

Figure 7

Alignments of the N-terminal parts of Cps2J and Cps2K. Identical amino acids are marked by bars. The amino acids shown in bold are also conserved in Cps14I, Cps14J of *S. pneumoniae* and several other glycosyltransferases (19). The aspartate residues marked by asterics are strongly conserved.

30

Figure 8

Transmission electron micrographs of thin sections of various *S. suis* strains.

35

(A) wild type strain 10;

WO 00/05378

PCT/NL99/00460

(B) mutant strain 10cpsB;
 (C) mutant strain 10cpsEF.
 Bar = 100 nm

5 Figure 9

(A) Kinetics of phagocytosis of wild type and mutant *S. suis* strains by porcine alveolar macrophages. Phagocytosis was determined as described in Materials and Methods. The Y-axis represents the number of CFU per milliliter in the supernatant
 10 fluids as determined by plate counting, the X-axis represents time in minutes.

- wild type strain 10;
- mutant strain 10cpsB;
- Δ mutant strain 10cpsEF.

15

(B) Kinetics of intracellular killing of wild type and mutant *S. suis* strains by porcine AM. The intracellular killing was determined as described in Material and Methods. The Y-axis represents the number of CFU per ml in the supernatant fluids
 20 after lysis of the macrophages as determined by plate counting, the X-axis represents time in minutes.

- wild type strain 10;
- mutant strain 10cpsB;
- Δ mutant strain 10cpsEF.

25

Figure 10

Nucleotide sequence alignment of the highly conserved 100-bp repeated element.

- 1) 100-bp repeat between cps2G and cps2H
- 30 2) 100-bp repeat within "cps2M"
- 3) 100-bp repeat between cps2O and cps2P

Figure 11. The cps2, cps9 and cps7 gene clusters of *S. suis*
 35 serotypes 2, 9 and 7.

WO 00/05378

PCT/NL99/00460

(A) Genetic organization of the *cps2* gene cluster [84]. The large arrows represent potential ORFs. Gene designations are indicated below the ORFs. Identically filled arrows represent ORFs which showed homology. The small closed arrows indicate the position of the potential promoter sequences. | indicates the position of the potential transcription regulator sequence.

(B) Physical map and genetic organization of the *cps9* gene cluster [15]. Restriction sites are as follows: B: BamHI; P: PstI; H: HindIII; X:XbaI. The DNA fragments cloned in the various plasmids are indicated. The open arrows represent potential ORFs.

(C) Physical map and genetic organization of the *cps7* gene cluster. Restriction sites are as follows: C: ClaI; P: PstI; Sc: ScaI. The DNA fragments cloned in the various plasmids are indicated. The open arrows represent potential ORFs.

Figure 12 (A) Ethidium bromide stained agarose gel showing PCR products obtained with chromosomal DNA of *S. suis* strains belonging to the serotypes 1, 2, 9 and 7 and the *cps7H* primer set. Strain designations are indicated above the lanes. C: negative control, no DNA present. M: molecular size marker (lambda digested with EcoRI and HindIII).

(B) Ethidium bromide stained agarose gel showing PCR products obtained with serotype 7 strains collected in different countries and from different organs. Bacterial DNA suitable for PCR was prepared by using the multiscreen method as described previously [89]. Strain designations are indicated above the lanes. M: molecular size marker (lambda digested with EcoRI and HindIII).

REFERENCES

1. Arends, J. P., and H. C. Zanen. 1988. Meningitis caused by *Streptococcus suis* in humans. Rev. Infect. Dis. 10:131-137.
2. Arrecubieta, C., E. Garcia, and R. Lopez. 1995. Sequence and transcriptional analysis of a DNA region involved in the
5 production of capsular polysaccharide in *Streptococcus pneumoniae* type 3. Gene 167: 1-7.
3. Arrecubieta, C., R. Lopez, and E. Garcia. 1994. Molecular characterization of *cap3A*, a gene from the operon required for the synthesis of the capsule of *Streptococcus pneumoniae*
10 type 3: sequencing of mutations responsible for the unencapsulated phenotype and localization of the capsular cluster on the pneumococcal chromosome. J. Bacteriol. 176: 6375-6383.
4. Clifton-Hadley, F.A. 1983. *Streptococcus suis* type 2
15 infections. Br. Vet. J. 139:1-5.
5. Charland, N., J. Harel, M. Kobisch, S. Lacasse, and M. Gottschalk. 1998. *Streptococcus suis* serotype 2 mutants deficient in capsular expression. Microbiol. 144:325-332.
6. Cross, A. S. 1990. The biological significance of
20 bacterial encapsulation. Curr. Top. Microbiol. Immunol. 150: 87-95.
7. Elliott, S. D. and J. Y. Tai . 1978. The type specific polysaccharide of *Streptococcus suis*. J. Exp. Med. 148: 1699-1704.
- 25 8. Feder, I., M. M. Chengappa, B. Fenwick, M. Rider and J. Staats. 1994. Partial characterization of *Streptococcus suis* type 2 hemolysin. J. Clin. Microbiol. 32:1256-1260.
9. Gottschalk, M., R. Higgins, M. Jacques, M. Beaudoin, and J. Henrichsen. 1991. Characterization of six new capsular
30 types (23 through 28) of *Streptococcus suis*. J. Clin. Microbiol. 29:2590-2594.
10. Gottschalk, M., S. Lacouture, and J. D. Dubreuil. 1995. Characterization of *Streptococcus suis* type 2 haemolysin.

Microbiology 141:189-195.

11. Gottschalk, M., A. Lebrun, M. Jacques, and R. Higgins. 1990. Haemagglutination properties of *Streptococcus suis*. J. Clin. Microbiol. 28: 2156-2158.
- 5 12. Guidolin, A., J. M. Morona, R. Morona, D. Hansman, and J. C. Paton. 1994. Nucleotide sequence analysis of genes essential for capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* type 19F. 1994. Infect. Immun. 62: 5384-5396.
- 10 13. Guitierrez, C., and J. C. Devedjian. 1989. Plasmid facilitating in vitro construction of PhoA fusions in *Escherichia coli*. Nucl. Acid. Res. 17: 3999.
14. Higgins, R., M. Gottschalk, M. Boudreau, A. Lebrun, and J. Henrichsen. 1995. Description of six new capsular types (28
15 through 34) of *Streptococcus suis*. J. Vet. Diagn. Invest. 7:405-406
15. Jacobs, A. A., P. L. W. Loeffen, A. J. G. van den Berg, and P. K. Storm. 1994. Identification, purification and characterization of a thiol-activated hemolysin (suilysin) of
20 *Streptococcus suis*. Infect. Immun. 62: 1742-1748.
16. Jacques, M., M. Gottschalk, B. Foirey B. and R. Higgins. 1990. Ultrastructural study of surface components of *Streptococcus suis*. J. Bacteriol. 172:2833-2838.
17. Klein P., M. Kanehisa and C. DeLisi. 1985. The detection
25 and classification of membrane spanning proteins. Biochim. Biophys. Acta. 851: 468-476.
18. Kolkman, M. A. B., D. A. Morrison, B. A. M. van der Zeijst, and P. J. M. Nuijten. 1996. The capsule polysaccharide synthesis locus of *Streptococcus pneumoniae* serotype 14:
30 identification of the glycosyl transferase gene *cps14E*. J. Bacteriol. 178: 3736-3541.
19. Kolkman, M. A. B., W. Wakarchuk, P. J. M. Nuijten, and B. A. M. van der Zeijst. 1997. Capsular polysaccharide synthesis in *Streptococcus pneumoniae* serotype 14: molecular analysis of
35 the complete *cps* locus and identification of genes encoding glycosyltransferases required for the biosynthesis of the

- tetrasaccharide subunit. Mol. Microbiol. 26: 197-208.
20. Kolkman, M. A. B., B. A. M. van der Zeijst and P. J. M. Nuijten. 1997. Functional analysis of glycosyltransferases encoded by the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae* serotype 14. J. Biol. Chem. 272: 19502-19508.
21. Konings, R. N. H., E. J. M. Verhoeven, and B. P. H. Peeters. 1987. pKUN vectors for the separate production of both DNA strands of recombinant plasmids. Methods Enzymol. 153: 12-34.
22. Korolik, V., B. N. Fry, M. R. Alderton, B. A. M. van der Zeijst, and P. J. Coloe. 1997. Expression of *Campylobacter hyoilei* lipo-oligosaccharide (LOS) antigens in *Escherichia coli*. Microbiol. 143: 3481-3489.
23. Leij, P. C. J., R. van Furth, and T. L. van Zwet. 1986. In vitro determination of phagocytosis and intracellular killing of polymorphonuclear and mononuclear phagocytes. In Handbook of Experimental Immunology, vol. 2. Cellular Immunology, pp. 46.1-46.21. Edited by D. M. Weir, L. A. Herzenberg, C. Blackwell and L. A. Herzenberg. Blackwell Scientific Publications, Oxford.
24. Lin, W. S., T. Cunneen, and C. Y. Lee. 1994. Sequence analysis and molecular characterization of genes required for the biosynthesis of type 1 capsular polysaccharide in *Staphylococcus aureus*. J. Bacteriol. 176: 7005-7016.
25. Liu, D., A. M. Haase, L. Lindqvist, A.A. Lindberg, and P. R. Reeves. 1993. Glycosyl transferases of O-antigen biosynthesis in *Salmonella enteritica*: Identification and characterization of transferase genes of group B, C2, and E1. J. Bacteriol. 175: 3408-3413.
26. Manoil, C., and J. Beckwith. 1985. A transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA 82: 8129-8133.
27. Marsh, J. L., M. Erfle, and E. J. Wykes. 1984. The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene

- 32:481-485.
28. Miller, J. 1972. Experiments in Molecular Genetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
29. Morona, J. K., R. Morona, and J. C. Paton. 1997.
- 5 Characterization of the locus encoding the *Streptococcus pneumoniae* type 19F capsular polysaccharide biosynthesis pathway. Mol. Microbiol. 23: 761-763.
30. Muñoz, R., M. Mollerach, R. López and E. Garcia. 1997. Molecular organization of the genes required for the synthesis
- 10 of type 1 capsular polysaccharide of *Streptococcus pneumoniae*; formation of binary encapsulated pneumococci and identification of cryptic dTDP-rhamnose biosynthesis genes. Mol. Microbiol. 25: 79-92.
31. Pearce B. J., Y. B. Yin, and H. R. Masure. 1993. Genetic
- 15 identification of exported proteins in *Streptococcus pneumoniae*. Mol. Microbiol. 9: 1037-1050.
32. Roberts, I. S. 1996. The biochemistry and genetics of capsular polysaccharide production in bacteria. Ann. Rev. Microbiol. 50: 285-315.
- 20 33. Rossbach, S., D. A. Kulpa, U. Rossbach, and F. J. de Bruin. 1994. Molecular and genetic characterization of the rhizopine catabolism (mocABRC) genes of *Rhizobium meliloti* L5-30. Mol. Gen. Genet. 245: 11-24.
34. Rubens, C. E., L. M. Heggen, R. F. Haft, and R. M.
- 25 Wessels. 1993. Identification of *cpsD*, a gene essential for type III capsule expression in group B streptococci. Mol. Microbiol. 8: 843-855.
- 35 35. Rubens, C. E., L. M. R. Wessels, L. M. Heggen, and D. L. Kasper. 1987. Transposon mutagenesis of type III group B *Streptococcus*: correlation of capsule expression with
- 30 virulence. Proc. Natl. Acad. Sci. USA 84:7208-7212.
36. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning. A laboratory manual. Second edition. Cold Spring Harbor Laboratory Press. Cold Spring Harbor. New York.
- 35 37. Smith, H. E., U. Vecht, H. J. Wisselink, N. Stockhofe-Zurwieden, Y. Biermann, and M. A. Smits. 1996. Mutants of

Streptococcus suis types 1 and 2 impaired in expression of muramidase-released protein and extracellular protein induce disease in newborn germfree pigs. Infect Immun. 64: 4409-4412.

38. Smith, H. E., H. J. Wisselink, U. Vecht, A. L. J. Gielkens and M. A. Smits. 1995. High-efficiency transformation and gene inactivation in *Streptococcus suis* type 2. Microbiol. 141: 181-188.
39. Sreenivasan, P. K., D. L. LeBlanc, L. N. Lee, and P. Fives-Taylor. 1991. Transformation of *Actinobacillus actinomycetemcomitans* by electroporation, utilizing constructed shuttle plasmids. Infect. Immun. 59: 4621-4627.
40. Stringele F., J.-R. Neeser, and B. Mollet. 1996. Identification and characterization of the *eps* (exopolysaccharide) gene cluster from *Streptococcus thermophilus* Sfi6. J. Bacteriol. 178: 1680-1690.
41. Stockhofe-Zurwieden, N., U. Vecht, H. J. Wisselink, H. van Lieshout, and H. E. Smith. 1996. Comparative studies on the pathogenicity of different *Streptococcus suis* serotype 1 strains. In Proceedings of the 14th IPVS Congress. pp. 299.
42. van Kranenburg, R., J. D. Marugg, I. I. van Swam, N. J. Willem and W. M. de Vos. 1997. Molecular characterization of the plasmid-encoded *eps* gene cluster essential for exopolysaccharide biosynthesis in *Lactococcus lactis* Mol. Microbiol. 24: 387-397.
43. van Leengoed, L. A., E. M. Kamp, and J. M. A. Pol. 1989. Toxicity of *Haemophilus pleuropneumoniae* to porcine lung macrophages. Vet. Microbiol. 19: 337-349.
44. van Leengoed, L. A. M. G., U. Vecht, and E. R. M. Verheyen. 1987. *Streptococcus suis* type 2 infections in pigs in The Netherlands (part two). Vet Quart. 9, 111-117.
45. Vecht, U., J. P. Arends, E. J. van der Molen, and L. A. M. G. van Leengoed. 1989. Differences in virulence between two strains of *Streptococcus suis* type 2 after experimentally induced infection of newborn germfree pigs. Am. J. Vet. Res. 50:1037-1043.
46. Vecht, U., L. A. M. G. van Leengoed, and E. R. M.

- Verheyen. 1985. *Streptococcus suis* infections in pigs in The Netherlands (part one). Vet. Quart. 7:315-321
47. Vecht, U., H. J. Wisselink, M. L. Jellema, and H. E. Smith. 1991. Identification of two proteins associated with virulence of *Streptococcus suis* type 2. Infect. Immun. 59:3156-3162.
48. Vecht, U., H. J. Wisselink, N. Stockhofe-Zurwieden, and H. E. Smith. 1996. Characterization of virulence of the *Streptococcus suis* serotype 2 reference strain Henrichsen S 735 in newborn gnotobiotic pigs. Vet. Microbiol. 51:125-136.
49. Vecht, U., H. J. Wisselink, J. E. van Dijk, and H. E. Smith. 1992. Virulence of *Streptococcus suis* type 2 strains in newborn germfree pigs depends on phenotype. Infect. Immun. 60:550-556.
50. Wagenaar, F., G. L. Kok, J. M. Broekhuijsen-Davies, and J. M. A. Pol. 1993. Rapid cold fixation of tissue samples by microwave irradiation for use in electron microscopy. Histochemical J. 25: 719-725.
51. Wessels, M. R. and M. S. Bronze. 1994. Critical role of the group A streptococcal capsule in pharyngeal colonization and infection in mice. Proc. Natl. Acad. Sci. USA 91: 12238-12242.
52. Wessels, M. R., A. E. Moses, J. B. Goldberg, and T. J. DiCesare. 1991. Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci. Proc. Natl. Acad. Sci. USA. 88: 8317-8321.
53. Yamane, K., M. Kumamano, and K. Kurita. 1996. The 25°-36° region of the *Bacillus subtilis* chromosome: determination of the sequence of a 146 kb segment and identification of 113 genes. Microbiol. 142: 3047-3056.
54. Butler, J. C., R. F. Breiman, H. B. Lipman, J. Hofmann, and R. R. Facklam. 1995. Serotype distribution of *Streptococcus pneumoniae* infections among preschool children in the United States, 1978-1994: implications for development of a conjugate vaccine. J. Infect. Dis. 171: 885-889.
55. Charland, N., M. Jacques, S. Lacoutre and M. Gottschalk.

WO 00/05378

PCT/NL99/00460

1997. Characterization and protective activity of a monoclonal antibody against a capsular epitope shared by *Streptococcus suis* serotypes 1, 2 and 1/2. *Microbiol.* 143:3607-3614.
56. Gottschalk, M., R. Higgins, M. Jacques, K. R. Mittal, and J. Henrichsen. Description of 14 new capsular types of *Streptococcus suis*. *J. Clin. Microbiol.* 27:2633-2636.
57. Heath, P. J., B. W. Hunt, and J. P. Duff. 1996. *Streptococcus suis* serotype 14 as a cause of pig disease in the UK. *Vet. Rec.* 2:450-451.
- 10 58. Hommeez, J., L. A. Devrieze, J. Henrichsen, and F. Castryck. 1986. Identification and characterization of *Streptococcus suis*. *Vet. Microbiol.* 16:349-355.
59. Killper-Balz, R., and K. H. Schleifer. 1987. *Streptococcus suis* sp. nov. nom. rev. *Int. J. Syst. Bacteriol.* 15 37:160-162.
60. Kolkman, M. A. B., B. A. M. van der Zeijst, and P. J. M. Nuijten. 1998. Diversity of capsular polysaccharide synthesis gene clusters in *Streptococcus pneumoniae*. Submitted for publication.
- 20 61. Lee, J. C., S. Xu, A. Albus, and P. J. Livolsi. 1994. Genetic analysis of type 5 capsular polysaccharide expression by *Staphylococcus aureus*. *J. Bacteriol.* 176:4883-4889.
62. Reek, F. H., M. A. Smits, E. M. Kamp, and H. E. Smith. 1995. Use of multiscreen plates for the preparation of bacterial DNA suitable for PCR. *BioTechniques* 19: 282-285.
- 25 63. Sau, S., N. Bhasin, E. R. Wann, J. C. Lee, T. J. Foster, and C. Y. Lee. 1997. The *Staphylococcus aureus* allelic genetic loci for serotype 5 and 8 capsule expression contain the type-specific genes flanked by common genes. *Microbiol.* 143: 2395-30 2405.
64. Sau, S., and C. Y. Lee. 1996. Cloning of type 8 capsule genes and analysis of gene clusters for the production of different capsular polysaccharides in *Staphylococcus aureus*. *J. Bacteriol.* 178: 2118-2126.
- 35 65. Sau, S., and C. Y. Lee. 1997. Molecular characterization and transcriptional analysis of type 8 capsule genes in

- Staphylococcus aureus*. J. Bacteriol. 179:1614-1621.
66. Smith, H. E., M. Rijnsburger, N. Stockhofe-Zurwieden, H. J. Wisselink, U. Vecht, and M. A. Smits. 1997. Virulent strains of *Streptococcus suis* serotype 2 and highly virulent strains of *Streptococcus suis* serotype 1 can be recognized by a unique ribotype profile. J. Clin. Microbiol. 35:1049-1053.
67. Yamazaki, M., L. Thorne, M. Mikolajczak, R. W. Armentrout, and T. J. Pollock. 1996. Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas* strain S88. J. Bacteriol. 178:2676-2687.
68. Zhang, L., A. Al-Hendy, P. Toivanen, and M. Skurnik. 1993. Genetic organization and sequence of the *rfb* gene cluster of *Yersinia enterocolitica* serotype O:3: similarities to the dTDP-L-rhamnose biosynthesis pathway of *Salmonella* and to the bacterial polysaccharide transport systems. Mol. Microbiol. 9:309-321.
- 69 Clifton-Hadley, F.A. (1983). *Streptococcus suis* type 2 infections. Br. Vet. J. 139, 1-5.
- 70 Vecht, U., van Leengoed, L. A. M. G. and Verheyen, E. R. M. (1985). *Streptococcus suis* infections in pigs in The Netherlands (part one). Vet. Quart. 7, 315-321
- 71 Arends, J. P. and Zanen, H. C. (1988). Meningitis caused by *Streptococcus suis* in humans. Rev. Infect. Dis. 10, 131-137.
- 72 Hommeze, J., Devrieze, L.A., Henrichsen, J. and Castryck, F. (1986). Identification and characterization of *Streptococcus suis*. Vet. Microbiol. 16, 349-355.
- 73 Killper-Balz, R. and Schleifer, K. H. (1987). *Streptococcus suis* sp. nov. nom. rev. Int. J. Syst. Bacteriol. 37, 160-162.
- 74 Gottschalk, M., Higgins, R. and Jacques, M. (1993). Production of capsular material by *Streptococcus suis* serotype 2 under different conditions. Can. J. Vet. Res. 57, 49-52.
- 75 Higgins, R. and Gottschalk, M. (1990). An update on *Streptococcus suis* identification. J. Vet. Diagn. Invest. 2, 249-252.

WO 00/05378

PCT/NL99/00460

- 76 Gottschalk, M., Higgins, R., Jacques, M., Beaudoin, M. and Henrichsen, J. (1991). Characterization of six new capsular types (23 through 28) of *Streptococcus suis*. J. Clin. Microbiol. 29, 2590-2594.
- 5 77 Gottschalk, M., Higgins, R., Jacques, M., Mittal, K. R. and Henrichsen, J. (1989) Description of 14 new capsular types of *Streptococcus suis*. J. Clin. Microbiol. 27, 2633-2636.
- 78 Higgins, R., Gottschalk, M., Boudreau, M., Lebrun, A. and Henrichsen, J. (1995). Description of six new capsular types
10 (28 through 34) of *Streptococcus suis*. J. Vet. Diagn. Invest. 7, 405-406
- 79 Aarestrup, F. M., Jorsal, S. E. and Jensen, N. E. (1998). Serological characterization and antimicrobial susceptibility of *Streptococcus suis* isolates from diagnostic samples in
15 Denmark during 1995 and 1996. Vet. Microbiol. 15, 59-66.
- 80 MacLennan, M., Foster, G., Dick, K., Smith, W. J. and Nielsen, B. (1996). *Streptococcus suis* serotypes 7, 8 and 14 from diseased pigs in Scotland. Vet Rec. 139, 423-424.
- 81 Sihvonen, L., Kurl, D. N. and Henrichsen, J. (1988).
20 *Streptococcus suis* isolates from pigs in Finland. Acta Vet. Scand. 29, 9-13.
- 82 Boetner, A. G., Binder, M. and Bille-Hansen, V. (1987). *Streptococcus suis* infections in Danish pigs and experimental infection with *Streptococcus suis* serotype 7. Acta Path.
25 Microbiol. Immunol. Scand. Sect. B, 95, 233-239.
- 83 Smith, H. E., Veenbergen, V., van der Velde, J., Damman, M., Wisselink, H. J. and Smits, M. A. (1999). The cps genes of *Streptococcus suis* serotypes 1, 2 and 9: development of rapid serotype-specific PCR assays. J. Clin. Microbiol. submitted
- 30 84 Smith, H. E., Damman, M., van der Velde, J., Wagenaar, F., Wisselink, H. J., Stockhofe-Zurwieden, N. and Smits, M. A. (1999). Identification and characterization of the cps locus of *Streptococcus suis* serotype 2: the capsule protects against phagocytosis and is an important virulence factor. Infect.
35 Immun. 67, 1750-1756.
- 85 Miller, J. (1972). Experiments in Molecular Genetics.

WO 00/05378

PCT/NL99/00460

Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

86 Sambrook, J., E. F. Fritsch, and T. Maniatis. (1989).

Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

- 5 87 Allen, A. and Maskell, D. (1996). The identification, cloning and mutagenesis of a genetic locus required for lipopolysaccharide biosynthesis in *Bordetella pertussis*. *Mol. Microbiol.* 19, 37-52.

- 88 Wang, L. and Reeves, P. R. (1998). Organization of
10 *Escherichia coli* O157 O antigen gene cluster and identification of its specific genes. *Infect. Immun.* 66, 3545-3551.

- 89 Wisselink, H. J., Reek, F. H., Vecht, U., Stockhofe-Zurwieden, N., Smits, M. A. and Smith, H. E. (1999).
15 Detection of virulent strains of *Streptococcus suis* type 2 and highly virulent strains of *Streptococcus suis* type 1 in tonsillar specimens of pigs by PCR. *Vet. Microbiol.* 67, 143-157.

- 90 Konings, R. N. H., Verhoeven, E. J. M. and Peeters, B. P. H. (1987). pKUN vectors for the separate production of both
20 DNA strands of recombinant plasmids. *Methods Enzymol.* 153, 12-34.

SEQUENCE LISTING

<110> ID-LELYSTAD, INSTITUUT VOOR DIERHOUDERIJ EN DIERGEZONDHEID B.V.

<120> Streptococcus suis vaccines and diagnostic tests

<130> PAT 48613W-1

<140> CA 2,341,268

<141> 2001-01-19

<150> EP 98202465.5

<151> 1998-07-22

<150> EP 98202467.1

<151> 1998-07-22

<160> 53

<170> PatentIn Ver. 2.1

<210> 1

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 1

caaacgcaag gaattacggt atc

23

<210> 2

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 2

gagtatctaa agaatgccta ttg

23

<210> 3

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 3

ggcgggtctag cagatgctcg

20

<210> 4

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 4

gcgaactggt agcaatgac

19

<210> 5

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 5

ggctacatat aatggaagcc c

21

<210> 6

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 6

cggaagtatc tgggctactg

20

<210> 7

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

83

<400> 7

agctctaaca cgaaataagg c

21

<210> 8

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 8

gtcaaacacc ctggatagcc g

21

<210> 9

<211> 26281

<212> DNA

<213> Streptococcus suis

<220>

<223> /note="CPS2 nucleotide sequence"

<400> 9

aagcttggat	attgatcaca	tgatggaggt	gatggaagca	tctaagtctg	cagcgggggtc	60
ggcgtgcccc	agtccgcagg	cttatcaggc	agcttttgag	ggagctgaga	acattatcgt	120
tgtgacgatt	acaggtgggc	tatcgggtag	ttttaatgcg	gcacgtgtag	ctagggatat	180
gtatatcgaa	gagcatccga	atgtcaatat	ccatttgata	gatagtttgt	cagccagtgg	240
ggaaatggat	ttacttgtac	accaaataca	tcgcttaatt	agtgcaggat	tagattttcc	300
acaagtagta	gaagcgataa	ctcactatcg	ggaacacagt	aagctcctct	ttgttttagc	360
gaaagttgat	aatcttgтта	agaatggaag	actgagcaaa	ttggtaggca	ctgtcgttgg	420
tcttctcaat	atccgtatgg	ttggtgaggc	aagtgctgaa	ggaaaattag	agttgcttca	480
aaaggcgcgt	ggtcataaga	aatctgtgac	agcagccttt	gaagaaatga	aaaaagcagg	540
ctatgatggт	ggtcgaattg	ttatggcccc	ccgcaacaat	gctaagttct	tccaacaatt	600
ctcagagttg	gtaaaagcaa	gttttccaac	ggctgttatt	gacgaagttg	caacatcagg	660
tctatgcagt	ttttatgctg	aagaaggтgг	acttttgatg	ggctacgaag	tgaaagcgtg	720
attcacagag	taataatttt	gggctgтаа	ttccgctata	gaataatccc	cctcttcttc	780
taagttcgag	ggggattgтт	tgtatgagac	tattggattt	cattcattca	aatatcttac	840
gaattgctcc	agtttatctg	caaaatcttg	ttcaaagaag	atctgтааgа	aatcagcttt	900
ctgtccgctg	aaataataac	attttccaaa	catgtgttgg	atgctaggag	aaagaatccc	960
cttgcttagc	tgaaaggтca	cgctcccctt	tggaattcga	tacgggatgt	ttaaagcgtа	1020
tttctctaga	cagtctttta	ttttattcca	ttgagcgtga	taaatgtgat	gaagatgctg	1080
tgtgttccgc	gcaaacatac	cgttatcaat	gtagagcgag	agagcttttt	gcatgataag	1140
attggtatcg	tagtcgatta	gactcttatg	tttgatgaag	atatcacgta	gctgattagg	1200
aaggctgatt	gcaccgattc	ggagggcagg	aaagagtgtc	ggtgтаааg	atттtаtаtа	1260
gatgacgcga	ttatctgtat	caagatagtг	taaaggtagg	ctatgactag	agtcgaaatc	1320
tgctaaatag	tcatcctcaa	tgatgtagac	atcgтattgc	tttgctaatt	ttacgatggc	1380
tgttttttgt	gctatatcat	aggттgaacc	gagagggттg	tgcaagcgag	gaattgtgta	1440
gaaaaactta	atттttccag	tttggaagat	actttccaat	tcttctaggt	caattccatc	1500
taaattccgt	tcaattgttt	gataggggгat	tccttgatgt	cgaatgagct	ctatcattcg	1560
tgaataggta	gggttctcta	tcaagatttc	cgtttttcca	gccaaгgttt	ccatttgтgt	1620

gagaatatat	agagcttggt	gactaccagc	tgtgataacc	agctgggtctt	tttttgtata	1680
gacatgatag	tccattaaca	gactttgaac	ggaggaaatc	aattctgcca	atccctcttg	1740
ctggtgatag	tagttgaata	ggtaattttc	ccgcccata	agactttctt	ttagacaaat	1800
ccgaaaatct	tcataggtaa	ttcttgaaag	tctgtaggat	tgagctctac	aggtatggtc	1860
ttggaaatct	ctatcctcta	agatataata	accgcttttt	tcgacagcgt	agatcttatt	1920
ttggtatttt	aattccaaca	tagccttttg	gacagtgtct	ttgctacaat	gatattgctc	1980
gcggagttga	cggatagaag	gtaatttctc	tccacgtttg	aatcgatgtt	cctctattcc	2040
agtcaaaaata	tcttgatga	taacttgata	ttttttcatc	taggtcccct	tttttataga	2100
ctatgttact	agctagtata	tagaaaaaat	tgaagaaaga	caatatatga	ataatggggt	2160
tgaggttcag	gaattaagct	actctatggt	ataattaagt	gatgaaaata	attataccta	2220
atgcaaaaaga	agtaaataca	aatctagaga	atgcctcgtt	ttatctcctg	tctgatcgaa	2280
gcaagccggt	gctggatgcc	ataagtcaat	ttgatgtaaa	aaagatggct	gccttttata	2340
aattgaatga	agcaaaggct	gagttagaag	ctgaccgttg	gtatcgaaatc	aggacaggtc	2400
aagcaaaaac	ctatccagcc	tggcagttat	atgatgggtct	catgtatcgt	tatatggata	2460
ggcgaggtat	agattcgaaa	gaagaaaatt	atttacgtga	ccacgttcgt	gtagcgacag	2520
ccttatacgg	attgattcat	ccttttgaat	tcatttcacc	tcaccgctta	gattttcaag	2580
ggagcttaaa	gataggcaat	cagtctttga	aacagtactg	gcgaccgtat	tatgaccaag	2640
aagttggtga	tgatgaactg	attctctcac	tggcttcgtc	agaatttgag	caggtgtttt	2700
ctccccagat	tcagaaaaga	ttagttaaaa	ttcttttcat	ggaagaaaaa	gcaggtcagc	2760
taaaagttca	ctcgactata	tcaaaaaaag	gcagaggaag	attgctgtcc	tggttggcta	2820
agaacaatat	tcaggaatta	tcggacattc	aagattttta	ggtaggatggc	tttgaatatt	2880
gtacttccga	atcaacggca	aaccaactta	ccttcatacg	atcaataaaa	atgtgaaatt	2940
atgaaaaaga	taacgttttc	cagcgctaaa	aagggtagaa	aaatatatta	ttctatgata	3000
taatggatgc	gttataggta	aaagtctagg	aaggttggtt	atgaaaaaga	gaagcggacg	3060
aagtaagtcg	tccaagttca	aattggtaaa	ttttgcgctt	ttgggacttt	attccattac	3120
tctatgtttg	ttcttagtga	ccatgtatcg	ctataacatc	ctagattttcc	ggtattttaa	3180
ctatatgttg	acgctttttg	tagtaggagt	ggcagtattg	gctggattat	tgatgtggcg	3240
taagaaagcg	cgcataattt	cagcgctctt	acttgttttt	tcactgggtca	tcacgtctgt	3300
tgggatctat	ggaatgcaag	aagttgtaaa	attttcaaca	cgactaaatt	caaattcgac	3360
attttcagaa	tatgaaatga	gtatccttgt	cccagcaa	agtgatatta	cggacgttcg	3420
tcagcttact	agtatccttg	ctccagccga	atacgacca	gataacatca	ccgctttatt	3480
ggatgacata	tccaaaatgg	aatctactca	actagcaact	agccccggga	cttcttacct	3540
gacagcatat	caatctatgt	tgaatggcga	gagtcaagcg	atgggtgttca	acggagtttt	3600
taccaatatt	ttagaaaatg	aagatccagg	cttttcttca	aaagtgaaaa	aaatatatag	3660
tttcaaagtg	actcagactg	ttgaaacagc	tactaagcag	gtgagtggag	atagctttta	3720
tatctatatt	agtgggtattg	atgcttatgg	accgatttct	acgggtctctc	gttcagatgt	3780
caatatcatt	atgactgtca	atcgtgcgac	acataagatt	ttattgacaa	ctactccacg	3840
agattcatac	gttgcttttcg	cagatggcgg	gcaaaatcaa	tacgataaac	taacacatgc	3900
tggatatttac	gggtgtcaatg	cttctgtgca	caccttagaa	aattttttatg	ggattgacat	3960
tagcaattat	gtgcggttga	acttcatttc	cttccttcaa	ttaatcgact	tgggtgggtgg	4020
aattgatgta	tataacgatc	aagaattttac	aagttttacat	gggaattatc	atttccctgt	4080
tggacaagtt	cattttaaact	cagaccaagc	attaggcttc	gttcgagagc	gctactcttt	4140
aacaggggggt	gacaatgacc	gtggtaaaaa	ccaggaaaaa	gtgattgctg	ccttgattaa	4200
aaagatgagt	acgccagaga	atctaataaa	ttaccaggca	atcctatctg	gattggaagg	4260
ctcaattcaa	acggatttga	gcttagaaac	gattatgagt	ttagtgaata	cccaactaga	4320
atcaggaaca	caatttacag	tagagtcaca	agcattgaca	ggaacaggac	gctcagactt	4380
atcttcttat	gcgatgcctg	gatcacaact	ttatatgatg	gaaattaacc	aagatagtct	4440
ggagcaatca	aaggcagcga	ttcagtccgt	acttggtgaa	aaataaagat	tttaggagaa	4500
aatatgaaca	atcaagaagt	aaatgcaatc	gaaatcgatg	ttttattctt	actaaaaaca	4560
atttggagaa	agaaattttt	aattctctta	actgcagtgt	tgactgcggg	gttggcattt	4620
gtctacagta	gttttttagt	gacacctcaa	tatgactcca	ctaccgctat	ctatgtagtg	4680
agtcaaaaatg	ttgaagccgg	tgcgggcttg	actaaccaag	agttacaagc	gggtacctat	4740
ttggcaaaaag	actatcgggg	aattatccta	tcacaagatg	tattgacaca	agtagcaacg	4800
gaattgaatc	tgaaagagag	tttgaaagaa	aaaatatcag	tttctattcc	tgttgatact	4860
cgtatcgttt	ctatttctgt	gcgtgatgcg	gatccaaatg	aagcggcacg	tattgcaaat	4920

agccttcgca	cctttgcagt	gcaaaagggt	gttgaggtca	ccaaggtaag	cgatgtgacg	4980
acacttgaag	aagcagtcct	agcggaagaa	ccaaccactc	caaatacaaa	acgaaatatc	5040
ttgcttggtt	tattagctgg	aggatatctg	gcaacagggtc	ttgtactggt	tatggagggtt	5100
ttggatgacc	gtgtaaaacg	tcctcaggac	atcgaagagg	taatgggatt	gacattgcta	5160
ggtatagtac	cagattcgaa	gaaattaaaa	taggagaaca	atatggcgat	gttagaaatt	5220
gcacgtacaa	aaagagaggg	agtaaataaa	accgaggagt	atttcaatgc	tatccgtacc	5280
aatattcagc	ttagcggagc	agatattaag	gttggttggt	ttacctctgt	taaatcgaat	5340
gaaggtaaga	gtacaactgc	ggctagtctc	gctattgcct	atgctcgttc	aggttataag	5400
accgtcttgg	tggatgcaga	tatccgaaat	tcagtcatgc	ctggtttctt	caagccaatt	5460
acaaagatta	caggtttgac	ggattaccta	gcagggacaa	cagacttgtc	tcaaggatta	5520
tgcgatacag	atattccaaa	cttgaccgta	attgagtcag	gaaagggttc	tcccaaccct	5580
actgcccttt	tacaaagtaa	gaattttgaa	aatctacttg	cgactcttcg	tcgctattat	5640
gattatgtta	tcgttgactg	tccaccatta	ggactggtaa	ttgatgcagc	tatcattgca	5700
caaaaatgtg	atgcgatggg	tgcagtagta	gaagcaggca	atgttaagtg	ctcatctttg	5760
aaaaaagtaa	aagagcagtt	ggaacaaaca	ggcacaccgt	tcttaggcgt	tatcttgaac	5820
aaatatgata	ttgccactga	gaagtatagt	gaatacggaa	attacggcaa	aaaagcctaa	5880
tttctcagat	aacataagtt	tgataagtag	gtattaatat	gattgatatc	cattcgcata	5940
tcatatattg	tgtggatgac	ggccccaaaa	ctattgaaga	gagcctgagt	ttgataagcg	6000
aagcttatcg	tcaagggtgt	cgctatatcg	tagcgacatc	tcatagacga	aaagggatgt	6060
ttgaaacacc	agaaaaaatc	atcatgatta	actttcttca	acttaaagag	gcagtagcag	6120
aagtttatcc	tgaaatacga	ttgtgctatg	gtgctgaatt	gtattatagt	aaagatatct	6180
taagcaaact	tgaaaaaaag	aaagtaccaa	cacttaatgg	ctcgtgctat	attctcttgg	6240
agttcagtac	ggatactcct	tggaaagaga	ttcaagaagc	agtgaacgaa	atgacgctac	6300
ttgggctaac	tcccgtactt	gcccataatg	agcgttatga	tgctctggca	tttcagtcag	6360
agagagtaga	aaagctaatt	gacaagggat	gctacactca	ggtaaatagt	aaccatgtgt	6420
tgaagcctgc	tttaattggc	gaacgagcaa	agaattttaa	aaaacgtact	cgatatTTTT	6480
tagagcagga	tttagtacat	tgtgttgcta	gcgatatgca	taattttatat	agtagacctc	6540
cgtttatgag	ggaggcgtat	cagcttgtaa	aaaaagagta	tggtgaggat	agagcgaagg	6600
ctttgttcaa	gaaaaatcct	ttgttgatat	tgaaaaatca	agtacagtaa	cctcatagaa	6660
atagtggagg	agctatgaat	attgaaatag	gatatcgcca	aacgaaattg	gcattgtttg	6720
atatgatagc	agttacgatt	tctgcaatct	taacaagtca	tataccaaat	gctgatttaa	6780
atcgttctgg	aattttttatc	ataatgatgg	ttcattatTT	tgcattTTTT	atatctcgta	6840
tgccgggtga	atttgagtat	agaggtaatc	tgatagagtt	tgaaaaaaca	tttaactata	6900
gtataatat	tgtaattttt	cttatggcag	tttcattttat	gttagagaat	aatttcgcac	6960
tttcaagacg	tggtgccgtg	tatttcacat	taataaactt	cgtttttggt	tacctattta	7020
acgtaattat	taagcagttt	aaggatagct	ttctatTTTT	gacaacctat	caaaaaaaga	7080
cgattcta	tacaacggct	gaactatggg	aaaatatgca	agtttttatt	gaatcagata	7140
tactatttca	aaaaaatcct	gttgcatagg	taatttttagg	tacagaaata	gataaaatta	7200
atttaccatt	accgctctat	tattctgttg	aagaagctat	agggttttca	acaagggag	7260
tggtcgacta	cgtctttata	aattttaccaa	gtgaatatTT	tgacttaaag	caattagttt	7320
cagactttga	gttggttaggt	attgatgtag	gcgttgatat	taattcattc	ggttttactg	7380
tggtgaagaa	taaaaaaatc	caaagtctag	gtgaccatag	catcgctact	ttttccacaa	7440
atttttataa	gcctagtcac	atctggatga	aacgactttt	agatatactt	ggagcagtag	7500
tcgggtta	tattagtggg	atagtttcta	ttttgtta	tccaattatt	cgtagagatg	7560
gtgggccagc	catttttgct	cagaaacgag	ttggacagaa	tggacgcata	tttacattct	7620
acaagtttcg	ttcgatgttt	gttgatgccg	aggtacgtaa	gaaagaatta	atggctcaaa	7680
accagatgca	agggtgggatg	ttcaaaatgg	acaacgatcc	tagaattact	ccaattggac	7740
acttcatacg	aaaaacaagt	ttagatgagt	taccacaatt	ttataatggt	ctaattggag	7800
atatgagtct	agtcggtacc	cgtccgccta	cagttgatga	atttgaaaaa	tatactccta	7860
gtcaaaagag	aagattgagt	tttaaaccag	ggattacagg	tctttggcaa	gtgagcggaa	7920
gaagtgat	cacagatttt	aatgaagtcg	ttaggctgga	cctaacatac	attgataatt	7980
ggaccatctg	gtcagacatt	aagattttat	tgaagacagt	gaaagttgta	ttgttgagag	8040
agggaggtca	gtaagactcc	tttaaaacaa	agaatagtag	taggggat	gagaacagtt	8100
tatatatttg	gttcaaaagg	aataccagca	aagtatgggtg	gtttcgagac	tttcgtagaa	8160
aaattaactg	agtatcagaa	agataaatca	attaattatt	ttgttgcatg	tacaagagaa	8220

aattcagcaa	aatcagatat	tacaggagaa	gtttttgaac	ataatggagc	aacatgtttt	8280
aatattgatg	tgccaaatat	tggttcagca	aaagccattc	tttatgatat	tatggctctc	8340
aagaaatcta	ttgaaattgc	caaagataga	aatgatacct	ctccaatttt	ctacattctt	8400
gcttgtcgga	ttggtccttt	catttatctt	tttaagaagc	agattgaatc	aattggaggt	8460
caacttttctg	taaacccaga	cggtcatgaa	tggtctacgtg	aaaagtggag	ttatcccgtc	8520
cgacagtatt	ggaaattttc	tgagagtttg	atgttaaaat	acgctgattt	actaatttgt	8580
gatagcaaaa	atattgaaaa	atatattcat	gaagattatc	gaaaatatgc	tcctgaaaca	8640
tcttatattg	cttatggaac	agacttagat	aaatcacgcc	tttctccgac	agatagtgtg	8700
gtacgtgagt	ggtataagga	gaaggaaatt	tcagaaaatg	attactattt	ggttggttga	8760
cgatttggtgc	ctgaaaataa	ctatgaagta	atgattcgag	agtttatgaa	atcatattca	8820
agaaaagatt	ttgttttgat	aacgaatgta	gagcataatt	ccttttatga	gaaattgaaa	8880
aaagaaacag	ggttcgataa	agataagcgt	ataaagtttg	ttggaacagt	ctataatcag	8940
gagctgttaa	aatatatctg	tgaaaatgca	tttgcttatt	ttcatgggtca	cgagggttga	9000
ggaacgaacc	catctttact	tgaagcactt	tcttctacta	aactaaatct	tcttctagat	9060
gtgggcttta	atagagaagt	aggggaagaa	ggagcgaaat	actggaataa	agataatctt	9120
cacagagtta	ttgacagtgtg	tgagcaatta	tcacaagaac	aaattaatga	tatggatagt	9180
ttatcaacaa	aacaagtcaa	agaaagattt	tcttgggatt	ttattgttga	tgagtatgag	9240
aagttgttta	aaggataagt	tatgaaaaag	attctatatc	tccatgctgg	agcagaatta	9300
tatggggcag	ataagggttct	cttggaaactt	ataaaaggct	tagataagaa	tgaatttgaa	9360
gcgcagtgtta	tcctacctaa	tgatggagtc	ctagtgccag	cattaagaga	agttgggtgcg	9420
caagttgaag	ttattaacta	tccaattcta	cgtaggaaat	attttaatcc	aaaagggtt	9480
tttgactact	tcatatcata	tcatcactat	tctaaacaga	ttgctcaata	tgccatagaa	9540
aataagggttg	acataattca	caataatact	accgctgtct	tagaaggcat	ttatctgaag	9600
cgaaaactca	aattaccttt	gttggtggcat	gttcatgaga	ttattgtcaa	acctaaattc	9660
atctctgatt	cgatcaattt	tttaatgggg	cgttttgctg	ataagattgt	gacagtttca	9720
caggctgtgg	caaaccatat	aaaacaatca	cctcatatca	aagatgacca	aatcagtgtg	9780
atctacaatg	gggtagataa	taaagtgttt	tatcagtcctg	atgctcggtc	tgttcgagaa	9840
agatttgaca	ttgacgaaga	ggctcttgct	attggtatgg	tcggctcgagt	caatgcgtgg	9900
aaaggacaag	gagatttttt	agaagcagtt	gctcctatac	tcgaacagaa	tccaaaagct	9960
atcgccctta	tagcaggaag	tgcttttgaa	ggagaagagt	ggcgagttagt	agaattagaa	10020
aagaagattt	ctcaattaaa	ggtctcttct	caagtcagac	gaatggatta	ttatgcaa	10080
accactgaat	tatataatat	gtttgatatt	tttgacttc	caagtactaa	tccagaccct	10140
ctaccaacgg	ttgtactaaa	agcaatggca	tgcggtaaac	ctggtgtcgg	ttaccgacat	10200
ggtggtgttt	gtgagatggt	gaaagaaggt	gttaacggtt	tcttagtcac	tccgaactca	10260
ccgttaaatt	tatcaaaagt	aattcttcag	ttatcgga	atataaatct	cagaaaaaaa	10320
attggttaata	attctataga	acgtcaaaaa	gaacattttt	cgttaaaaag	ctatgtaaaa	10380
aatttttcga	aagtctacac	ctccctcaaa	gtatactgat	tggtctgaagt	gaatgcttta	10440
gtatagcgat	ttatcgat	ctcattcgat	aaaacaaatg	ttcagaaaca	gttataagtt	10500
atttctaaag	ggcacctcta	taaactccca	aaattgcgaa	tttggtggtta	cgaaagcctt	10560
gttaaatcaa	catttttaaat	tttagaaaat	tagtttttag	agctccccta	aaatagaaga	10620
taacagaagg	gagccttcaa	aaacttcatt	tttaattgga	ttgtagaaaa	actgttaaat	10680
caatatttag	atttttagga	gttcagtttt	tggtggggaga	gcttaataat	ctatgcacta	10740
tatttcgaaa	aatatatggt	gtaaaatcag	aactgatggt	cgtggcaaaa	aagagaatga	10800
ggaatttatg	aaaattat	ctttttacaat	ggttaataac	gaaagtgaga	taatagagtc	10860
atttatacgg	tataattata	actttattga	cgagatgggtc	attattgata	atgggtgtac	10920
agataaacag	atgcaaatta	tttttaattt	gattaaagag	ggatataaaa	tatccgtata	10980
tgatgagtc	ttagaggcat	ataatcagta	tcgacttgat	aataaatatc	taacgaaaat	11040
aattgctgaa	aaaaatccag	atttgataat	acctttggat	gcggatgaat	ttttaacagc	11100
cgattcaaat	ccacggaaac	ttttggaaca	actggactta	gaaaagatac	attatgtgaa	11160
ttggcaatgg	tttggttatga	ctaaaaaaga	tgatattaat	gattcggttta	taccacgtag	11220
aatgcaatat	tgttttgaaa	aacctgtttg	gcattcattct	gatggtaaac	cagttactaa	11280
atgtataatt	tccgctaagt	attacaaaaa	aatgaattta	aagctatcga	tgggacatca	11340
cactgttttt	ggtaacccaa	atgtaaggat	agaacatcat	aatgatttga	aatttgcaca	11400
ttatcgagct	attagccaag	agcaattaat	ttataaaaca	atttggttaca	ctattcgcca	11460
tattgctact	atggagaaca	atatcgaaac	agctcaaaga	acaaatcaga	tggcgctcat	11520

tgaatctggc	gtggatatgt	gggaaacggc	gagagaagcc	tcttattcag	gttatgattg	11580
taatgttata	catgcaccaa	ttgattttaag	tttttgtaaa	gaaaatattg	taataaaaata	11640
taacgaacta	tccagagaaa	cagtagcaga	acgcgtgatg	aaaacgggaa	gagaaatggc	11700
tgttcgtgca	tataatgtgg	agcgaaaaca	aaaagaaaag	aaattttctaa	aacctattat	11760
at ttgtatta	gatgggttaa	aaggagatga	gtatat tcat	cccaatccat	caaatcattt	11820
gacgatctta	actgaaatgt	ataacgtcag	aggcttactt	accgataatc	accaaattaa	11880
at ttctcaaa	gttaattata	gattaattat	aactccagat	tttgctaagt	ttttaccgca	11940
tgaatttatt	gttgtagcag	ataccttgga	tatagagcaa	gttaaaagcc	agtatgttgg	12000
tacagggtga	gacttggtcaa	agattat ttc	tttaaaagag	tatcgaaaag	agataggctt	12060
tattggtaat	ttgtatgcgc	ttttaggatt	tgttccgaat	atgctcaata	gaatttatct	12120
atatattcag	agaaacggta	ttgcaaacac	tattataaaa	atcaagtcga	gattgtgaga	12180
gttggtttact	tttattttgta	at tttaaaaag	taatgcaggc	agataggaga	aaaacgtttg	12240
gaaaaatgag	aataagaatt	aataatttgt	tttttgttgc	catagcgttt	atgggcataa	12300
ttattagtaa	ttcgcaagtt	gttctagcga	taggcaaagc	ttctgtgatt	cagtatctat	12360
cttatttagt	tttgatttta	tgtatagtta	atgatttatt	aaaaaataac	aaacatattg	12420
tagtttataa	attagggtat	ttgtttctta	ttatat tttt	atttactatc	ggaatatgtc	12480
agcaaattct	tcctataaca	actaaaatat	at ttatcaat	ttcaatgatg	attat tttcag	12540
ttttagcaac	gttgccaata	agtttgataa	aagatattga	tgatttttaga	cggattttcaa	12600
atcatttggt	attcgctctt	tttataactt	cgatattagg	aataaagatg	ggggcaacga	12660
tgttcacggg	ggcagtagaa	ggtatcgggt	ttagtcaggg	ttttaatgga	ggattgacgc	12720
ataagaactt	ttttggaata	actat tttta	tgggggttcgt	attaacttac	ttggcgtata	12780
agtatggttc	ctataaaaaga	acggatcgtt	ttat ttttagg	attagaattg	tttttgattc	12840
ttat tttcaaa	cacacgctca	gtttat ttaa	tactattgct	ttttctat t	cttggttaatc	12900
ttgacaaaat	caaaatagaa	caaagacaat	ggagtacgct	taaatatatt	tccatgctat	12960
tttggtgctat	tttttttatac	tat tttctttg	gtttttttaat	aacacatagt	gattccttacg	13020
ctcatcgcgt	taatgggtctt	attaat tttt	ttgagtatta	tagaaatgat	tggttccatc	13080
taatgttttg	tgcagcggat	ttggcatatg	gggattttaac	tttagactat	gctataaggg	13140
ttagacgcgt	tttaggttgg	aatggaacgc	ttgaaatgcc	cttactgagt	attatgttaa	13200
aaaatggttt	tatcggtctg	gtaggggtatg	ggattgtttt	atataaactt	tatcgtaatg	13260
taagaatatt	aaaaacagat	aataataaaa	caataggaaa	gtctgtat t	atcattgtag	13320
tcctatctgc	aacagtagaa	aatttatattg	taaat ttaag	ttttgtat t	atgccaatat	13380
gtttttgttt	attaaattct	atatctacta	tggaatcaac	tattaacaaa	caactgcaaa	13440
cataaattgg	caggaataga	gttttgagtt	gctattaatt	tggtagagca	tatgttctat	13500
agggtggcaag	ataaagatag	tat ttttttac	atgatgat t	ttatgatagc	aaagcaagtt	13560
acggcataaa	aggaattaga	ggatggaaaa	agtcagcatt	attgtaccta	tttttaatac	13620
ggaaaagtac	ttaagagagt	gttttagatag	cattat tttcc	caatcgtata	ctaactctaga	13680
gattc ttttg	atagatgacg	gttcttcaga	ttcatcaacg	gatatatgtt	tggaatacgc	13740
agagcaagat	ggtagaataa	aact tttccg	gttaccaa at	ggtggtgttt	caaacgcaag	13800
gaattacggt	atcaaaaata	gcacagcaaa	ttatat tttatg	tttgtagatt	ctgatgatat	13860
tgttgacggc	aacattgttg	agtccttata	cacctgttta	aaagagaatg	atagtgat t	13920
gtcgggaggg	ttacttgcta	cttttgatgg	aaattatcaa	gaatctgagc	tgcaaaagtg	13980
tcaaattgat	ttggaagaga	taaaagaggt	gcgagactta	ggaaatgaaa	at tttcccaa	14040
tcattatatg	agcggtatct	ttaatagccc	ttgttgcaaa	ctttataaga	atatatatat	14100
aaaccaaggt	tttgacactg	aacagtgggt	aggagaggac	ttattat tta	atctaaatta	14160
tttaaagaat	ataaaaaaag	tccgctatgt	taacagaaat	ctttat ttttg	ccagaagaag	14220
tttacaaagt	actacaaata	cgttttaata	tgatgt tttt	attcaattag	aaaat ttaga	14280
agaaaaaact	tttgatttgt	ttgttaaaat	at ttggtgga	caatatgaat	tttctgtttt	14340
taaagagacg	ctacagtggc	atattat tta	ttatagctta	ttaatgttca	aaaatggaga	14400
tgaatcgctt	ccaaagaaat	tgcatatatt	taagtat tta	tacaataggc	attc ttttaga	14460
tactctaagt	attaaacgaa	cgtcctctgt	ttttaaaaga	atatgtaaat	taattgttgc	14520
taataatttg	tttaaaat t	ttttaaatac	tttaattagg	gaagaaaaaa	ataatgatta	14580
acatttctat	catcg tccca	at ttacaatg	ttgaacaata	tctatccaag	tgtataaata	14640
gcattgtaaa	tcagacctac	aaacatatag	agattcttct	ggtgaatgac	ggtagtacgg	14700
ataattcgga	agaaatttgt	ttagcatatg	cgaagaaaga	tagtcgcatt	cgttat tttta	14760
aaaaagagaa	cggcgggcta	tcagatgccc	gtaattatgg	cataagtcgc	gccaagggtg	14820

actacttagc	ttttatagac	tcagatgatt	ttattcattc	ggagttcatc	caacgtttac	14880
acgaagcaat	tgagagagag	aatgcccttg	tggcagttgc	tggttatgat	agggtagatg	14940
cttcggggca	tttcttaaca	gcagagccgc	ttcctacaaa	tcaggctggt	ctgagcggca	15000
ggaatgtttg	taaaaagctg	ctagaggcgg	atgggtcatcg	ctttgtggtg	gcctggaata	15060
aactctataa	aaaagaacta	tttgaagatt	ttcgatttga	aaagggtgtaag	attcatgaag	15120
atgaatactt	cacttatcgc	ttgctctatg	agttagaaaa	agttgcaata	gttaaggagt	15180
gcttgactta	ttatgttgac	cgagaaaata	gtatcataac	ttctagtatg	actgaccatc	15240
gcttccattg	cctactggaa	tttcaaaatg	aacgaatgga	cttctatgaa	agtagaggag	15300
ataaagagct	cttactagag	tggtatcgtt	cattttttagc	ctttgctggt	ttgttttttag	15360
gcaaataata	tcattgggtg	agcaaacagc	aaaagaagct	tctccaaacg	ctatttagaa	15420
ttgtatataa	acaattgaag	caaaataagc	gacttgcttt	actaatgaat	gcttattatt	15480
tggtaggggtg	tcttcatctt	aatttttagtg	tctttctgaa	aacggggaaa	gataaaaattc	15540
aagaaagatt	gagaagaagt	gaaagtagta	ctcggtaaga	atgttgtaat	aaatgggtga	15600
aagaaaaggg	gattaaaatg	aatccaacaa	atagtagaat	agcactcttt	gatacgatta	15660
aatgtatcat	ggtactttgt	gttattttta	cacatctgga	ttggctctgtt	gagcagcgtc	15720
aatggtttat	ctttccgtat	ttcgttgaca	tggctgttcc	aatttttctg	ttgctttctg	15780
cctattttctg	aacgaataag	tggaatacaa	aacaagagac	gctaaagctc	aagttcagca	15840
gtgggtataaa	agaaagtata	aacatgcttt	gtctctatgc	tatcgtgatg	gctgttaatg	15900
ttttattgag	ctattcgaga	accatctgat	aggagtaaag	cctttttcag	gttcttcatc	15960
gctccgttca	tttgctcctgt	ggctactttc	tggagaatcg	ggcccaggga	gttgggagtt	16020
actatgttcc	gttggttgatt	caggtagttt	ttttattacc	aattttgtat	gttcttttctg	16080
agaaaaataa	atgggtgggc	ttgcttactt	gttttttagt	aaacttttca	gtggatgccca	16140
tatttgctaa	catggctgaa	cacggcatat	atatatagac	taatatcact	tcgttatctt	16200
tttgttctag	ggcttggttt	tttctttcaa	agcaggatgt	gcgttccaag	gtagatactt	16260
tcattgcgac	cctatttggg	attattggag	caattctgat	ttttgtgaat	cattctatag	16320
agcccttctc	ctgggttttat	ggttggaagt	ctacttcctt	tctatgcgtc	ccatttgctg	16380
atgctatgct	attttttatg	ataaagtatg	gacagaagat	tccagcaata	ctgttgtcaa	16440
aattgggagt	tgcttcttat	catatctact	tgacccagat	gctgtatttt	tcagtagtcg	16500
caccattttt	agcagtgcaa	tttaagggtat	cttcgttgaa	tttgtggaac	ggcttggtta	16560
cctttctaat	ttgectgttt	gggtggctata	ttttctacaa	agtggatctg	tttatgagag	16620
tacgtggaaa	acgataatga	ctcattttcag	attagcagat	gccatttctg	ttattagcag	16680
attcgcgatgt	taatattccg	acaaagaaat	tcaaataagg	tgacgagaga	ggagtgggtat	16740
ctgtttctaa	accccagtat	ccccctttat	tttcaaagct	atatttatta	actgaacaag	16800
gagaattttt	aagagaactg	tttgtttaaat	cccagcacga	tctggttcga	aaggcttacc	16860
gaataaaaac	atgctatttt	tggacgggaa	acccatgatt	tttcacacga	ttgatgtggc	16920
aattgaatca	ggttgttttg	agaaagaaga	catctatgtc	agtacggatt	cagaaatgta	16980
taaggggggc	acctctataa	attcccaaaa	ttgcgaattt	ggagttacga	aagccttggt	17040
aatcaacat	cttaaatttt	agaaaattag	tttttagagg	tccccaaggg	gatttgcgag	17100
acaagaggca	tcaatgtatt	gttaagaccc	aaagaactat	ctacttatca	tactccatcg	17160
aatgaagtca	gtacgcactt	ttttacgaat	ctggatttta	tgaagattgt	atatttgttc	17220
ttctgcaagt	cacctcaccg	ttacggactg	gcgaacagat	aaaagaagcc	atgaatatgt	17280
acttacaggg	ggactcagaa	aatgtttttgc	atttcaatga	tgaagggcaa	gaaagagtga	17340
atcagtacat	tatcgaagct	gtacaggggt	tataaaaagg	ggttacttat	ccttaaagtc	17400
tgtatgtaga	aggagaaaaa	ttgagacgaa	tttatatttg	ccatacgatg	tatcagatcc	17460
tgatttccctt	gttaaagatg	gacgttgaga	gagatagttt	gatgtccggt	gatatcatcg	17520
ggcattttcc	agatgtcagg	gagcaactgc	agcagcatgt	tcatctaatc	gagggagacg	17580
gagcgttcat	ttgatctata	ttctttgata	gctagatcaa	aaacaaaaga	acgcctttcc	17640
ttgttacaga	gctatgacga	gggtgatcatt	tttcaagatc	accgtcaagt	cggtcatttt	17700
ttaaataaac	atcggattcc	ctattctctt	ttggaggatg	gttataattt	tttcaaggat	17760
aaaagagtgt	gcgatttgga	gtcaattcaa	tcatctgtct	ggaaaagact	cttttatcaa	17820
tggtatttta	aaccaacata	tttgattgggt	tcaagtctct	attgtcaatc	cattgaggtc	17880
aatgatctgt	cgctcgtaca	atttgactag	gcttataaac	cctttgtaga	agttccgaga	17940
aagcaattat	ttgatcaagc	atcgccagag	aagggtgcaag	cgctgctgca	gatatttgga	18000
gcaagggcga	tagtagcgga	tgaagagtct	tctcaaaaac	gattgctatt	attgaccag	18060
cccttgctctt	gggattatca	tgtgaccgaa	gagagttggt	ggagatttat	gtagcaggtc	18120

ttgcccctta	tcggaagac	tatacaatct	acataaaacc	gcacccacga	gatgggggttg	18180
attattcatt	tctgggtaag	gctgtgggtgc	ttctgcctca	aggatttccg	tttgagttgt	18240
tcgaaatggc	aggtaatatc	cgttttgata	tcggtatgac	ctatagttcg	tctgcttttag	18300
atTTTTTaaa	ttgttttgaa	gagaaagtgt	atttaaagga	cacttttctc	cttctttcaa	18360
aaaatgatat	tttgctgag	gggatagaat	aggaggattc	atgtctaaaa	aatcaatagt	18420
tgtctcaggt	ctcgtctata	cgattggaac	catcctcggt	cagggattag	ccttcattac	18480
cctccccatc	tatactcgtg	tcatttctca	ggaagtatat	gggcagttta	gcttgtataa	18540
ttcgtgggtg	gggctagttg	gtctctttat	cggctctacag	ttaggtgggg	cttttggccc	18600
gggatgggta	cacttccgcg	agaaatttga	tgatttcgta	tccaccttga	tgggtctcttc	18660
tatcgctttc	tttttaccaa	tttttgggct	atcttttctc	ctcagtcagc	ccctatcgct	18720
cctattttggt	ttgcctgatt	gggtcgttcc	gctttacttt	ttgcaaagtt	ttatgagtgt	18780
tgtgcaagga	ttttttacga	cctattttagt	gcagcggcag	cagtcctatgt	ggactttact	18840
cctatcggtg	ctgagcgtcg	ttatcaacac	tgctttatct	ttatttctca	tcttttctgat	18900
ggagaatgat	ttcatcgctc	gtgtaatggc	aaactcggca	acgactgggtg	tttttgcttg	18960
tgtgtccttg	ttgtttttct	ataagaagat	tgggcttcat	tttcgaaagg	actatcttcg	19020
gtatggttta	agtatatcga	ttcctcttat	ttttcatgga	ttaggtcata	atgtactcaa	19080
tcaatttgac	agaatcatgc	tcggcaagat	gctaacactg	tcagatgtag	ccctatacag	19140
tttcggctac	acacttgctg	ctatcttaca	aattgtgttt	tcgagcttga	atacgggatg	19200
gtgtccgtgg	tatttttgaga	aaaagagagg	tgagataaaa	gatttgctca	gttatgtccg	19260
ttactatctg	gcgattggcc	tgtttgtgac	ttttggattt	ctaacaattt	accctgaatt	19320
agcgatgttg	ttaggtggat	ctgagtatcg	tttcagtatg	ggatttattc	ccatgattat	19380
tgtcgggggtg	ttctttgtat	ttcttttatag	ttttccagcc	aatatccagt	tttatagtgg	19440
aaatacaaaag	tttttgccaa	ttgggtacttt	tatagcaggt	gtactaaata	tttccgtcca	19500
ctttgttttg	ataccgacaa	agaatttatg	gtgctgcttt	gcaacgactg	cttcctatct	19560
gttggttgcta	gtcttgcat	attttggtgc	taagaaaaag	tatgcttacg	atgaagttgc	19620
gatttcaaca	tttggttaagg	taattgctct	tgttgtcgtc	tatacaggct	tgatgacagt	19680
atttgtcgg	tcaatctgga	ttcgttggtc	actaggaata	gcggttctag	tcgttttatgc	19740
ctacattttt	agaaaggaat	taacagttgc	cctcaataca	ttcagggaaa	aacggtctaa	19800
ataagggcac	ctctataaac	tcccaaaatt	gcgaatttgg	agttacgaaa	gccttggttaa	19860
atcaaacatt	ttaaatttta	gaaaattagt	tttttagaggt	ccccatataa	aaacgtccca	19920
aatgagaggt	gctcataaga	attgaccatc	actgccatct	acccaaagtt	caagtattct	19980
ctaccatgaa	aattgtgcta	taatcaagta	taaagaagg	aatgtttctt	aaaggacgta	20040
tgcgcctctg	cttatgccag	aagtcatgag	gtaaatctcc	ctaaaaattg	ggtagaaaag	20100
cagattaaac	ttccaccaat	ctattgaaga	tcgtgttgaa	gagcaggctt	tagaagcaac	20160
aagccctgag	actattcgaa	agaaatctag	ggctattttt	tctaatcggc	tatcagaagt	20220
gaagtagcga	tctttattag	tggttctttta	ctacttaagg	aaaaccaagc	tgctccctca	20280
agactttatg	ggagcgattt	acagtcattt	ttagaaagga	aataaaatgg	tttatattat	20340
tgcagaaatt	ggttgtaatc	acaacggtga	tgttcatcta	gcacggaaaa	tggtagaagt	20400
tgccgttgat	tgtggtgtgg	atgccgttaa	atttcagaca	tttaaggcag	atttggtgat	20460
ttcaaaatac	gcaccaaagg	ccgaatacca	aaaaattaca	acaggagagt	cagattctca	20520
gctcgaaatg	actcgtcgtt	tggaattgag	ctttgaagag	tatcttgatt	tgctgatta	20580
ctgtcttgaa	aaggaggttg	atgtgttttc	gacacctttt	gatgaggaat	cattggactt	20640
cttgattagc	acagatatgc	ccgtttataa	gattccatct	ggtgagatta	ccaatcttcc	20700
ctatttgga	aaaattgggtc	gtcaagctaa	gaaagttatt	ctttcaactg	gtatggctgt	20760
tatggatgaa	attcatcaag	cgggtgaagat	tttgcaggaa	aatggaacga	ccgatatttc	20820
gattttgcat	tgtacaaccg	agtatccaac	cccttaccct	gctttgaatt	tgaatgtctt	20880
gcataccttg	aaaaaagaat	ttccaaactt	aacaattggc	tattcagacc	atagtgttgg	20940
ttcagaagta	cccatcgctg	ctgcagcaat	gggagctgaa	ttgattgaaa	agcactttac	21000
tctggacaat	gaaatggaag	gaccagatca	taaagcgagt	gctactcctg	atatcttagc	21060
agccttggtg	aaaggagtga	ggatagtggg	acaatctctt	ggtaaatttg	aaaaagagcc	21120
agaagaagtt	gaagtacgaa	ataaaattgt	agctagaaaa	tctattgttg	ccaaaaaagc	21180
aattgctaaa	ggcgaagtct	ttacagaaga	aaacatcact	gtcaaaagac	caggaaatgg	21240
aatttcgcca	atggaatgg	acaaagtctt	ggggcagggtg	agtgagcagg	attttgagga	21300
agacccaaat	atttgccata	gtgcttttga	aaatcaaagt	taagcggagt	aaggatgaaa	21360
aaaatttggt	ttgtgacagg	ctctcgtgcc	gaatatggga	ttatgcgtcg	cttattgagc	21420

tatctacagg	atgatccaga	aatggagctg	gatcttgtag	tgacagccat	gcatctagaa	21480
gaaaaatatg	ggatgacggg	caaagacatc	gaagcggaca	agcgtaggat	tgtcaagcgg	21540
attccattgc	atgtgacgga	tacgtctaag	cagacaatcg	tcaaactctt	agcgaccttg	21600
acagagcaac	tcacggttct	ttttgaagaa	gtccagtatg	acttggtggt	gattctgggg	21660
gatcgctatg	agatgctacc	agttgccaat	gctgcgttgc	tttataatat	tcctatttgc	21720
catattcatg	gtggtgaaaa	aaccatggga	aattttgatg	agtcgattcg	ccatgccatt	21780
accaagatga	gtcaccttca	tctgacatca	acggatgaat	ttagaaatcg	tgtcattcaa	21840
ctaggagaaa	atccaaccat	gtactgaaca	tcggagctat	gggtgttgaa	aatgttttaa	21900
aacaagactt	tttgacaaga	gaagagttgg	cgatggaact	tgggaattgat	tttgccgagg	21960
attactatgt	tgtactcttt	cacctgttta	ccttgaggga	taacacagcc	gaagaacaaa	22020
cgcaggcctt	attagatgct	ctaaaagaag	atggtagcca	gtgtttgata	attggatcca	22080
attcggatac	acatgccgat	aagataatgg	aattgatgca	tgaatttgta	aaacaagact	22140
ctgattctta	catctttact	tcgcttccaa	ctcgttatta	ccattccttg	gtcaagcatt	22200
cacaagggtt	aatagggaat	tcttcgtcag	gtttgattga	agtgcctcca	ttacagggttc	22260
cgaccttaaa	tattggaaat	cgccaatttg	gacgtttgtc	aggaccgagt	gtggtacatg	22320
ttggaacttc	taaggaagcg	attgttggtg	gtttggggca	attacgtgat	gtgatagatt	22380
ttaccaatcc	atgtgaacaa	cctgattctg	ctttacaagg	ttatcgagct	atcaaggaat	22440
ttttatctgt	acaggcctca	accatgaaag	agttttatga	tagatagggg	agaaagtttg	22500
atgaaaaaag	tagcctttct	aggagcgggt	accttttcag	atggtgtcct	tccttggttg	22560
gatagaactc	gatatgaact	cattggatat	tttgaagata	aaccgatcag	tgactatcgt	22620
ggctatcctg	tatttggtcc	cttgcaagat	gtcctaacct	atgttgatga	tggaaaagta	22680
gatgctgtct	tcgtcactat	aggtgacaat	gtcaagcgca	aggaaatctt	tgacttgctt	22740
gccaaagatc	attatgatgc	tttgttcaac	atcattagcg	agcaagccaa	tattttttcc	22800
ccagatagta	tcaagggacg	aggggttttc	atagggtttt	caagttttgt	aggagccgat	22860
tcctatgtct	atgacaattg	tatcatcaat	acgggtgcca	ttgtggaaca	tcataccacg	22920
gtggaggccc	attgtaacat	tactccagga	gtgaccataa	atggcttggtg	ccgtatcgga	22980
gaaagcactt	atattggaag	tggttcaaca	gtgattcaat	gtatcgagat	tgcaccttat	23040
acaacattgg	gggcagggac	agttgttttg	aaatcgttga	cggagtcagg	gacctatggt	23100
ggtgtacctg	ctagaaagat	taaataggtg	aattgatgga	accaatttgt	ctgattcctg	23160
ctcggtcagg	atcaaaaagg	ttaccaaaata	aaaacatggt	attttttagat	ggtgtaccga	23220
tgattttcca	taccattcga	gctgcgattg	agtctggatg	ttttaagaaa	gaaaatatat	23280
atgtcagtac	tgattcagag	gtttacaagg	aaatttggtg	aacaactggg	gttcaagtcc	23340
tcattgcgtcc	agctgacttg	gcgacagatt	ttacaacctc	ttttcaactg	aacgaacatt	23400
ttttacaaga	tttttctgat	gaccaagtat	ttgttctcct	gcaagttagc	tccccattaa	23460
gatcgggaaa	acatgtcaag	gaggcgatgg	agttatatgg	gaaagggtcaa	gctgaccacg	23520
ttgttagctt	taccaaagtc	gataagtctc	caacattggt	ttcaacttta	gacgaaaacg	23580
gattcgctaa	ggatattgca	ggattaggtg	gcagttatcg	tcgtcaagat	gagaaaacac	23640
tctactatcc	taatggagcg	atttatatatt	cttctaagca	ggcttatatta	gcggataaaa	23700
cttatttttc	tgaaaaaaca	gcggcctatg	tgatgacgaa	ggaagattcg	attgatgtag	23760
atgatcactt	tgatttttact	ggtgtttattg	gtcgaattta	ctttgattac	cagcgtcgtg	23820
agcaacaaaa	caaaccattt	tataaaagag	agttaaagcg	tttatgtgag	caacgagtcc	23880
atgatagtct	tgtgattggc	gatagtcgtc	tgttagcctt	gttactggat	ggttttcgata	23940
atatcagcat	cgggtgggatg	acagcttcga	cagcacttga	aaaccaagggt	ctcttttttg	24000
ctactccgat	aaagaaagtt	ttgctttctc	ttggtgtgaa	tgatttgatt	actgactatc	24060
ccttgcatat	gattgaggat	actattcgcc	agctgatgga	aagtcttggt	tccaaagcag	24120
agcagggttt	tgtgacgacg	attgcctaca	cgctgtttcg	tgatagcgtt	tccaatgaag	24180
aaattgtgca	gctgaatgac	gttattgttc	agtcagcaag	tgaactgggt	atttcagtga	24240
ttgatctaaa	tgaagttggt	gaaaaagagg	cgatgcttga	ctatcagtat	accaatgatg	24300
gattgcattt	caatcagatt	ggacaagagc	gtgtgaatca	gctgattttg	acaagtttga	24360
caagataatt	tggtgataga	agctatttca	gtggctagac	tatgttggtg	tgtgttttag	24420
agcccaggaa	taacatctgt	agaggatgct	agccttgaga	attgacaacc	atttagttgt	24480
tttaattata	taaggggacc	tctaaaaact	ccctaaattt	cccaaaaatg	agataataga	24540
ataaaaagta	atgaggagag	ctgtcatgca	tttattcaca	gacgatgaaa	aatcttgttc	24600
aaaactatca	gagaaaggca	atcccttaga	acgtttggat	gccgttatgg	attggaatat	24660
ctttcttcca	ttgttgtcag	agttattcag	tcgtaaagat	aaagtcatca	gtcgtggcgg	24720

```

tcgtcctcac ctagactatc tcatgatgtt caaagcgctc ttgcttcaac gtcttcataa 24780
cctatctgac gatgccatgg aatatcaact gctggatcgt atatcttttc gtcgttttgt 24840
tggttgatcat gaagacactg ttcccgatgc gaaaactatc tggctctatc gtgagaaatt 24900
aaccaagtca ggtcgtgaaa aggagttgtt cgatttggtc tatgcccac tcacagatga 24960
aggggtgatt gcccatcag gtcagattgt ggatgctacc tttgtcgaat gccctaaaca 25020
acgcaattca cgtgaggaca atcagaaaat caaaacttat cgaaaattat gaggtcacaa 25080
cagctagtgt acacgactcc aatgtcctag ctctcttttg tgatgccaat gaagcggttt 25140
ttgatgacag tgcttatgtt ggaaaatcag taccagaagg ttgtcgccac cacacgattc 25200
gtcgtgcttt tagaaataaa ccgttgactg agactgataa ggtcattaat cgacatatta 25260
ccaaagtcg ttgtcgcgtt gagcatgggt ttggcttcat tgaaactaac atgaaaggta 25320
acatctgtcg agcaattggg aaggcacgag ctgaaaccaa tgtgacctta accaacctgc 25380
tctacaatat ctgtcgtttt gagcaaatca aacgactggg attaccatcc gtgggcttag 25440
tgcgccc aaaataggaaa ataagcaaaa agaggctggg caaaaactag tttctcacia 25500
taaaaaaacg gctctttgtc aactgtagtg ggtagacgaa aagctaacac ctagagagga 25560
cgaaattcgt tctctcattt ttgatgttta aagcgtaacc gcctaataac aaggatatcta 25620
tccaatcaca cattcctcca ttatatagtt aaatgaaaca aaaacagtac atctatgata 25680
taatgtattt atggcatatt cattagattt tcgtaaaaaa gttctcgcac actgtgagaa 25740
aaccggcagt attactgaag catcagctat tttccaagtt tcacgtaaca ctatctatca 25800
atggctaaaa ttaaaagaga aaaccggcga gttcatcac caagttaaag gaaccaagcc 25860
aagaaaagtg gatagagata aattaaagaa ttatcttgaa actcatccag atgcttattt 25920
gactgaaata gcttctgaat ttgactgtca tccaacagct attcattacc ccctcaaagc 25980
tatgggatat actcgaaaaa aaagagctgt acctactatg aacaagaccc tgaaaaagta 26040
gaactgttcc ttaaagaatt gaataactta agccacttga ctctgtttta tattgacgag 26100
acagggtttg agacatatth tcatcgaaaa tatggtcgct ctttgaaagg tcagttgata 26160
aaaggtaagg tctctggaag aagataccag cggatatctt tagtagcagg tctcataaat 26220
gggtgcgctta tagccccgat gacatacaaa gatactatga cgagtggctt tttcgaagct 26280
t

```

<210> 10

<211> 239

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="ORF2Z amino acid sequence"

<400> 10

```

Ser Leu Asp Ile Asp His Met Met Glu Val Met Glu Ala Ser Lys Ser
 1             5             10             15
Ala Ala Gly Ser Ala Cys Pro Ser Pro Gln Ala Tyr Gln Ala Ala Phe
          20             25             30
Glu Gly Ala Glu Asn Ile Ile Val Val Thr Ile Thr Gly Gly Leu Ser
          35             40             45
Gly Ser Phe Asn Ala Ala Arg Val Ala Arg Asp Met Tyr Ile Glu Glu
          50             55             60
His Pro Asn Val Asn Ile His Leu Ile Asp Ser Leu Ser Ala Ser Gly
          65             70             75             80

```

92

Glu Met Asp Leu Leu Val His Gln Ile Asn Arg Leu Ile Ser Ala Gly
 85 90 95
 Leu Asp Phe Pro Gln Val Val Glu Ala Ile Thr His Tyr Arg Glu His
 100 105 110
 Ser Lys Leu Leu Phe Val Leu Ala Lys Val Asp Asn Leu Val Lys Asn
 115 120 125
 Gly Arg Leu Ser Lys Leu Val Gly Thr Val Val Gly Leu Leu Asn Ile
 130 135 140
 Arg Met Val Gly Glu Ala Ser Ala Glu Gly Lys Leu Glu Leu Leu Gln
 145 150 155 160
 Lys Ala Arg Gly His Lys Lys Ser Val Thr Ala Ala Phe Glu Glu Met
 165 170 175
 Lys Lys Ala Gly Tyr Asp Gly Gly Arg Ile Val Met Ala His Arg Asn
 180 185 190
 Asn Ala Lys Phe Phe Gln Gln Phe Ser Glu Leu Val Lys Ala Ser Phe
 195 200 205
 Pro Thr Ala Val Ile Asp Glu Val Ala Thr Ser Gly Leu Cys Ser Phe
 210 215 220
 Tyr Ala Glu Glu Gly Gly Leu Leu Met Gly Tyr Glu Val Lys Ala
 225 230 235

<210> 11

<211> 419

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="ORF2Y amino acid sequence"

<400> 11

Met Lys Lys Tyr Gln Val Ile Ile Gln Asp Ile Leu Thr Gly Ile Glu
 1 5 10 15
 Glu His Arg Phe Lys Arg Gly Glu Lys Leu Pro Ser Ile Arg Gln Leu
 20 25 30
 Arg Glu Gln Tyr His Cys Ser Lys Asp Thr Val Gln Lys Ala Met Leu
 35 40 45
 Glu Leu Lys Tyr Gln Asn Lys Ile Tyr Ala Val Glu Lys Ser Gly Tyr
 50 55 60
 Tyr Ile Leu Glu Asp Arg Asp Phe Gln Asp His Thr Cys Arg Ala Gln
 65 70 75 80

Ser	Tyr	Arg	Leu	Ser	Arg	Ile	Thr	Tyr	Glu	Asp	Phe	Arg	Ile	Cys	Leu
				85					90					95	
Lys	Glu	Ser	Leu	Ile	Gly	Arg	Glu	Asn	Tyr	Leu	Phe	Asn	Tyr	Tyr	His
			100					105					110		
Gln	Gln	Glu	Gly	Leu	Ala	Glu	Leu	Ile	Ser	Ser	Val	Gln	Ser	Leu	Leu
		115					120					125			
Met	Asp	Tyr	His	Val	Tyr	Thr	Lys	Lys	Asp	Gln	Leu	Val	Ile	Thr	Ala
	130					135					140				
Gly	Ser	Gln	Gln	Ala	Leu	Tyr	Ile	Leu	Thr	Gln	Met	Glu	Thr	Leu	Ala
145					150					155					160
Gly	Lys	Thr	Glu	Ile	Leu	Ile	Glu	Asn	Pro	Thr	Tyr	Ser	Arg	Met	Ile
			165					170						175	
Glu	Leu	Ile	Arg	His	Gln	Gly	Ile	Pro	Tyr	Gln	Thr	Ile	Glu	Arg	Asn
			180					185					190		
Leu	Asp	Gly	Ile	Asp	Leu	Glu	Glu	Leu	Glu	Ser	Ile	Phe	Gln	Thr	Gly
		195					200					205			
Lys	Ile	Lys	Phe	Phe	Tyr	Thr	Ile	Pro	Arg	Leu	His	Asn	Pro	Leu	Gly
	210					215					220				
Ser	Thr	Tyr	Asp	Ile	Ala	Thr	Lys	Thr	Ala	Ile	Val	Lys	Leu	Ala	Lys
225					230					235					240
Gln	Tyr	Asp	Val	Tyr	Ile	Ile	Glu	Asp	Asp	Tyr	Leu	Ala	Asp	Phe	Asp
				245				250						255	
Ser	Ser	His	Ser	Leu	Pro	Leu	His	Tyr	Leu	Asp	Thr	Asp	Asn	Arg	Val
			260					265					270		
Ile	Tyr	Ile	Lys	Ser	Phe	Thr	Pro	Thr	Leu	Phe	Pro	Ala	Leu	Arg	Ile
		275					280					285			
Gly	Ala	Ile	Ser	Leu	Pro	Asn	Gln	Leu	Arg	Asp	Ile	Phe	Ile	Lys	His
	290					295					300				
Lys	Ser	Leu	Ile	Asp	Tyr	Asp	Thr	Asn	Leu	Ile	Met	Gln	Lys	Ala	Leu
305				310						315					320
Ser	Leu	Tyr	Ile	Asp	Asn	Gly	Met	Phe	Ala	Arg	Asn	Thr	Gln	His	Leu
				325					330					335	
His	His	Ile	Tyr	His	Ala	Gln	Trp	Asn	Lys	Ile	Lys	Asp	Cys	Leu	Glu
			340					345					350		
Lys	Tyr	Ala	Leu	Asn	Ile	Pro	Tyr	Arg	Ile	Pro	Lys	Gly	Ser	Val	Thr
		355					360					365			

94

Phe Gln Leu Ser Lys Gly Ile Leu Ser Pro Ser Ile Gln His Met Phe
 370 375 380

Gly Lys Cys Tyr Tyr Phe Ser Gly Gln Lys Ala Asp Phe Leu Gln Ile
 385 390 395 400

Phe Phe Glu Gln Asp Phe Ala Asp Lys Leu Glu Gln Phe Val Arg Tyr
 405 410 415

Leu Asn Glu

<210> 12

<211> 244

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="ORF2X amino acid sequence"

<400> 12

Met Lys Ile Ile Ile Pro Asn Ala Lys Glu Val Asn Thr Asn Leu Glu
 1 5 10 15

Asn Ala Ser Phe Tyr Leu Leu Ser Asp Arg Ser Lys Pro Val Leu Asp
 20 25 30

Ala Ile Ser Gln Phe Asp Val Lys Lys Met Ala Ala Phe Tyr Lys Leu
 35 40 45

Asn Glu Ala Lys Ala Glu Leu Glu Ala Asp Arg Trp Tyr Arg Ile Arg
 50 55 60

Thr Gly Gln Ala Lys Thr Tyr Pro Ala Trp Gln Leu Tyr Asp Gly Leu
 65 70 75 80

Met Tyr Arg Tyr Met Asp Arg Arg Gly Ile Asp Ser Lys Glu Glu Asn
 85 90 95

Tyr Leu Arg Asp His Val Arg Val Ala Thr Ala Leu Tyr Gly Leu Ile
 100 105 110

His Pro Phe Glu Phe Ile Ser Pro His Arg Leu Asp Phe Gln Gly Ser
 115 120 125

Leu Lys Ile Gly Asn Gln Ser Leu Lys Gln Tyr Trp Arg Pro Tyr Tyr
 130 135 140

Asp Gln Glu Val Gly Asp Asp Glu Leu Ile Leu Ser Leu Ala Ser Ser
 145 150 155 160

Glu Phe Glu Gln Val Phe Ser Pro Gln Ile Gln Lys Arg Leu Val Lys
 165 170 175

95

Ile Leu Phe Met Glu Glu Lys Ala Gly Gln Leu Lys Val His Ser Thr
 180 185 190

Ile Ser Lys Lys Gly Arg Gly Arg Leu Leu Ser Trp Leu Ala Lys Asn
 195 200 205

Asn Ile Gln Glu Leu Ser Asp Ile Gln Asp Phe Lys Val Asp Gly Phe
 210 215 220

Glu Tyr Cys Thr Ser Glu Ser Thr Ala Asn Gln Leu Thr Phe Xaa Arg
 225 230 235 240

Ser Ile Lys Met

<210> 13

<211> 481

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS2A amino acid sequence"

<400> 13

Met Lys Lys Arg Ser Gly Arg Ser Lys Ser Ser Lys Phe Lys Leu Val
 1 5 10 15

Asn Phe Ala Leu Leu Gly Leu Tyr Ser Ile Thr Leu Cys Leu Phe Leu
 20 25 30

Val Thr Met Tyr Arg Tyr Asn Ile Leu Asp Phe Arg Tyr Leu Asn Tyr
 35 40 45

Ile Val Thr Leu Leu Leu Val Gly Val Ala Val Leu Ala Gly Leu Leu
 50 55 60

Met Trp Arg Lys Lys Ala Arg Ile Phe Thr Ala Leu Leu Leu Val Phe
 65 70 75 80

Ser Leu Val Ile Thr Ser Val Gly Ile Tyr Gly Met Gln Glu Val Val
 85 90 95

Lys Phe Ser Thr Arg Leu Asn Ser Asn Ser Thr Phe Ser Glu Tyr Glu
 100 105 110

Met Ser Ile Leu Val Pro Ala Asn Ser Asp Ile Thr Asp Val Arg Gln
 115 120 125

Leu Thr Ser Ile Leu Ala Pro Ala Glu Tyr Asp Gln Asp Asn Ile Thr
 130 135 140

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

96

Ser Pro Gly Thr Ser Tyr Leu Thr Ala Tyr Gln Ser Met Leu Asn Gly
 165 170 175
 Glu Ser Gln Ala Met Val Phe Asn Gly Val Phe Thr Asn Ile Leu Glu
 180 185 190
 Asn Glu Asp Pro Gly Phe Ser Ser Lys Val Lys Lys Ile Tyr Ser Phe
 195 200 205
 Lys Val Thr Gln Thr Val Glu Thr Ala Thr Lys Gln Val Ser Gly Asp
 210 215 220
 Ser Phe Asn Ile Tyr Ile Ser Gly Ile Asp Ala Tyr Gly Pro Ile Ser
 225 230 235 240
 Thr Val Ser Arg Ser Asp Val Asn Ile Ile Met Thr Val Asn Arg Ala
 245 250 255
 Thr His Lys Ile Leu Leu Thr Thr Thr Pro Arg Asp Ser Tyr Val Ala
 260 265 270
 Phe Ala Asp Gly Gly Gln Asn Gln Tyr Asp Lys Leu Thr His Ala Gly
 275 280 285
 Ile Tyr Gly Val Asn Ala Ser Val His Thr Leu Glu Asn Phe Tyr Gly
 290 295 300
 Ile Asp Ile Ser Asn Tyr Val Arg Leu Asn Phe Ile Ser Phe Leu Gln
 305 310 315 320
 Leu Ile Asp Leu Val Gly Gly Ile Asp Val Tyr Asn Asp Gln Glu Phe
 325 330 335
 Thr Ser Leu His Gly Asn Tyr His Phe Pro Val Gly Gln Val His Leu
 340 345 350
 Asn Ser Asp Gln Ala Leu Gly Phe Val Arg Glu Arg Tyr Ser Leu Thr
 355 360 365
 Gly Gly Asp Asn Asp Arg Gly Lys Asn Gln Glu Lys Val Ile Ala Ala
 370 375 380
 Leu Ile Lys Lys Met Ser Thr Pro Glu Asn Leu Lys Asn Tyr Gln Ala
 385 390 395 400
 Ile Leu Ser Gly Leu Glu Gly Ser Ile Gln Thr Asp Leu Ser Leu Glu
 405 410 415
 Thr Ile Met Ser Leu Val Asn Thr Gln Leu Glu Ser Gly Thr Gln Phe
 420 425 430
 Thr Val Glu Ser Gln Ala Leu Thr Gly Thr Gly Arg Ser Asp Leu Ser
 435 440 445

97

Ser Tyr Ala Met Pro Gly Ser Gln Leu Tyr Met Met Glu Ile Asn Gln
 450 455 460

Asp Ser Leu Glu Gln Ser Lys Ala Ala Ile Gln Ser Val Leu Val Glu
 465 470 475 480

Lys

<210> 14

<211> 229

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS2B amino acid sequence"

<400> 14

Met Asn Asn Gln Glu Val Asn Ala Ile Glu Ile Asp Val Leu Phe Leu
 1 5 10 15

Leu Lys Thr Ile Trp Arg Lys Lys Phe Leu Ile Leu Leu Thr Ala Val
 20 25 30

Leu Thr Ala Gly Leu Ala Phe Val Tyr Ser Ser Phe Leu Val Thr Pro
 35 40 45

Gln Tyr Asp Ser Thr Thr Arg Ile Tyr Val Val Ser Gln Asn Val Glu
 50 55 60

Ala Gly Ala Gly Leu Thr Asn Gln Glu Leu Gln Ala Gly Thr Tyr Leu
 65 70 75 80

Ala Lys Asp Tyr Arg Glu Ile Ile Leu Ser Gln Asp Val Leu Thr Gln
 85 90 95

Val Ala Thr Glu Leu Asn Leu Lys Glu Ser Leu Lys Glu Lys Ile Ser
 100 105 110

Val Ser Ile Pro Val Asp Thr Arg Ile Val Ser Ile Ser Val Arg Asp
 115 120 125

Ala Asp Pro Asn Glu Ala Ala Arg Ile Ala Asn Ser Leu Arg Thr Phe
 130 135 140

Ala Val Gln Lys Val Val Glu Val Thr Lys Val Ser Asp Val Thr Thr
 145 150 155 160

Leu Glu Glu Ala Val Pro Ala Glu Glu Pro Thr Thr Pro Asn Thr Lys
 165 170 175

Arg Asn Ile Leu Leu Gly Leu Leu Ala Gly Gly Ile Leu Ala Thr Gly
 180 185 190

98

Leu Val Leu Val Met Glu Val Leu Asp Asp Arg Val Lys Arg Pro Gln
 195 200 205

Asp Ile Glu Glu Val Met Gly Leu Thr Leu Leu Gly Ile Val Pro Asp
 210 215 220

Ser Lys Lys Leu Lys
 225

<210> 15

<211> 225

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS2C amino acid sequence"

<400> 15

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

Thr Glu Glu Tyr Phe Asn Ala Ile Arg Thr Asn Ile Gln Leu Ser Gly
 20 25 30

Ala Asp Ile Lys Val Val Gly Ile Thr Ser Val Lys Ser Asn Glu Gly
 35 40 45

Lys Ser Thr Thr Ala Ala Ser Leu Ala Ile Ala Tyr Ala Arg Ser Gly
 50 55 60

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

Gly Phe Phe Lys Pro Ile Thr Lys Ile Thr Gly Leu Thr Asp Tyr Leu
 85 90 95

Ala Gly Thr Thr Asp Leu Ser Gln Gly Leu Cys Asp Thr Asp Ile Pro
 100 105 110

Asn Leu Thr Val Ile Glu Ser Gly Lys Val Ser Pro Asn Pro Thr Ala
 115 120 125

Leu Leu Gln Ser Lys Asn Phe Glu Asn Leu Leu Ala Thr Leu Arg Arg
 130 135 140

Tyr Tyr Asp Tyr Val Ile Val Asp Cys Pro Pro Leu Gly Leu Val Ile
 145 150 155 160

Asp Ala Ala Ile Ile Ala Gln Lys Cys Asp Ala Met Val Ala Val Val
 165 170 175

Glu Ala Gly Asn Val Lys Cys Ser Ser Leu Lys Lys Val Lys Glu Gln
 180 185 190

99

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

Asp Ile Ala Thr Glu Lys Tyr Ser Glu Tyr Gly Asn Tyr Gly Lys Lys
 210 215 220

Ala
 225

<210> 16
 <211> 243
 <212> PRT
 <213> Streptococcus suis

<220>
 <223> /note="CPS2D amino acid sequence"

<400> 16

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

Lys Thr Ile Glu Glu Ser Leu Ser Leu Ile Ser Glu Ala Tyr Arg Gln
 20 25 30

Gly Val Arg Tyr Ile Val Ala Thr Ser His Arg Arg Lys Gly Met Phe
 35 40 45

Glu Thr Pro Glu Lys Ile Ile Met Ile Asn Phe Leu Gln Leu Lys Glu
 50 55 60

Ala Val Ala Glu Val Tyr Pro Glu Ile Arg Leu Cys Tyr Gly Ala Glu
 65 70 75 80

Leu Tyr Tyr Ser Lys Asp Ile Leu Ser Lys Leu Glu Lys Lys Lys Val
 85 90 95

Pro Thr Leu Asn Gly Ser Cys Tyr Ile Leu Leu Glu Phe Ser Thr Asp
 100 105 110

Thr Pro Trp Lys Glu Ile Gln Glu Ala Val Asn Glu Met Thr Leu Leu
 115 120 125

Gly Leu Thr Pro Val Leu Ala His Ile Glu Arg Tyr Asp Ala Leu Ala
 130 135 140

Phe Gln Ser Glu Arg Val Glu Lys Leu Ile Asp Lys Gly Cys Tyr Thr
 145 150 155 160

Gln Val Asn Ser Asn His Val Leu Lys Pro Ala Leu Ile Gly Glu Arg
 165 170 175

Ala Lys Glu Phe Lys Lys Arg Thr Arg Tyr Phe Leu Glu Gln Asp Leu
 180 185 190

100

Val His Cys Val Ala Ser Asp Met His Asn Leu Tyr Ser Arg Pro Pro
 195 200 205

Phe Met Arg Glu Ala Tyr Gln Leu Val Lys Lys Glu Tyr Gly Glu Asp
 210 215 220

Arg Ala Lys Ala Leu Phe Lys Lys Asn Pro Leu Leu Ile Leu Lys Asn
 225 230 235 240

Gln Val Gln

<210> 17

<211> 459

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS2E amino acid sequence"

<400> 17

Met Asn Ile Glu Ile Gly Tyr Arg Gln Thr Lys Leu Ala Leu Phe Asp
 1 5 10 15

Met Ile Ala Val Thr Ile Ser Ala Ile Leu Thr Ser His Ile Pro Asn
 20 25 30

Ala Asp Leu Asn Arg Ser Gly Ile Phe Ile Ile Met Met Val His Tyr
 35 40 45

Phe Ala Phe Phe Ile Ser Arg Met Pro Val Glu Phe Glu Tyr Arg Gly
 50 55 60

Asn Leu Ile Glu Phe Glu Lys Thr Phe Asn Tyr Ser Ile Ile Phe Val
 65 70 75 80

Ile Phe Leu Met Ala Val Ser Phe Met Leu Glu Asn Asn Phe Ala Leu
 85 90 95

Ser Arg Arg Gly Ala Val Tyr Phe Thr Leu Ile Asn Phe Val Leu Val
 100 105 110

Tyr Leu Phe Asn Val Ile Ile Lys Gln Phe Lys Asp Ser Phe Leu Phe
 115 120 125

Ser Thr Thr Tyr Gln Lys Lys Thr Ile Leu Ile Thr Thr Ala Glu Leu
 130 135 140

Trp Glu Asn Met Gln Val Leu Phe Glu Ser Asp Ile Leu Phe Gln Lys
 145 150 155 160

Asn Leu Val Ala Leu Val Ile Leu Gly Thr Glu Ile Asp Lys Ile Asn
 165 170 175

Leu	Pro	Leu	Pro	Leu	Tyr	Tyr	Ser	Val	Glu	Glu	Ala	Ile	Gly	Phe	Ser	
			180				185						190			
Thr	Arg	Glu	Val	Val	Asp	Tyr	Val	Phe	Ile	Asn	Leu	Pro	Ser	Glu	Tyr	
			195				200						205			
Phe	Asp	Leu	Lys	Gln	Leu	Val	Ser	Asp	Phe	Glu	Leu	Leu	Gly	Ile	Asp	
			210				215						220			
Val	Gly	Val	Asp	Ile	Asn	Ser	Phe	Gly	Phe	Thr	Val	Leu	Lys	Asn	Lys	
225						230						235			240	
Lys	Ile	Gln	Met	Leu	Gly	Asp	His	Ser	Ile	Val	Thr	Phe	Ser	Thr	Asn	
			245						250						255	
Phe	Tyr	Lys	Pro	Ser	His	Ile	Trp	Met	Lys	Arg	Leu	Leu	Asp	Ile	Leu	
			260						265						270	
Gly	Ala	Val	Val	Gly	Leu	Ile	Ile	Ser	Gly	Ile	Val	Ser	Ile	Leu	Leu	
			275						280						285	
Ile	Pro	Ile	Ile	Arg	Arg	Asp	Gly	Gly	Pro	Ala	Ile	Phe	Ala	Gln	Lys	
			290						295						300	
Arg	Val	Gly	Gln	Asn	Gly	Arg	Ile	Phe	Thr	Phe	Tyr	Lys	Phe	Arg	Ser	
305						310						315			320	
Met	Phe	Val	Asp	Ala	Glu	Val	Arg	Lys	Lys	Glu	Leu	Met	Ala	Gln	Asn	
			325						330						335	
Gln	Met	Gln	Gly	Gly	Met	Phe	Lys	Met	Asp	Asn	Asp	Pro	Arg	Ile	Thr	
			340						345						350	
Pro	Ile	Gly	His	Phe	Ile	Arg	Lys	Thr	Ser	Leu	Asp	Glu	Leu	Pro	Gln	
			355						360						365	
Phe	Tyr	Asn	Val	Leu	Ile	Gly	Asp	Met	Ser	Leu	Val	Gly	Thr	Arg	Pro	
			370						375						380	
Pro	Thr	Val	Asp	Glu	Phe	Glu	Lys	Tyr	Thr	Pro	Ser	Gln	Lys	Arg	Arg	
385						390						395			400	
Leu	Ser	Phe	Lys	Pro	Gly	Ile	Thr	Gly	Leu	Trp	Gln	Val	Ser	Gly	Arg	
			405						410						415	
Ser	Asp	Ile	Thr	Asp	Phe	Asn	Glu	Val	Val	Arg	Leu	Asp	Leu	Thr	Tyr	
			420						425						430	
Ile	Asp	Asn	Trp	Thr	Ile	Trp	Ser	Asp	Ile	Lys	Ile	Leu	Leu	Lys	Thr	
			435						440						445	
Val	Lys	Val	Val	Leu	Leu	Arg	Glu	Gly	Gly	Gln						
450						455										

```
<210> 18
<211> 389
<212> PRT
<213> Streptococcus suis
```

```
<220>  
<223> /note="CPS2F amino acid sequence"
```

<400> 18

Met 1	Arg	Thr	Val	Tyr 5	Ile	Ile	Gly	Ser	Lys 10	Gly	Ile	Pro	Ala	Lys 15	Tyr
Gly	Gly	Phe	Glu 20	Thr	Phe	Val	Glu	Lys 25	Leu	Thr	Glu	Tyr	Gln 30	Lys	Asp
Lys	Ser	Ile 35	Asn	Tyr	Phe	Val	Ala 40	Cys	Thr	Arg	Glu	Asn 45	Ser	Ala	Lys
Ser	Asp 50	Ile	Thr	Gly	Glu	Val 55	Phe	Glu	His	Asn	Gly 60	Ala	Thr	Cys	Phe
Asn 65	Ile	Asp	Val	Pro	Asn 70	Ile	Gly	Ser	Ala	Lys 75	Ala	Ile	Leu	Tyr	Asp 80
Ile	Met	Ala	Leu	Lys 85	Lys	Ser	Ile	Glu	Ile 90	Ala	Lys	Asp	Arg	Asn 95	Asp
Thr	Ser	Pro	Ile 100	Phe	Tyr	Ile	Leu	Ala 105	Cys	Arg	Ile	Gly	Pro 110	Phe	Ile
Tyr	Leu	Phe 115	Lys	Lys	Gln	Ile	Glu	Ser 120	Ile	Gly	Gly	Gln 125	Leu	Phe	Val
Asn 130	Pro	Asp	Gly	His	Glu	Trp 135	Leu	Arg	Glu	Lys 140	Trp	Ser	Tyr	Pro	Val
Arg 145	Gln	Tyr	Trp	Lys	Phe 150	Ser	Glu	Ser	Leu	Met 155	Leu	Lys	Tyr	Ala	Asp 160
Leu	Leu	Ile	Cys 165	Asp	Ser	Lys	Asn	Ile	Glu 170	Lys	Tyr	Ile	His	Glu	Asp 175
Tyr	Arg	Lys 180	Tyr	Ala	Pro	Glu	Thr	Ser 185	Tyr	Ile	Ala	Tyr	Gly 190	Thr	Asp
Leu	Asp 195	Lys	Ser	Arg	Leu	Ser	Pro 200	Thr	Asp	Ser	Val	Val 205	Arg	Glu	Trp
Tyr 210	Lys	Glu	Lys	Glu	Ile	Ser 215	Glu	Asn	Asp	Tyr	Tyr 220	Leu	Val	Val	Gly
Arg 225	Phe	Val	Pro	Glu	Asn 230	Asn	Tyr	Glu	Val	Met 235	Ile	Arg	Glu	Phe	Met 240

Arg Lys Tyr Phe Asn Pro Lys Gly Ile Phe Asp Tyr Phe Ile Ser Tyr
65 70 75 80

104

His	His	Tyr	Ser	Lys	Gln	Ile	Ala	Gln	Tyr	Ala	Ile	Glu	Asn	Lys	Val	85	90	95	
Asp	Ile	Ile	His	Asn	Asn	Thr	Thr	Ala	Val	Leu	Glu	Gly	Ile	Tyr	Leu	100	105	110	
Lys	Arg	Lys	Leu	Lys	Leu	Pro	Leu	Leu	Trp	His	Val	His	Glu	Ile	Ile	115	120	125	
Val	Lys	Pro	Lys	Phe	Ile	Ser	Asp	Ser	Ile	Asn	Phe	Leu	Met	Gly	Arg	130	135	140	
Phe	Ala	Asp	Lys	Ile	Val	Thr	Val	Ser	Gln	Ala	Val	Ala	Asn	His	Ile	145	150	155	160
Lys	Gln	Ser	Pro	His	Ile	Lys	Asp	Asp	Gln	Ile	Ser	Val	Ile	Tyr	Asn	165	170	175	
Gly	Val	Asp	Asn	Lys	Val	Phe	Tyr	Gln	Ser	Asp	Ala	Arg	Ser	Val	Arg	180	185	190	
Glu	Arg	Phe	Asp	Ile	Asp	Glu	Glu	Ala	Leu	Val	Ile	Gly	Met	Val	Gly	195	200	205	
Arg	Val	Asn	Ala	Trp	Lys	Gly	Gln	Gly	Asp	Phe	Leu	Glu	Ala	Val	Ala	210	215	220	
Pro	Ile	Leu	Glu	Gln	Asn	Pro	Lys	Ala	Ile	Ala	Phe	Ile	Ala	Gly	Ser	225	230	235	240
Ala	Phe	Glu	Gly	Glu	Glu	Trp	Arg	Val	Val	Glu	Leu	Glu	Lys	Lys	Ile	245	250	255	
Ser	Gln	Leu	Lys	Val	Ser	Ser	Gln	Val	Xaa	Arg	Met	Asp	Tyr	Tyr	Ala	260	265	270	
Asn	Thr	Thr	Glu	Leu	Tyr	Asn	Met	Phe	Asp	Ile	Phe	Val	Leu	Pro	Ser	275	280	285	
Thr	Asn	Pro	Asp	Pro	Leu	Pro	Thr	Val	Val	Leu	Lys	Ala	Met	Ala	Cys	290	295	300	
Gly	Lys	Pro	Val	Val	Gly	Tyr	Arg	His	Gly	Gly	Val	Cys	Glu	Met	Val	305	310	315	320
Lys	Glu	Gly	Val	Asn	Gly	Phe	Leu	Val	Thr	Pro	Asn	Ser	Pro	Leu	Asn	325	330	335	
Leu	Ser	Lys	Val	Ile	Leu	Gln	Leu	Ser	Glu	Asn	Ile	Asn	Leu	Arg	Lys	340	345	350	
Lys	Ile	Gly	Asn	Asn	Ser	Ile	Glu	Arg	Gln	Lys	Glu	His	Phe	Ser	Leu	355	360	365	

105

Lys Ser Tyr Val Lys Asn Phe Ser Lys Val Tyr Thr Ser Leu Lys Val
 370 375 380

Tyr
 385

<210> 20

<211> 456

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS2H amino acid sequence"

<400> 20

Met Lys Ile Ile Ser Phe Thr Met Val Asn Asn Glu Ser Glu Ile Ile
 1 5 10 15

Glu Ser Phe Ile Arg Tyr Asn Tyr Asn Phe Ile Asp Glu Met Val Ile
 20 25 30

Ile Asp Asn Gly Cys Thr Asp Asn Thr Met Gln Ile Ile Phe Asn Leu
 35 40 45

Ile Lys Glu Gly Tyr Lys Ile Ser Val Tyr Asp Glu Ser Leu Glu Ala
 50 55 60

Tyr Asn Gln Tyr Arg Leu Asp Asn Lys Tyr Leu Thr Lys Ile Ile Ala
 65 70 75 80

Glu Lys Asn Pro Asp Leu Ile Ile Pro Leu Asp Ala Asp Glu Phe Leu
 85 90 95

Thr Ala Asp Ser Asn Pro Arg Lys Leu Leu Glu Gln Leu Asp Leu Glu
 100 105 110

Lys Ile His Tyr Val Asn Trp Gln Trp Phe Val Met Thr Lys Lys Asp
 115 120 125

Asp Ile Asn Asp Ser Phe Ile Pro Arg Arg Met Gln Tyr Cys Phe Glu
 130 135 140

Lys Pro Val Trp His His Ser Asp Gly Lys Pro Val Thr Lys Cys Ile
 145 150 155 160

Ile Ser Ala Lys Tyr Tyr Lys Lys Met Asn Leu Lys Leu Ser Met Gly
 165 170 175

His His Thr Val Phe Gly Asn Pro Asn Val Arg Ile Glu His His Asn
 180 185 190

Asp Leu Lys Phe Ala His Tyr Arg Ala Ile Ser Gln Glu Gln Leu Ile
 195 200 205

106

Tyr Lys Thr Ile Cys Tyr Thr Ile Arg Asp Ile Ala Thr Met Glu Asn
 210 215 220
 Asn Ile Glu Thr Ala Gln Arg Thr Asn Gln Met Ala Leu Ile Glu Ser
 225 230 235 240
 Gly Val Asp Met Trp Glu Thr Ala Arg Glu Ala Ser Tyr Ser Gly Tyr
 245 250 255
 Asp Cys Asn Val Ile His Ala Pro Ile Asp Leu Ser Phe Cys Lys Glu
 260 265 270
 Asn Ile Val Ile Lys Tyr Asn Glu Leu Ser Arg Glu Thr Val Ala Glu
 275 280 285
 Arg Val Met Lys Thr Gly Arg Glu Met Ala Val Arg Ala Tyr Asn Val
 290 295 300
 Glu Arg Lys Gln Lys Glu Lys Lys Phe Leu Lys Pro Ile Ile Phe Val
 305 310 315 320
 Leu Asp Gly Leu Lys Gly Asp Glu Tyr Ile His Pro Asn Pro Ser Asn
 325 330 335
 His Leu Thr Ile Leu Thr Glu Met Tyr Asn Val Arg Gly Leu Leu Thr
 340 345 350
 Asp Asn His Gln Ile Lys Phe Leu Lys Val Asn Tyr Arg Leu Ile Ile
 355 360 365
 Thr Pro Asp Phe Ala Lys Phe Leu Pro His Glu Phe Ile Val Val Pro
 370 375 380
 Asp Thr Xaa Asp Ile Glu Gln Val Lys Ser Gln Tyr Val Gly Thr Gly
 385 390 395 400
 Val Asp Leu Ser Lys Ile Ile Ser Leu Lys Glu Tyr Arg Lys Glu Ile
 405 410 415
 Gly Phe Ile Gly Asn Leu Tyr Ala Leu Leu Gly Phe Val Pro Asn Met
 420 425 430
 Leu Asn Arg Ile Tyr Leu Tyr Ile Gln Arg Asn Gly Ile Ala Asn Thr
 435 440 445
 Ile Ile Lys Ile Lys Ser Arg Leu
 450 455

<210> 21

<211> 410

<212> PRT

<213> Streptococcus suis

```
<220>
<223> /note="CPS2I amino acid sequence"
```

Met	Gln	Ala	Asp	Arg	Arg	Lys	Thr	Phe	Gly	Lys	Met	Arg	Ile	Arg	Ile	
1				5					10						15	
Asn	Asn	Leu	Phe	Phe	Val	Ala	Ile	Ala	Phe	Met	Gly	Ile	Ile	Ile	Ile	Ser
			20					25					30			
Asn	Ser	Gln	Val	Val	Leu	Ala	Ile	Gly	Lys	Ala	Ser	Val	Ile	Gln	Tyr	
		35					40					45				
Leu	Ser	Tyr	Leu	Val	Leu	Ile	Leu	Cys	Ile	Val	Asn	Asp	Leu	Leu	Lys	
	50					55					60					
Asn	Asn	Lys	His	Ile	Val	Val	Tyr	Lys	Leu	Gly	Tyr	Leu	Phe	Leu	Ile	
65					70					75					80	
Ile	Phe	Leu	Phe	Thr	Ile	Gly	Ile	Cys	Gln	Gln	Ile	Leu	Pro	Ile	Thr	
				85					90					95		
Thr	Lys	Ile	Tyr	Leu	Ser	Ile	Ser	Met	Met	Ile	Ile	Ser	Val	Leu	Ala	
			100					105					110			
Thr	Leu	Pro	Ile	Ser	Leu	Ile	Lys	Asp	Ile	Asp	Asp	Phe	Arg	Arg	Ile	
		115					120					125				
Ser	Asn	His	Leu	Leu	Phe	Ala	Leu	Phe	Ile	Thr	Ser	Ile	Leu	Gly	Ile	
	130					135					140					
Lys	Met	Gly	Ala	Thr	Met	Phe	Thr	Gly	Ala	Val	Glu	Gly	Ile	Gly	Phe	
145					150					155					160	
Ser	Gln	Gly	Phe	Asn	Gly	Gly	Leu	Thr	His	Lys	Asn	Phe	Phe	Gly	Ile	
				165					170					175		
Thr	Ile	Leu	Met	Gly	Phe	Val	Leu	Thr	Tyr	Leu	Ala	Tyr	Lys	Tyr	Gly	
			180					185					190			
Ser	Tyr	Lys	Arg	Thr	Asp	Arg	Phe	Ile	Leu	Gly	Leu	Glu	Leu	Phe	Leu	
		195					200					205				
Ile	Leu	Ile	Ser	Asn	Thr	Arg	Ser	Val	Tyr	Leu	Ile	Leu	Leu	Leu	Phe	
	210					215					220					
Leu	Phe	Leu	Val	Asn	Leu	Asp	Lys	Ile	Lys	Ile	Glu	Gln	Arg	Gln	Trp	
225					230					235					240	
Ser	Thr	Leu	Lys	Tyr	Ile	Ser	Met	Leu	Phe	Cys	Ala	Ile	Phe	Leu	Tyr	
				245					250					255		
Tyr	Phe	Phe	Gly	Phe	Leu	Ile	Thr	His	Ser	Asp	Ser	Tyr	Ala	His	Arg	
			260					265					270			

108

Val Asn Gly Leu Ile Asn Phe Phe Glu Tyr Tyr Arg Asn Asp Trp Phe
 275 280 285
 His Leu Met Phe Gly Ala Ala Asp Leu Ala Tyr Gly Asp Leu Thr Leu
 290 295 300
 Asp Tyr Ala Ile Arg Val Arg Arg Val Leu Gly Trp Asn Gly Thr Leu
 305 310 315 320
 Glu Met Pro Leu Leu Ser Ile Met Leu Lys Asn Gly Phe Ile Gly Leu
 325 330 335
 Val Gly Tyr Gly Ile Val Leu Tyr Lys Leu Tyr Arg Asn Val Arg Ile
 340 345 350
 Leu Lys Thr Asp Asn Ile Lys Thr Ile Gly Lys Ser Val Phe Ile Ile
 355 360 365
 Val Val Leu Ser Ala Thr Val Glu Asn Tyr Ile Val Asn Leu Ser Phe
 370 375 380
 Val Phe Met Pro Ile Cys Phe Cys Leu Leu Asn Ser Ile Ser Thr Met
 385 390 395 400
 Glu Ser Thr Ile Asn Lys Gln Leu Gln Thr
 405 410

<210> 22

<211> 332

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS2J amino acid sequence"

<400> 22

Met Glu Lys Val Ser Ile Ile Val Pro Ile Phe Asn Thr Glu Lys Tyr
 1 5 10 15
 Leu Arg Glu Cys Leu Asp Ser Ile Ile Ser Gln Ser Tyr Thr Asn Leu
 20 25 30
 Glu Ile Leu Leu Ile Asp Asp Gly Ser Ser Asp Ser Ser Thr Asp Ile
 35 40 45
 Cys Leu Glu Tyr Ala Glu Gln Asp Gly Arg Ile Lys Leu Phe Arg Leu
 50 55 60
 Pro Asn Gly Gly Val Ser Asn Ala Arg Asn Tyr Gly Ile Lys Asn Ser
 65 70 75 80
 Thr Ala Asn Tyr Ile Met Phe Val Asp Ser Asp Asp Ile Val Asp Gly
 85 90 95

109

```

Asn Ile Val Glu Ser Leu Tyr Thr Cys Leu Lys Glu Asn Asp Ser Asp
      100                      105                      110

Leu Ser Gly Gly Leu Leu Ala Thr Phe Asp Gly Asn Tyr Gln Glu Ser
      115                      120                      125

Glu Leu Gln Lys Cys Gln Ile Asp Leu Glu Glu Ile Lys Glu Val Arg
      130                      135                      140

Asp Leu Gly Asn Glu Asn Phe Pro Asn His Tyr Met Ser Gly Ile Phe
145                      150                      155                      160

Asn Ser Pro Cys Cys Lys Leu Tyr Lys Asn Ile Tyr Ile Asn Gln Gly
      165                      170                      175

Phe Asp Thr Glu Gln Trp Leu Gly Glu Asp Leu Leu Phe Asn Leu Asn
      180                      185                      190

Tyr Leu Lys Asn Ile Lys Lys Val Arg Tyr Val Asn Arg Asn Leu Tyr
      195                      200                      205

Phe Ala Arg Arg Ser Leu Gln Ser Thr Thr Asn Thr Phe Lys Tyr Asp
      210                      215                      220

Val Phe Ile Gln Leu Glu Asn Leu Glu Glu Lys Thr Phe Asp Leu Phe
225                      230                      235                      240

Val Lys Ile Phe Gly Gly Gln Tyr Glu Phe Ser Val Phe Lys Glu Thr
      245                      250                      255

Leu Gln Trp His Ile Ile Tyr Tyr Ser Leu Leu Met Phe Lys Asn Gly
      260                      265                      270

Asp Glu Ser Leu Pro Lys Lys Leu His Ile Phe Lys Tyr Leu Tyr Asn
      275                      280                      285

Arg His Ser Leu Asp Thr Leu Ser Ile Lys Arg Thr Ser Ser Val Phe
      290                      295                      300

Lys Arg Ile Cys Lys Leu Ile Val Ala Asn Asn Leu Phe Lys Ile Phe
305                      310                      315                      320

Leu Asn Thr Leu Ile Arg Glu Glu Lys Asn Asn Asp
      325                      330

```

```

<210> 23
<211> 334
<212> PRT
<213> Streptococcus suis

```

```

<220>
<223> /note="CPS2K amino acid sequence"

```

```

<400> 23

```

110

Met	Ile	Asn	Ile	Ser	Ile	Ile	Val	Pro	Ile	Tyr	Asn	Val	Glu	Gln	Tyr	1	5	10	15
Leu	Ser	Lys	Cys	Ile	Asn	Ser	Ile	Val	Asn	Gln	Thr	Tyr	Lys	His	Ile	20	25	30	
Glu	Ile	Leu	Leu	Val	Asn	Asp	Gly	Ser	Thr	Asp	Asn	Ser	Glu	Glu	Ile	35	40	45	
Cys	Leu	Ala	Tyr	Ala	Lys	Lys	Asp	Ser	Arg	Ile	Arg	Tyr	Phe	Lys	Lys	50	55	60	
Glu	Asn	Gly	Gly	Leu	Ser	Asp	Ala	Arg	Asn	Tyr	Gly	Ile	Ser	Arg	Ala	65	70	75	80
Lys	Gly	Asp	Tyr	Leu	Ala	Phe	Ile	Asp	Ser	Asp	Asp	Phe	Ile	His	Ser	85	90	95	
Glu	Phe	Ile	Gln	Arg	Leu	His	Glu	Ala	Ile	Glu	Arg	Glu	Asn	Ala	Leu	100	105	110	
Val	Ala	Val	Ala	Gly	Tyr	Asp	Arg	Val	Asp	Ala	Ser	Gly	His	Phe	Leu	115	120	125	
Thr	Ala	Glu	Pro	Leu	Pro	Thr	Asn	Gln	Ala	Val	Leu	Ser	Gly	Arg	Asn	130	135	140	
Val	Cys	Lys	Lys	Leu	Leu	Glu	Ala	Asp	Gly	His	Arg	Phe	Val	Val	Ala	145	150	155	160
Trp	Asn	Lys	Leu	Tyr	Lys	Lys	Glu	Leu	Phe	Glu	Asp	Phe	Arg	Phe	Glu	165	170	175	
Lys	Gly	Lys	Ile	His	Glu	Asp	Glu	Tyr	Phe	Thr	Tyr	Arg	Leu	Leu	Tyr	180	185	190	
Glu	Leu	Glu	Lys	Val	Ala	Ile	Val	Lys	Glu	Cys	Leu	Tyr	Tyr	Tyr	Val	195	200	205	
Asp	Arg	Glu	Asn	Ser	Ile	Ile	Thr	Ser	Ser	Met	Thr	Asp	His	Arg	Phe	210	215	220	
His	Cys	Leu	Leu	Glu	Phe	Gln	Asn	Glu	Arg	Met	Asp	Phe	Tyr	Glu	Ser	225	230	235	240
Arg	Gly	Asp	Lys	Glu	Leu	Leu	Leu	Glu	Cys	Tyr	Arg	Ser	Phe	Leu	Ala	245	250	255	
Phe	Ala	Val	Leu	Phe	Leu	Gly	Lys	Tyr	Asn	His	Trp	Leu	Ser	Lys	Gln	260	265	270	
Gln	Lys	Lys	Leu	Leu	Gln	Thr	Leu	Phe	Arg	Ile	Val	Tyr	Lys	Gln	Leu	275	280	285	

111

Lys Gln Asn Lys Arg Leu Ala Leu Leu Met Asn Ala Tyr Tyr Leu Val
 290 295 300

Gly Cys Leu His Leu Asn Phe Ser Val Phe Leu Lys Thr Gly Lys Asp
 305 310 315 320

Lys Ile Gln Glu Arg Leu Arg Arg Ser Glu Ser Ser Thr Arg
 325 330

<210> 24

<211> 467

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS20 amino acid sequence"

<400> 24

Met Ser Lys Lys Ser Ile Val Val Ser Gly Leu Val Tyr Thr Ile Gly
 1 5 10 15

Thr Ile Leu Val Gln Gly Leu Ala Phe Ile Thr Leu Pro Ile Tyr Thr
 20 25 30

Arg Val Ile Ser Gln Glu Val Tyr Gly Gln Phe Ser Leu Tyr Asn Ser
 35 40 45

Trp Val Gly Leu Val Gly Leu Phe Ile Gly Leu Gln Leu Gly Gly Ala
 50 55 60

Phe Gly Pro Gly Trp Val His Phe Arg Glu Lys Phe Asp Asp Phe Val
 65 70 75 80

Ser Thr Leu Met Val Ser Ser Ile Ala Phe Phe Leu Pro Ile Phe Gly
 85 90 95

Leu Ser Phe Leu Leu Ser Gln Pro Leu Ser Leu Leu Phe Gly Leu Pro
 100 105 110

Asp Trp Val Val Pro Leu Ile Phe Leu Gln Ser Leu Met Ile Val Val
 115 120 125

Gln Gly Phe Phe Thr Thr Tyr Leu Val Gln Arg Gln Gln Ser Met Trp
 130 135 140

Thr Leu Pro Leu Ser Val Leu Ser Ala Val Ile Asn Thr Ala Leu Ser
 145 150 155 160

Leu Phe Leu Thr Phe Pro Met Glu Asn Asp Phe Ile Ala Arg Val Met
 165 170 175

Ala Asn Pro Ala Thr Thr Gly Val Leu Ala Cys Val Ser Xaa Trp Phe
 180 185 190

112

Ser Gln Lys Lys Asn Gly Leu His Phe Arg Lys Asp Tyr Leu Arg Tyr
 195 200 205
 Gly Leu Ser Ile Ser Ile Pro Leu Ile Phe His Gly Leu Gly His Asn
 210 215 220
 Val Leu Asn Gln Phe Asp Arg Ile Met Leu Gly Lys Met Leu Thr Leu
 225 230 235 240
 Ser Asp Val Ala Leu Tyr Ser Phe Gly Tyr Thr Leu Ala Ser Ile Leu
 245 250 255
 Gln Ile Val Phe Ser Ser Leu Asn Thr Val Trp Cys Pro Trp Tyr Phe
 260 265 270
 Glu Lys Lys Arg Gly Ala Asp Lys Asp Leu Leu Ser Tyr Val Arg Tyr
 275 280 285
 Tyr Leu Ala Ile Gly Leu Phe Val Thr Phe Gly Phe Leu Thr Ile Tyr
 290 295 300
 Pro Glu Leu Ala Met Leu Leu Gly Gly Ser Glu Tyr Arg Phe Ser Met
 305 310 315 320
 Gly Phe Ile Pro Met Ile Ile Val Gly Val Phe Phe Val Phe Leu Tyr
 325 330 335
 Ser Phe Pro Ala Asn Ile Gln Phe Tyr Ser Gly Asn Thr Lys Phe Leu
 340 345 350
 Pro Ile Gly Thr Phe Ile Ala Gly Val Leu Asn Ile Ser Val His Phe
 355 360 365
 Val Leu Ile Pro Thr Lys Asn Leu Trp Cys Cys Phe Ala Thr Thr Ala
 370 375 380
 Ser Tyr Leu Leu Leu Leu Val Leu His Tyr Phe Val Ala Lys Lys Lys
 385 390 395 400
 Tyr Ala Tyr Asp Glu Val Ala Ile Ser Thr Phe Val Lys Val Ile Ala
 405 410 415
 Leu Val Val Val Tyr Thr Gly Leu Met Thr Val Phe Val Gly Ser Ile
 420 425 430
 Trp Ile Arg Trp Ser Leu Gly Ile Ala Val Leu Val Val Tyr Ala Tyr
 435 440 445
 Ile Phe Arg Lys Glu Leu Thr Val Ala Leu Asn Thr Phe Arg Glu Lys
 450 455 460
 Arg Ser Lys
 465

113

<210> 25
 <211> 338
 <212> PRT
 <213> Streptococcus suis

<220>
 <223> /note="CPS2P amino acid sequence"

<400> 25

Met	Val	Tyr	Ile	Ile	Ala	Glu	Ile	Gly	Cys	Asn	His	Asn	Gly	Asp	Val	1	5	10	15
His	Leu	Ala	Arg	Lys	Met	Val	Glu	Val	Ala	Val	Asp	Cys	Gly	Val	Asp	20	25	30	
Ala	Val	Lys	Phe	Gln	Thr	Phe	Lys	Ala	Asp	Leu	Leu	Ile	Ser	Lys	Tyr	35	40	45	
Ala	Pro	Lys	Ala	Glu	Tyr	Gln	Lys	Ile	Thr	Thr	Gly	Glu	Ser	Asp	Ser	50	55	60	
Gln	Leu	Glu	Met	Thr	Arg	Arg	Leu	Glu	Leu	Ser	Phe	Glu	Glu	Tyr	Leu	65	70	75	80
Asp	Leu	Arg	Asp	Tyr	Cys	Leu	Glu	Lys	Gly	Val	Asp	Val	Phe	Ser	Thr	85	90	95	
Pro	Phe	Asp	Glu	Glu	Ser	Leu	Asp	Phe	Leu	Ile	Ser	Thr	Asp	Met	Pro	100	105	110	
Val	Tyr	Lys	Ile	Pro	Ser	Gly	Glu	Ile	Thr	Asn	Leu	Pro	Tyr	Leu	Glu	115	120	125	
Lys	Ile	Gly	Arg	Gln	Ala	Lys	Lys	Val	Ile	Leu	Ser	Thr	Gly	Met	Ala	130	135	140	
Val	Met	Asp	Glu	Ile	His	Gln	Ala	Val	Lys	Ile	Leu	Gln	Glu	Asn	Gly	145	150	155	160
Thr	Thr	Asp	Ile	Ser	Ile	Leu	His	Cys	Thr	Thr	Glu	Tyr	Pro	Thr	Pro	165	170	175	
Tyr	Pro	Ala	Leu	Asn	Leu	Asn	Val	Leu	His	Thr	Leu	Lys	Lys	Glu	Phe	180	185	190	
Pro	Asn	Leu	Thr	Ile	Gly	Tyr	Ser	Asp	His	Ser	Val	Gly	Ser	Glu	Val	195	200	205	
Pro	Ile	Ala	Ala	Ala	Ala	Met	Gly	Ala	Glu	Leu	Ile	Glu	Lys	His	Phe	210	215	220	
Thr	Leu	Asp	Asn	Glu	Met	Glu	Gly	Pro	Asp	His	Lys	Ala	Ser	Ala	Thr	225	230	235	240

Gln Met

```
<210> 26
<211> 170
<212> PRT
<213> Streptococcus suis

<220>
<223> /note="CPS2Q amino acid sequence"

<400> 26
```

Met 1	Lys	Lys	Ile	Cys 5	Phe	Val	Thr	Gly	Ser 10	Arg	Ala	Glu	Tyr	Gly 15	Ile
Met	Arg	Arg	Leu 20	Leu	Ser	Tyr	Leu	Gln 25	Asp	Asp	Pro	Glu	Met 30	Glu	Leu
Asp	Leu	Val 35	Val	Thr	Ala	Met	His 40	Leu	Glu	Glu	Lys	Tyr 45	Gly	Met	Thr
Val 50	Lys	Asp	Ile	Glu	Ala	Asp 55	Lys	Arg	Arg	Ile	Val 60	Lys	Arg	Ile	Pro
Leu 65	His	Leu	Thr	Asp	Thr 70	Ser	Lys	Gln	Thr	Ile 75	Val	Lys	Ser	Leu	Ala 80
Thr	Leu	Thr	Glu	Gln 85	Leu	Thr	Val	Leu	Phe 90	Glu	Glu	Val	Gln	Tyr 95	Asp
Leu	Val	Leu	Ile 100	Leu	Gly	Asp	Arg	Tyr 105	Glu	Met	Leu	Pro	Val 110	Ala	Asn
Ala	Ala	Leu	Leu	Tyr	Asn	Ile	Pro 120	Ile	Cys	His	Ile	His 125	Gly	Gly	Glu

115

Lys Thr Met Gly Asn Phe Asp Glu Ser Ile Arg His Ala Ile Thr Lys
 130 135 140

Met Ser His Leu His Leu Thr Ser Thr Asp Glu Phe Arg Asn Arg Val
 145 150 155 160

Ile Gln Leu Gly Glu Asn Pro Thr Met Tyr
 165 170

<210> 27

<211> 184

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS2R amino acid sequence"

<400> 27

Met Glu Leu Gly Ile Asp Phe Ala Glu Asp Tyr Tyr Val Val Leu Phe
 1 5 10 15

His Pro Val Thr Leu Glu Asp Asn Thr Ala Glu Glu Gln Thr Gln Ala
 20 25 30

Leu Leu Asp Ala Leu Lys Glu Asp Gly Ser Gln Cys Leu Ile Ile Gly
 35 40 45

Ser Asn Ser Asp Thr His Ala Asp Lys Ile Met Glu Leu Met His Glu
 50 55 60

Phe Val Lys Gln Asp Ser Asp Ser Tyr Ile Phe Thr Ser Leu Pro Thr
 65 70 75 80

Arg Tyr Tyr His Ser Leu Val Lys His Ser Gln Gly Leu Ile Gly Asn
 85 90 95

Ser Ser Ser Gly Leu Ile Glu Val Pro Ser Leu Gln Val Pro Thr Leu
 100 105 110

Asn Ile Gly Asn Arg Gln Phe Gly Arg Leu Ser Gly Pro Ser Val Val
 115 120 125

His Val Gly Thr Ser Lys Glu Ala Ile Val Gly Gly Leu Gly Gln Leu
 130 135 140

Arg Asp Val Ile Asp Phe Thr Asn Pro Phe Glu Gln Pro Asp Ser Ala
 145 150 155 160

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

Thr Met Lys Glu Phe Tyr Asp Arg
 180

116

<210> 28
 <211> 208
 <212> PRT
 <213> Streptococcus suis

<220>
 <223> /note="CPS2S amino acid sequence"

<400> 28

Met	Lys	Lys	Val	Ala	Phe	Leu	Gly	Ala	Gly	Thr	Phe	Ser	Asp	Gly	Val	1	5	10	15
Leu	Pro	Trp	Leu	Asp	Arg	Thr	Arg	Tyr	Glu	Leu	Ile	Gly	Tyr	Phe	Glu	20	25	30	
Asp	Lys	Pro	Ile	Ser	Asp	Tyr	Arg	Gly	Tyr	Pro	Val	Phe	Gly	Pro	Leu	35	40	45	
Gln	Asp	Val	Leu	Thr	Tyr	Leu	Asp	Asp	Gly	Lys	Val	Asp	Ala	Val	Phe	50	55	60	
Val	Thr	Ile	Gly	Asp	Asn	Val	Lys	Arg	Lys	Glu	Ile	Phe	Asp	Leu	Leu	65	70	75	80
Ala	Lys	Asp	His	Tyr	Asp	Ala	Leu	Phe	Asn	Ile	Ile	Ser	Glu	Gln	Ala	85	90	95	
Asn	Ile	Phe	Ser	Pro	Asp	Ser	Ile	Lys	Gly	Arg	Gly	Val	Phe	Ile	Gly	100	105	110	
Phe	Ser	Ser	Phe	Val	Gly	Ala	Asp	Ser	Tyr	Val	Tyr	Asp	Asn	Cys	Ile	115	120	125	
Ile	Asn	Thr	Gly	Ala	Ile	Val	Glu	His	His	Thr	Thr	Val	Glu	Ala	His	130	135	140	
Cys	Asn	Ile	Thr	Pro	Gly	Val	Thr	Ile	Asn	Gly	Leu	Cys	Arg	Ile	Gly	145	150	155	160
Glu	Ser	Thr	Tyr	Ile	Gly	Ser	Gly	Ser	Thr	Val	Ile	Gln	Cys	Ile	Glu	165	170	175	
Ile	Ala	Pro	Tyr	Thr	Thr	Leu	Gly	Ala	Gly	Thr	Val	Val	Leu	Lys	Ser	180	185	190	
Leu	Thr	Glu	Ser	Gly	Thr	Tyr	Val	Gly	Val	Pro	Ala	Arg	Lys	Ile	Lys	195	200	205	

Met 1	Glu	Pro	Ile	Cys 5	Leu	Ile	Pro	Ala	Arg 10	Ser	Gly	Ser	Lys	Gly 15	Leu
Pro	Asn	Lys	Asn 20	Met	Leu	Phe	Leu	Asp 25	Gly	Val	Pro	Met	Ile 30	Phe	His
Thr	Ile	Arg 35	Ala	Ala	Ile	Glu	Ser 40	Gly	Cys	Phe	Lys	Lys 45	Glu	Asn	Ile
Tyr	Val 50	Ser	Thr	Asp	Ser	Glu 55	Val	Tyr	Lys	Glu	Ile 60	Cys	Glu	Thr	Thr
Gly 65	Val	Gln	Val	Leu	Met 70	Arg	Pro	Ala	Asp	Leu 75	Ala	Thr	Asp	Phe	Thr 80
Thr	Ser	Phe	Gln	Leu 85	Asn	Glu	His	Phe	Leu 90	Gln	Asp	Phe	Ser	Asp 95	Asp
Gln	Val	Phe	Val 100	Leu	Leu	Gln	Val	Thr 105	Ser	Pro	Leu	Arg	Ser 110	Gly	Lys
His	Val 115	Lys	Glu	Ala	Met	Glu	Leu 120	Tyr	Gly	Lys	Gly 125	Gln	Ala	Asp	His
Val 130	Val	Ser	Phe	Thr	Lys	Val 135	Asp	Lys	Ser	Pro	Thr 140	Leu	Phe	Ser	Thr
Leu 145	Asp	Glu	Asn	Gly	Phe 150	Ala	Lys	Asp	Ile	Ala 155	Gly	Leu	Gly	Gly	Ser 160
Tyr	Arg	Arg	Gln	Asp 165	Glu	Lys	Thr	Leu	Tyr 170	Tyr	Pro	Asn	Gly	Ala 175	Ile
Tyr	Ile	Ser	Ser 180	Lys	Gln	Ala	Tyr	Leu 185	Ala	Asp	Lys	Thr	Tyr 190	Phe	Ser
Glu	Lys 195	Thr	Ala	Ala	Tyr	Val	Met 200	Thr	Lys	Glu	Asp	Ser 205	Ile	Asp	Val
Asp 210	Asp	His	Phe	Asp	Phe	Thr 215	Gly	Val	Ile	Gly	Arg 220	Ile	Tyr	Phe	Asp
Tyr 225	Gln	Arg	Arg	Glu	Gln 230	Gln	Asn	Lys	Pro	Phe 235	Tyr	Lys	Arg	Glu	Leu 240

118

Lys Arg Leu Cys Glu Gln Arg Val His Asp Ser Leu Val Ile Gly Asp
 245 250 255
 Ser Arg Leu Leu Ala Leu Leu Leu Asp Gly Phe Asp Asn Ile Ser Ile
 260 265 270
 Gly Gly Met Thr Ala Ser Thr Ala Leu Glu Asn Gln Gly Leu Phe Leu
 275 280 285
 Ala Thr Pro Ile Lys Lys Val Leu Leu Ser Leu Gly Val Asn Asp Leu
 290 295 300
 Ile Thr Asp Tyr Pro Leu His Met Ile Glu Asp Thr Ile Arg Gln Leu
 305 310 315 320
 Met Glu Ser Leu Val Ser Lys Ala Glu Gln Val Phe Val Thr Thr Ile
 325 330 335
 Ala Tyr Thr Leu Phe Arg Asp Ser Val Ser Asn Glu Glu Ile Val Gln
 340 345 350
 Leu Asn Asp Val Ile Val Gln Ser Ala Ser Glu Leu Gly Ile Ser Val
 355 360 365
 Ile Asp Leu Asn Glu Val Val Glu Lys Glu Ala Met Leu Asp Tyr Gln
 370 375 380
 Tyr Thr Asn Asp Gly Leu His Phe Asn Gln Ile Gly Gln Glu Arg Val
 385 390 395 400
 Asn Gln Leu Ile Leu Thr Ser Leu Thr Arg
 405 410

<210> 30

<211> 6992

<212> DNA

<213> Streptococcus suis

<220>

<223> /note="CPS1 nucleotide sequence"

<400> 30

```

atcgccaaac gaaattggca ttatttgata tgatagcagt tgcaatttct gcaatcttaa      60
caagtcatat accaaatgct gattttaaadc gttctggaat ttttatcata atgatggttc      120
attattttgc attttttata tctcgtatgc cagttgaatt tgagtataga ggtaatctga      180
tagagtttga aaaaacattt aactatagta taatatttgc aatttttctt acggcagtat      240
catttttgtt ggagaataat ttcgcacttt caagacgtgg tgccgtgtat ttcacattaa      300
taaacttcgt tttggtatac ctattttaacg taattattaa gcagttttaag gatagctttc      360
tattttcgac aatctatcaa aaaaagacga ttctaattac aacggctgaa cgatgggaaa      420
atatgcaagt tttatttgaa tcacataaac aaattcaaaa aaatcttggt gcattggtag      480
ttttaggtac agaaatagat aaaattaatt tatcattacc gctctattat tctgtggaag      540
aagctataga gttttcaaca agggaagtgg tcgaccacgt ctttataaat ctaccaagtg      600
agttttttaga cgtaaagcaa ttcgtttcag atttttgagtt gttaggtatt gatgtaagcg      660

```

ttgatattaa	ttcattcggg	tttactgcgt	tgaaaaacaa	aaaaatccaa	ctgctaggtg	720
accatagcat	tgtaactttt	tccacaaatt	tttataagcc	tagtcatatc	atgatgaaac	780
gactttttgga	tatactcgga	gcggtagtcg	ggttaattat	ttgtgggtata	gtttctattt	840
tgtagttcc	aattattcgt	agagatgggtg	gaccggctat	ttttgctcag	aaacgagttg	900
gacagaatgg	acgcataatt	acattctaca	agtttcgata	gatgtatggt	gatgctgagg	960
agcgcaaaaa	agacttgctc	agccaaaacc	agatgcaagg	gtgggtatgt	tttaaaatgg	1020
gaaaaacgat	cctagaatta	ctccaattgg	acatttcata	cgcaaaaaca	agtttagacg	1080
agttaccaca	gtttttataat	gttttaattg	gcgatatgag	tctagttggg	acacgtccac	1140
ctacagttga	tgaatttgaa	aaatatactc	ctgggtcaaaa	gagacgattg	agtttttaaac	1200
cagggattac	aggtctctgg	caggttagtg	gtcgtagtaa	tatcacagac	ttcgacgacg	1260
tagttcgggt	ggacttagca	tacattgata	attggactat	ctgggtcagat	attaaaattt	1320
tattaaagac	agtgaaggtt	gtattgttga	gagaggggag	taagtaaaag	tatatgaaag	1380
tttgtttggg	cggttcttca	gggggacatt	tgactcactt	gtatttggtt	aaaccgtttt	1440
ggaaggaaga	agaacgtttt	tgggtaacat	ttgataaaga	ggatgcaaga	agtcttttga	1500
agaatgaaaa	aatgtatcca	tggttactttc	caacaaatcg	caatctcatt	aatttagtga	1560
aaaatacttt	cttagctttc	aaaattttac	gtgatgagaa	accagatggt	attatttcat	1620
ctgggtgcggc	cgttgctgctc	cccttctttt	acatcggaaa	actattttgga	gcaaagacga	1680
tttatattga	agtatttgat	cgagttaata	aatctacatt	aactggaaaa	ctagttttatc	1740
ccgtaacaga	tattttttatt	gttcagtggg	aagaaatgaa	gaaggatat	cctaaatcta	1800
ttaacttggg	gagtattttt	taatgatttt	tgtaacagta	ggaactcatg	aacaacagtt	1860
taatcgattg	ataaaagaga	ttgattttatt	gaaaaaaaat	ggaagtataa	ccgacgaaat	1920
atttattcaa	acaggatatt	ctgactatat	tccagaatat	tgcaagtata	aaaaatttct	1980
cagttacaaa	gaaatggaac	aatatattaa	caaatcagaa	gtagttattt	gccacggagg	2040
ccccgctact	tttatgaatt	cattatccaa	aggaaaaaaa	caattattgt	ttcctagaca	2100
aaaaaagtat	ggtgaacatg	taaatgatca	tcaagtagag	tttgtaagaa	gaattttaca	2160
agataataat	atttttattta	tagaaaatat	agatgatttg	tttgaaaaaa	ttattgaagt	2220
ttctaagcaa	actaacttta	catcaaataa	taattttttt	tgtgaaagat	taaaacaaat	2280
agttgaaaaa	tttaatgagg	atcaagaaaa	tgaataataa	aaaagatgca	tatttgataa	2340
tggcttatca	taatttttct	cagattttac	tggagagggg	tacagatatt	atcatcttct	2400
ctcaggagaa	tgcacaccat	tagttccttc	agaataacctg	tataattatt	ttaaatattc	2460
tcaggattta	tatgttgaat	ttacaaaaga	tgagcaaaaa	tataaagaaa	ataggatata	2520
tgaacgagtt	aaatgttaca	gattattttc	taatatatca	gaaaaaacta	ttgataatgt	2580
actgttttaga	atttttattaa	gaatgtatcg	agctttttgaa	tactattttac	aaagattggt	2640
gtttattgat	agaataaaaa	acatgggtcta	agaataagat	ttgggttctaa	ttgggtttcg	2700
cttccacatg	atttttgtggc	aattcttttta	tcaaatagaa	acgaaacagc	ttattttattt	2760
aagtaatcta	aatgtccaga	tgaactattt	atacagacaa	ttatagaaaa	atatgaattt	2820
tcaaatagat	tatctaaata	tggaaatttta	agatatataa	agtggaaaaa	atcaacatct	2880
tctcctattg	tctttacaga	tgattctatt	gatgaattgc	taaatgcaag	aaatttaggt	2940
tttttatttg	ctagaaagtt	aaaaatagaa	aataaatcta	aatttaaaga	aattattact	3000
aaaaaataaa	atagttgatt	ttgtgagagt	aatgtatggt	taaattattt	aaatatgacc	3060
cggaatattt	tattttttaag	tacttctggg	tgattatttt	tattccagag	caaaagtatg	3120
tattttttatt	aattttttatg	aattttaattt	tatttcatat	aaaatttttg	aaaactaagc	3180
taatattaaa	aaatgaaatt	ttattgtttt	tattatgggc	tatattatgt	tttgtttcag	3240
tagtcacaag	tatgtttggt	gaaataaatt	ttgaaagatt	atttgcagat	tttactgctc	3300
ccataatttg	gattattgca	ataatgtatt	ataatttgta	ttcattttata	aatattgatt	3360
ataaaaaatt	aaaaaatagt	atcttttttta	gtttttttagt	tttattaggt	atatctgcat	3420
tgtatattat	tcaaaatggg	aaagatatgt	tattttttaga	cagacacctt	ataggactag	3480
actatcttat	aacaggcgctc	aaaacaagggt	tgggtggctt	tatgaactat	cctacgttaa	3540
ataccactac	aattatagtt	tcaattccgt	taatctttgc	acttataaaa	aataaaaatgc	3600
aacaattttt	tttcttgtgt	cttgcttttta	taccgatcta	tttaagtgga	tcgagaattg	3660
gtagtttatc	gctagcaata	ttaattatat	gcttggttatg	gagatatata	ggtggaaaat	3720
ttgcttggat	aaaaaagcta	atagtaatat	ttgtaatact	acttattatt	ttaaatactg	3780
aattgcttta	ccatgaaatt	ttggctgttt	ataattctag	agaatcaagt	aacgaagcta	3840
gatttattat	ttatcaagga	agtattgata	aagtattaga	aaacaatatt	ttatttggat	3900
atggaatatc	cgaatattca	gttacgggaa	cttggctcgg	aagtcattca	ggctatatat	3960

catttttttta	taaatcagga	atagttgggt	tgatttttact	gatgttttct	tttttttatg	4020
ttataaaaaa	aagttatgga	gttaatgggg	aaacagcact	attttatttt	acatcattag	4080
ccatatTTTT	catatatgaa	acaatagatc	cgattattat	tatattagta	ctattctttt	4140
cttcaatagg	tatttggaat	aatataaatt	ttaaaaagga	tatggagaca	aaaaatgaat	4200
gatttaattt	cagttattgt	accaatttat	aatgtccaag	attatcttga	taaatgtatt	4260
aacagtatta	ttaaccaaac	atatacta	at	tttagaggta	ttctcgtaaa	4320
actgatgatt	ctgagaaaat	ttgcttaaac	tatatgaaga	acgatggaag	aattaaatat	4380
tacaagaaaa	ttaatggcgg	tctagcagat	gctcgaaatt	tcggactaga	acatgcaaca	4440
ggtaaatata	ttgcttttgt	cgattctgat	gactatatag	aagttgcaat	gttcgagaga	4500
atgcatgata	atataactga	gtataatgcc	gatatagcag	agatagattt	ttgttttagta	4560
gacgaaaacg	ggtatacaaa	gaaaaaaaaga	aatagtaatt	ttcatgtctt	aacgagagaa	4620
gagactgtaa	aagaattttt	gtcaggatct	aatatagaaa	ataatgtttg	gtgcaagctt	4680
tattcacgag	atattataaa	agatataaaa	ttccaaatta	ataatagaag	tattggtgag	4740
gatttgcttt	ttaatttgga	ggtcttgaac	aatgtaacac	gtgtagtagt	tgatactaga	4800
gaatattatt	ataattatgt	cattcgtaac	agttcgctta	ttaatcagaa	attctctata	4860
aataatattg	atttagtcac	aagattggag	aattaccctt	ttaagttaaa	aagagagttt	4920
agtcattatt	ttgatgcaaa	agttattaaa	gagaagggtta	aatgttttaa	caaaatgtat	4980
tcaacagatt	gtttggataa	tgagttcttg	ccaatattag	agtcttatcg	aaaagaaata	5040
cgtagatatc	cattttattaa	agcgaaaaga	tatttatcaa	gaaagcattt	agttacgttg	5100
tatttgatga	aatttttcgcc	taaactatat	gtaatgttat	ataagaaatt	tcaaaagcag	5160
tagaggtaaa	aatggataaa	attagtgtta	ttgttccagt	ttataatgta	gataaatatt	5220
taagtagttg	tatagaaagc	attattaatc	aaaattataa	aaatatagaa	atattattga	5280
tagatgatgg	ctctgtagat	gattctgcta	aaatatgcaa	ggaatatgca	gaaaaagata	5340
aaagagtaaa	aatttttttc	actaatcata	gtggagtatc	aaatgctaga	aatcatggaa	5400
taaagcggag	tacagctgaa	tatattatgt	ttgttgactc	tgatgatggt	gttgatagta	5460
gattagtaga	aaaattatat	tttaataatta	taaaaagtag	aagtgattta	tctggttggt	5520
tgtacgctac	tttttcagaa	aatataaata	at	tttgaagt	gaataatcca	5580
ttgaagcaat	taataccgtg	caggacatgg	gagaaaaaaa	ttttatgaat	ttgtatataa	5640
ataatatttt	ttctactcct	gtttgtaaac	tatataagaa	aagatacata	acagatcttt	5700
ttcaagagaa	tcaatgggtta	ggagaagatt	tacttttttaa	tctgcattat	ttaaagaata	5760
tagatagagt	tagttatttg	actgaacatc	tttatttttta	taggagaggt	atactaagta	5820
cagtaaattc	ttttaagaa	ggtgtgtttt	tgcaattgga	aaatttgcaa	aaacaagtga	5880
tagtattggt	taagcaaata	tatggtgagg	at	tttgacgt	atcaattggt	5940
tacgttggca	agtattttat	tatagcttac	taatgtttta	atacggaaaa	cagtctattt	6000
ttgacaaatt	tttaattttt	agaaatcttt	ataaaaaata	ttatttttaac	ttgtttaaag	6060
tatctaacaa	aaattctttg	tctaaaaaatt	tttgtataag	aattgttttcg	aacaaagttt	6120
ttaaaaaaat	attatgggtta	taataggaag	atatcatgga	tactattagt	aaaatttcta	6180
taattgtacc	tatatataat	gtagaaaaat	at	ttatctaa	atgtatagat	6240
atcagaccta	caaacatata	gagattcttc	tggtgaatga	cggtagtacg	gataattcgg	6300
aagaaatttg	tttagcatat	gcgaagaaag	atagtcgcat	tcgttatttt	aaaaaagaga	6360
acggcgggct	atcagatgcc	cgtaattatg	gcataagtcg	cgccaagggt	gactacttag	6420
cttttataga	ctcagatgat	tttattcatt	cggagttcat	ccaacgttta	cacgaagcaa	6480
ttgagagaga	gaatgccctt	gtggcagttg	ctggttatga	tagggtagat	gcttcggggc	6540
at	ttcttaac	agcagagccg	cttcctacaa	atcaggctgt	tctgagcggc	6600
gtaaaaagct	gctagaggcg	gatgggtcatc	gctttgtggt	ggcctgtaat	aaactctata	6660
aaaaagaact	at	tttgaagat	tttcgatttg	aaaagggtaa	gattcatgaa	6720
tcacttatcg	cttgctctat	gagttagaaa	aagttgcaat	agttaaggag	tgcttggtact	6780
attatgttga	ccgagaaaat	agtatcacia	cttctagcat	gactgaccat	cgcttccatt	6840
gcctactgga	at	ttcaaaat	gaacgaatgg	acttctatga	aagtagagga	6900
tcttactaga	gtgttatcgt	tcattttttag	cctttgctgt	tttgttttta	ggcaaatata	6960
atcattgggt	gagcaaacag	caaaagaagc	tt			6992

<210> 31

<211> 455

121

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS1E amino acid sequence"

<400> 31

Arg	Gln	Thr	Lys	Leu	Ala	Leu	Phe	Asp	Met	Ile	Ala	Val	Ala	Ile	Ser	1	5	10	15
Ala	Ile	Leu	Thr	Ser	His	Ile	Pro	Asn	Ala	Asp	Leu	Asn	Arg	Ser	Gly	20	25	30	
Ile	Phe	Ile	Ile	Met	Met	Val	His	Tyr	Phe	Ala	Phe	Phe	Ile	Ser	Arg	35	40	45	
Met	Pro	Val	Glu	Phe	Glu	Tyr	Arg	Gly	Asn	Leu	Ile	Glu	Phe	Glu	Lys	50	55	60	
Thr	Phe	Asn	Tyr	Ser	Ile	Ile	Phe	Ala	Ile	Phe	Leu	Thr	Ala	Val	Ser	65	70	75	80
Phe	Leu	Leu	Glu	Asn	Asn	Phe	Ala	Leu	Ser	Arg	Arg	Gly	Ala	Val	Tyr	85	90	95	
Phe	Thr	Leu	Ile	Asn	Phe	Val	Leu	Val	Tyr	Leu	Phe	Asn	Val	Ile	Ile	100	105	110	
Lys	Gln	Phe	Lys	Asp	Ser	Phe	Leu	Phe	Ser	Thr	Ile	Tyr	Gln	Lys	Lys	115	120	125	
Thr	Ile	Leu	Ile	Thr	Thr	Ala	Glu	Arg	Trp	Glu	Asn	Met	Gln	Val	Leu	130	135	140	
Phe	Glu	Ser	His	Lys	Gln	Ile	Gln	Lys	Asn	Leu	Val	Ala	Leu	Val	Val	145	150	155	160
Leu	Gly	Thr	Glu	Ile	Asp	Lys	Ile	Asn	Leu	Ser	Leu	Pro	Leu	Tyr	Tyr	165	170	175	
Ser	Val	Glu	Glu	Ala	Ile	Glu	Phe	Ser	Thr	Arg	Glu	Val	Val	Asp	His	180	185	190	
Val	Phe	Ile	Asn	Leu	Pro	Ser	Glu	Phe	Leu	Asp	Val	Lys	Gln	Phe	Val	195	200	205	
Ser	Asp	Phe	Glu	Leu	Leu	Gly	Ile	Asp	Val	Ser	Val	Asp	Ile	Asn	Ser	210	215	220	
Phe	Gly	Phe	Thr	Ala	Leu	Lys	Asn	Lys	Lys	Ile	Gln	Leu	Leu	Gly	Asp	225	230	235	240
His	Ser	Ile	Val	Thr	Phe	Ser	Thr	Asn	Phe	Tyr	Lys	Pro	Ser	His	Ile	245	250	255	

122

Met Met Lys Arg Leu Leu Asp Ile Leu Gly Ala Val Val Gly Leu Ile
 260 265 270
 Ile Cys Gly Ile Val Ser Ile Leu Leu Val Pro Ile Ile Arg Arg Asp
 275 280 285
 Gly Gly Pro Ala Ile Phe Ala Gln Lys Arg Val Gly Gln Asn Gly Arg
 290 295 300
 Ile Phe Thr Phe Tyr Lys Phe Arg Ser Met Tyr Val Asp Ala Glu Glu
 305 310 315 320
 Arg Lys Lys Asp Leu Leu Ser Gln Asn Gln Met Gln Gly Trp Val Cys
 325 330 335
 Phe Lys Met Gly Lys Thr Ile Leu Glu Leu Leu Gln Leu Asp Ile Ser
 340 345 350
 Tyr Ala Lys Thr Ser Leu Asp Glu Leu Pro Gln Phe Tyr Asn Val Leu
 355 360 365
 Ile Gly Asp Met Ser Leu Val Gly Thr Arg Pro Pro Thr Val Asp Glu
 370 375 380
 Phe Glu Lys Tyr Thr Pro Gly Gln Lys Arg Arg Leu Ser Phe Lys Pro
 385 390 395 400
 Gly Ile Thr Gly Leu Trp Gln Val Ser Gly Arg Ser Asn Ile Thr Asp
 405 410 415
 Phe Asp Asp Val Val Arg Leu Asp Leu Ala Tyr Ile Asp Asn Trp Thr
 420 425 430
 Ile Trp Ser Asp Ile Lys Ile Leu Leu Lys Thr Val Lys Val Val Leu
 435 440 445
 Leu Arg Glu Gly Ser Lys Asn
 450 455

<210> 32

<211> 149

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS1F amino acid sequence"

<400> 32

Met Lys Val Cys Leu Val Gly Ser Ser Gly Gly His Leu Thr His Leu
 1 5 10 15
 Tyr Leu Leu Lys Pro Phe Trp Lys Glu Glu Glu Arg Phe Trp Val Thr
 20 25 30

```
<210> 33
<211> 164
<212> PRT
<213> Streptococcus suis

<220>
<223> /note="CPS1G amino acid sequence"

<400> 33
```

Met	Ile	Phe	Val	Thr	Val	Gly	Thr	His	Glu	Gln	Gln	Phe	Asn	Arg	Leu
1				5					10					15	
Ile	Lys	Glu	Ile	Asp	Leu	Leu	Lys	Lys	Asn	Gly	Ser	Ile	Thr	Asp	Glu
			20					25					30		
Ile	Phe	Ile	Gln	Thr	Gly	Tyr	Ser	Asp	Tyr	Ile	Pro	Glu	Tyr	Cys	Lys
		35					40					45			
Tyr	Lys	Lys	Phe	Leu	Ser	Tyr	Lys	Glu	Met	Glu	Gln	Tyr	Ile	Asn	Lys
	50					55					60				
Ser	Glu	Val	Val	Ile	Cys	His	Gly	Gly	Pro	Ala	Thr	Phe	Met	Asn	Ser
65					70					75					80
Leu	Ser	Lys	Gly	Lys	Lys	Gln	Leu	Leu	Phe	Pro	Arg	Gln	Lys	Lys	Tyr
				85					90					95	
Gly	Glu	His	Val	Asn	Asp	His	Gln	Val	Glu	Phe	Val	Arg	Arg	Ile	Leu
			100				105						110		

124

Gln Asp Asn Asn Ile Leu Phe Ile Glu Asn Ile Asp Asp Leu Phe Glu
 115 120 125

Lys Ile Ile Glu Val Ser Lys Gln Thr Asn Phe Thr Ser Asn Asn Asn
 130 135 140

Phe Phe Cys Glu Arg Leu Lys Gln Ile Val Glu Lys Phe Asn Glu Asp
 145 150 155 160

Gln Glu Asn Glu

<210> 34

<211> 389

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS1H amino acid sequence"

<400> 34

Met Phe Lys Leu Phe Lys Tyr Asp Pro Glu Tyr Phe Ile Phe Lys Tyr
 1 5 10 15

Phe Trp Leu Ile Ile Phe Ile Pro Glu Gln Lys Tyr Val Phe Leu Leu
 20 25 30

Ile Phe Met Asn Leu Ile Leu Phe His Ile Lys Phe Leu Lys Thr Lys
 35 40 45

Leu Ile Leu Lys Asn Glu Ile Leu Leu Phe Leu Leu Trp Ser Ile Leu
 50 55 60

Cys Phe Val Ser Val Val Thr Ser Met Phe Val Glu Ile Asn Phe Glu
 65 70 75 80

Arg Leu Phe Ala Asp Phe Thr Ala Pro Ile Ile Trp Ile Ile Ala Ile
 85 90 95

Met Tyr Tyr Asn Leu Tyr Ser Phe Ile Asn Ile Asp Tyr Lys Lys Leu
 100 105 110

Lys Asn Ser Ile Phe Phe Ser Phe Leu Val Leu Leu Gly Ile Ser Ala
 115 120 125

Leu Tyr Ile Ile Gln Asn Gly Lys Asp Ile Val Phe Leu Asp Arg His
 130 135 140

Leu Ile Gly Leu Asp Tyr Leu Ile Thr Gly Val Lys Thr Arg Leu Val
 145 150 155 160

Gly Phe Met Asn Tyr Pro Thr Leu Asn Thr Thr Thr Ile Ile Val Ser
 165 170 175

125

Ile Pro Leu Ile Phe Ala Leu Ile Lys Asn Lys Met Gln Gln Phe Phe
 180 185 190
 Phe Leu Cys Leu Ala Phe Ile Pro Ile Tyr Leu Ser Gly Ser Arg Ile
 195 200 205
 Gly Ser Leu Ser Pro Leu Ala Ile Leu Ile Ile Cys Leu Leu Trp Arg
 210 215 220
 Tyr Ile Gly Gly Lys Phe Ala Trp Ile Lys Lys Leu Ile Val Ile Phe
 225 230 235 240
 Val Ile Leu Leu Ile Ile Leu Asn Thr Glu Leu Leu Tyr His Glu Ile
 245 250 255
 Leu Ala Val Tyr Asn Ser Arg Glu Ser Ser Asn Glu Ala Arg Phe Ile
 260 265 270
 Ile Tyr Gln Gly Ser Ile Asp Lys Val Leu Glu Asn Asn Ile Leu Phe
 275 280 285
 Gly Tyr Gly Ile Ser Glu Tyr Ser Val Thr Gly Thr Trp Leu Gly Ser
 290 295 300
 His Ser Gly Tyr Ile Ser Phe Phe Tyr Lys Ser Gly Ile Val Gly Leu
 305 310 315 320
 Ile Leu Leu Met Phe Ser Phe Phe Tyr Val Ile Lys Lys Ser Tyr Gly
 325 330 335
 Val Asn Gly Glu Thr Ala Leu Phe Tyr Phe Thr Ser Leu Ala Ile Phe
 340 345 350
 Phe Ile Tyr Glu Thr Ile Asp Pro Ile Ile Ile Ile Leu Val Leu Phe
 355 360 365
 Phe Ser Ser Ile Gly Ile Trp Asn Asn Ile Asn Phe Lys Lys Asp Met
 370 375 380
 Glu Thr Lys Asn Glu
 385

<210> 35

<211> 322

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS1I amino acid sequence"

<400> 35

Met Asn Asp Leu Ile Ser Val Ile Val Pro Ile Tyr Asn Val Gln Asp
 1 5 10 15

126

Tyr Leu Asp Lys Cys Ile Asn Ser Ile Ile Asn Gln Thr Tyr Thr Asn
 20 25 30
 Leu Glu Val Ile Leu Val Asn Asp Gly Ser Thr Asp Asp Ser Glu Lys
 35 40 45
 Ile Cys Leu Asn Tyr Met Lys Asn Asp Gly Arg Ile Lys Tyr Tyr Lys
 50 55 60
 Lys Ile Asn Gly Gly Leu Ala Asp Ala Arg Asn Phe Gly Leu Glu His
 65 70 75 80
 Ala Thr Gly Lys Tyr Ile Ala Phe Val Asp Ser Asp Asp Tyr Ile Glu
 85 90 95
 Val Ala Met Phe Glu Arg Met His Asp Asn Ile Thr Glu Tyr Asn Ala
 100 105 110
 Asp Ile Ala Glu Ile Asp Phe Cys Leu Val Asp Glu Asn Gly Tyr Thr
 115 120 125
 Lys Lys Lys Arg Asn Ser Asn Phe His Val Leu Thr Arg Glu Glu Thr
 130 135 140
 Val Lys Glu Phe Leu Ser Gly Ser Asn Ile Glu Asn Asn Val Trp Cys
 145 150 155 160
 Lys Leu Tyr Ser Arg Asp Ile Ile Lys Asp Ile Lys Phe Gln Ile Asn
 165 170 175
 Asn Arg Ser Ile Gly Glu Asp Leu Leu Phe Asn Leu Glu Val Leu Asn
 180 185 190
 Asn Val Thr Arg Val Val Val Asp Thr Arg Glu Tyr Tyr Tyr Asn Tyr
 195 200 205
 Val Ile Arg Asn Ser Ser Leu Ile Asn Gln Lys Phe Ser Ile Asn Asn
 210 215 220
 Ile Asp Leu Val Thr Arg Leu Glu Asn Tyr Pro Phe Lys Leu Lys Arg
 225 230 235 240
 Glu Phe Ser His Tyr Phe Asp Ala Lys Val Ile Lys Glu Lys Val Lys
 245 250 255
 Cys Leu Asn Lys Met Tyr Ser Thr Asp Cys Leu Asp Asn Glu Phe Leu
 260 265 270
 Pro Ile Leu Glu Ser Tyr Arg Lys Glu Ile Arg Arg Tyr Pro Phe Ile
 275 280 285
 Lys Ala Lys Arg Tyr Leu Ser Arg Lys His Leu Val Thr Leu Tyr Leu
 290 295 300

127

Met Lys Phe Ser Pro Lys Leu Tyr Val Met Leu Tyr Lys Lys Phe Gln
 305 310 315 320

Lys Gln

<210> 36

<211> 322

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS1J amino acid sequence"

<400> 36

Met Asn Asp Leu Ile Ser Val Ile Val Pro Ile Tyr Asn Val Gln Asp
 1 5 10 15

Tyr Leu Asp Lys Cys Ile Asn Ser Ile Ile Asn Gln Thr Tyr Thr Asn
 20 25 30

Leu Glu Val Ile Leu Val Asn Asp Gly Ser Thr Asp Asp Ser Glu Lys
 35 40 45

Ile Cys Leu Asn Tyr Met Lys Asn Asp Gly Arg Ile Lys Tyr Tyr Lys
 50 55 60

Lys Ile Asn Gly Gly Leu Ala Asp Ala Arg Asn Phe Gly Leu Glu His
 65 70 75 80

Ala Thr Gly Lys Tyr Ile Ala Phe Val Asp Ser Asp Asp Tyr Ile Glu
 85 90 95

Val Ala Met Phe Glu Arg Met His Asp Asn Ile Thr Glu Tyr Asn Ala
 100 105 110

Asp Ile Ala Glu Ile Asp Phe Cys Leu Val Asp Glu Asn Gly Tyr Thr
 115 120 125

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

Val Lys Glu Phe Leu Ser Gly Ser Asn Ile Glu Asn Asn Val Trp Cys
 145 150 155 160

Lys Leu Tyr Ser Arg Asp Ile Ile Lys Asp Ile Lys Phe Gln Ile Asn
 165 170 175

Asn Arg Ser Ile Gly Glu Asp Leu Leu Phe Asn Leu Glu Val Leu Asn
 180 185 190

Asn Val Thr Arg Val Val Val Asp Thr Arg Glu Tyr Tyr Tyr Asn Tyr
 195 200 205

128

Val Ile Arg Asn Ser Ser Leu Ile Asn Gln Lys Phe Ser Ile Asn Asn
 210 215 220
 Ile Asp Leu Val Thr Arg Leu Glu Asn Tyr Pro Phe Lys Leu Lys Arg
 225 230 235 240
 Glu Phe Ser His Tyr Phe Asp Ala Lys Val Ile Lys Glu Lys Val Lys
 245 250 255
 Cys Leu Asn Lys Met Tyr Ser Thr Asp Cys Leu Asp Asn Glu Phe Leu
 260 265 270
 Pro Ile Leu Glu Ser Tyr Arg Lys Glu Ile Arg Arg Tyr Pro Phe Ile
 275 280 285
 Lys Ala Lys Arg Tyr Leu Ser Arg Lys His Leu Val Thr Leu Tyr Leu
 290 295 300
 Met Lys Phe Ser Pro Lys Leu Tyr Val Met Leu Tyr Lys Lys Phe Gln
 305 310 315 320
 Lys Gln

<210> 37
 <211> 278
 <212> PRT
 <213> Streptococcus suis

<220>
 <223> /note="CPS1K amino acid sequence"

<400> 37

Met Asp Thr Ile Ser Lys Ile Ser Ile Ile Val Pro Ile Tyr Asn Val
 1 5 10 15
 Glu Lys Tyr Leu Ser Lys Cys Ile Asp Ser Ile Val Asn Gln Thr Tyr
 20 25 30
 Lys His Ile Glu Ile Leu Leu Val Asn Asp Gly Ser Thr Asp Asn Ser
 35 40 45
 Glu Glu Ile Cys Leu Ala Tyr Ala Lys Lys Asp Ser Arg Ile Arg Tyr
 50 55 60
 Phe Lys Lys Glu Asn Gly Gly Leu Ser Asp Ala Arg Asn Tyr Gly Ile
 65 70 75 80
 Ser Arg Ala Lys Gly Asp Tyr Leu Ala Phe Ile Asp Ser Asp Asp Phe
 85 90 95
 Ile His Ser Glu Phe Ile Gln Arg Leu His Glu Ala Ile Glu Arg Glu
 100 105 110

129

Asn Ala Leu Val Ala Val Ala Gly Tyr Asp Arg Val Asp Ala Ser Gly
 115 120 125
 His Phe Leu Thr Ala Glu Pro Leu Pro Thr Asn Gln Ala Val Leu Ser
 130 135 140
 Gly Arg Asn Val Cys Lys Lys Leu Leu Glu Ala Asp Gly His Arg Phe
 145 150 155 160
 Val Val Ala Cys Asn Lys Leu Tyr Lys Lys Glu Leu Phe Glu Asp Phe
 165 170 175
 Arg Phe Glu Lys Gly Lys Ile His Glu Asp Glu Tyr Phe Thr Tyr Arg
 180 185 190
 Leu Leu Tyr Glu Leu Glu Lys Val Ala Ile Val Lys Glu Cys Leu Tyr
 195 200 205
 Tyr Tyr Val Asp Arg Glu Asn Ser Ile Thr Thr Ser Ser Met Thr Asp
 210 215 220
 His Arg Phe His Cys Leu Leu Glu Phe Gln Asn Glu Arg Met Asp Phe
 225 230 235 240
 Tyr Glu Ser Arg Gly Asp Lys Glu Leu Leu Leu Glu Cys Tyr Arg Ser
 245 250 255
 Phe Leu Ala Phe Ala Val Leu Phe Leu Gly Lys Tyr Asn His Trp Leu
 260 265 270
 Ser Lys Gln Gln Lys Lys
 275

<210> 38

<211> 4519

<212> DNA

<213> Streptococcus suis

<220>

<223> /note="CPS9 nucleotide sequence"

<400> 38

```

aagcttatcg tcaaggtggt cgctatatcg tggcgacatc tcatagacga aaagggatgt      60
ttgaaacacc agaaaaagtt atcatgacta actttcttca atttaaagac gcagtagcag      120
aagtttatcc tgaaatacga ttgtgctatg gtgctgaatt gtattatagt aaagatatat      180
taagcaaact tgaaaaaaag aaagtaccca cacttaatgg ctgcgcgtat attcttttgg      240
agttcagtag tgatactcct tggaaagaga ttcaagaagc agtgaacgaa gtgacgctac      300
ttgggctaac tcccgtaact gcccatatag aacgatatga cgccctagcg tttcatgcag      360
agagagtaga agagttaatt gacaagggat gctatactca ggtaaatagt aatcatgtgc      420
tgaagcccac tttaattggt gatcgagcaa aagaatttaa aaaacgtact cggtattttt      480
tagagcagga tttagtacat tgtgttgcta gcgatatgca taatttatct agtagacctc      540
cgtttatgag ggaggcttat aagttgctaa cagaggaatt tggcaaagat aaagcgaaag      600
cgttgctaaa aaagaatcct cttatgctat taaaaacca ggcgatttaa actgggttact      660

```

ctagattgtg	gagagaaaaa	tggatttagg	aactgttact	gataaactgt	tagaacgcaa	720
cagtaaacga	ttgatactcg	tgtgcatgga	tacgtgtctt	cttatagttt	ccatgatttt	780
gagcagactg	tttttggatg	ttattattga	cataccagat	gaacgcttca	ttcttgcaagt	840
tttattcgta	tcaattttat	atttgattct	atcgtttaga	ttaaaagtct	tttcattaat	900
tacgcgttac	acaggggtatc	agagttatgt	aaaaatagga	cttagtttaa	tatctgcgca	960
ttcattgttt	ttaattatct	caatgggtgt	gtggcaggct	tttagttatc	gtttcatctt	1020
agtatcctta	tttttgtcgt	atgtaatgct	cattactccg	aggattgttt	ggaaagtctt	1080
acatgagacg	agaaaaaatg	ctatccgtaa	gaaggatagc	ccactaagaa	tcttagtagt	1140
aggtgctgga	gatgggtgga	atatttttat	caatactgtc	aaagatcgaa	aattgaattt	1200
tgaaattgtc	ggtatcgttg	atcgtgatcc	aaataaactt	ggaacattta	tccgtacggc	1260
taaagtttta	ggaaaccgta	atgatattcc	acgactggta	gaggaattag	ctgttgacca	1320
agtgcagatt	gccatccctt	ctttaaatgg	taaggagcga	gagaagattg	ttgaaatctg	1380
taacactaca	ggagtgaccg	tcaataatat	gccgagtatt	gaagacatta	tggcggggaa	1440
catgtctgtc	agtgcctttc	aggaaattga	cgtagcagac	cttcttggtc	gaccagaggt	1500
tgttttggat	caggatgaat	tgaatcagtt	tttccaaggg	aaaacaatcc	ttgtcacagg	1560
agcagggtggc	tctatcgggt	cagagctatg	tcgtcaaatt	gctaagttta	cgcctaaacg	1620
cttggtgttg	cttggacatg	gagaaaattc	aatctatctc	attcatcgag	agttactgga	1680
aaagtaccaa	ggtaagattg	agttgggtccc	tctcattgca	gatattcaag	atagagaatt	1740
gatttttagc	ataatggctg	aatatcaacc	cgatgttggt	tatcatgctg	cagcacataa	1800
gcatgttcct	ttgatggaat	ataatccaca	tgaagcagtg	aagaataata	tttttggaac	1860
gaagaatgtg	gctgaggcgg	ctaaaactgc	aaagggtgcc	aaatttggtta	tggtttcaac	1920
agataaagct	gttaatccac	caaatgtcat	gggagcgact	aaacgtggtg	cagaaatgat	1980
tgttacaggt	ttaaacgagc	caggtcagac	tcaatttgcg	gcagtccggt	ttgggaatgt	2040
tctaggtagt	cgtggaagtg	ttgttccgct	attcaaagag	caaattagaa	aaggtggacc	2100
tgttacgggt	accgacttta	ggatgactcg	ttatttcatg	acgattcctg	aggcaagtcg	2160
tttggttatc	caagctggac	atltggcaaa	aggtggagaa	atatttgtct	tggatatggg	2220
cgagccagta	caaatcctgg	aattggcaag	aaaagttatc	ttgttaagtg	gacacacaga	2280
ggaagaaatc	gggattgtag	aatctggaat	cagaccaggc	gagaaactct	acgaggaatt	2340
attatcaaca	gaagaacgtg	tcagcgaaca	gattcatgaa	aaaatatttg	tgggtcgcgt	2400
tacaaataag	cagtcggaca	ttgtcaattc	atlttatcaat	ggattactcc	aaaaagatag	2460
aatgaatta	aaaaatatgt	tgattgaatt	tgcaaaacaa	gaataagaaa	gtaaaaaata	2520
tttttacttt	cctagagttt	aaacgatgtt	taagttctag	gaaggttaga	atacctaatt	2580
aacaacaata	ttactattta	ttaagagtca	gataatagca	actaagtgct	acaaactatc	2640
tttataataa	gtatatattg	tcaaaaggga	gatgtgaaat	gtatccaatt	tgtaaacgta	2700
tttttagcaat	tattatctca	gggattgcta	ttgttggtct	gagtccaatt	ttattattga	2760
ttgcattggc	aattaaatta	gattctaaag	gtccgggtatt	atlttaacaa	aagcgggttg	2820
gtaaaaacaa	gtcatacttt	atgatlttata	aattccggtc	tatgtacgtt	gacgcaccaa	2880
gtgatatgcc	gactcatcta	ttaaaggatc	ctaaggcgat	gattaccaag	gtgggcgcgt	2940
ttctcagaaa	aacaagttta	gatgaactgc	cacagctttt	taatattttt	aaaggtgaaa	3000
tggcgattgt	tgggtccacgc	ccagccttat	ggaatcaata	tgacttaatt	gaagagcgag	3060
ataaatatgg	tgcaaatgat	attcgtcctg	gactaaccgg	ttgggctcaa	attaatgggtc	3120
gtgatgaatt	ggaaattgat	gaaaagtcaa	aattagatgg	atattatggt	caaaatatga	3180
gtctaggttt	ggatatttaa	tglttcttag	gtacattcct	cagtgtagcc	agaagcgaag	3240
gtgttggtga	aggtggaaca	gggcagaaag	gaaaaggatg	aaattttcag	tattaatgtc	3300
ggtctatgag	aaagaaaaac	cagagtttct	tagggaatct	ttggaaagca	tccttgtcaa	3360
tcaaacaatg	attccaacgg	aggttggtct	ggtagaggat	gggccactca	atcagagctt	3420
atatagtatt	ttagaagaat	ttaaaagtgc	atlttctatt	tttaaaacga	tagccttgga	3480
aaagaattcg	ggtttaggaa	ttgcactgaa	tgaaggtttg	aaacattgta	attatgagtg	3540
ggtttgcacg	aaatggattc	tgatgatgtt	gcatatacat	acacgttttg	aaaagcaagt	3600
taactttata	aaacaaaacc	cgactataga	tattgagata	gatgagttct	taaattctac	3660
tagtgaaata	gtttctcata	aaaatgttcc	aaccagcac	gatgaaatat	taaagatggc	3720
aaggcgggag	aaatccatgt	gccacatgac	tgtaatgttt	aaaaagaaaa	gtgtcgagag	3780
agcagggggg	tatcaaacac	ttccgtacgt	agaagattat	ttcctttggg	tgcgcagtgt	3840
tgcttcagga	tcgaaatttg	caaacattga	tgaaacacta	gttcttgcac	gtgttggaag	3900
tgggatgttc	aataggaggg	ggaacagaga	acaaattaac	agttggacat	tactaattga	3960

131

```

atttatgtta gctcaaggaa ttgttacacc actagatgta tttattaatc aaattttacat 4020
tagggctcttt gtttatatgc caacttggat aaagaaactc atttatggaa aaatcttaag 4080
gaaatagtat gattacagta ttgatggcta catataatgg aagcccatTT ataataaaac 4140
agttagattc aattcgaaat caaagtgtat cagcagacaa agttattatt tgggatgatt 4200
gctcgacaga tgatacaata aaaataataa aagattatat aaaaaaatat tctttggatt 4260
catgggttgt ctctcaaaat aaatctaatac aggggcatta tcaaacattt ataaatttga 4320
caaagttagt tcaggaagga atagtctttt ttccagatca agatgatatt tgggactgtc 4380
ataaaattga gacaatgctt ccaatctttg acagagaaaa tgtatcaatg gtgttttgca 4440
aatccagatt gattgatgaa aacggaaata ttatcagtag cccagatact tcggatagaa 4500
tcaatacgta ctctctaga 4519

```

<210> 39

<211> 215

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS9D amino acid sequence"

<400> 39

```

Ala Tyr Arg Gln Gly Val Arg Tyr Ile Val Ala Thr Ser His Arg Arg
 1              5              10              15

Lys Gly Met Phe Glu Thr Pro Glu Lys Val Ile Met Thr Asn Phe Leu
      20              25              30

Gln Phe Lys Asp Ala Val Ala Glu Val Tyr Pro Glu Ile Arg Leu Cys
      35              40              45

Tyr Gly Ala Glu Leu Tyr Tyr Ser Lys Asp Ile Leu Ser Lys Leu Glu
      50              55              60

Lys Lys Lys Val Pro Thr Leu Asn Gly Ser Arg Tyr Ile Leu Leu Glu
      65              70              75              80

Phe Ser Ser Asp Thr Pro Trp Lys Glu Ile Gln Glu Ala Val Asn Glu
      85              90              95

Val Thr Leu Leu Gly Leu Thr Pro Val Leu Ala His Ile Glu Arg Tyr
      100              105              110

Asp Ala Leu Ala Phe His Ala Glu Arg Val Glu Glu Leu Ile Asp Lys
      115              120              125

Gly Cys Tyr Thr Gln Val Asn Ser Asn His Val Leu Lys Pro Thr Leu
      130              135              140

Ile Gly Asp Arg Ala Lys Glu Phe Lys Lys Arg Thr Arg Tyr Phe Leu
      145              150              155              160

Glu Gln Asp Leu Val His Cys Val Ala Ser Asp Met His Asn Leu Ser
      165              170              175

```

132

Ser Arg Pro Pro Phe Met Arg Glu Ala Tyr Lys Leu Leu Thr Glu Glu
 180 185 190

Phe Gly Lys Asp Lys Ala Lys Ala Leu Leu Lys Lys Asn Pro Leu Met
 195 200 205

Leu Leu Lys Asn Gln Ala Ile
 210 215

<210> 40

<211> 608

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS9E amino acid sequence"

<400> 40

Met Asp Leu Gly Thr Val Thr Asp Lys Leu Leu Glu Arg Asn Ser Lys
 1 5 10 15

Arg Leu Ile Leu Val Cys Met Asp Thr Cys Leu Leu Ile Val Ser Met
 20 25 30

Ile Leu Ser Arg Leu Phe Leu Asp Val Ile Ile Asp Ile Pro Asp Glu
 35 40 45

Arg Phe Ile Leu Ala Val Leu Phe Val Ser Ile Leu Tyr Leu Ile Leu
 50 55 60

Ser Phe Arg Leu Lys Val Phe Ser Leu Ile Thr Arg Tyr Thr Gly Tyr
 65 70 75 80

Gln Ser Tyr Val Lys Ile Gly Leu Ser Leu Ile Ser Ala His Ser Leu
 85 90 95

Phe Leu Ile Ile Ser Met Val Leu Trp Gln Ala Phe Ser Tyr Arg Phe
 100 105 110

Ile Leu Val Ser Leu Phe Leu Ser Tyr Val Met Leu Ile Thr Pro Arg
 115 120 125

Ile Val Trp Lys Val Leu His Glu Thr Arg Lys Asn Ala Ile Arg Lys
 130 135 140

Lys Asp Ser Pro Leu Arg Ile Leu Val Val Gly Ala Gly Asp Gly Gly
 145 150 155 160

Asn Ile Phe Ile Asn Thr Val Lys Asp Arg Lys Leu Asn Phe Glu Ile
 165 170 175

Val Gly Ile Val Asp Arg Asp Pro Asn Lys Leu Gly Thr Phe Ile Arg
 180 185 190

133

Thr Ala Lys Val Leu Gly Asn Arg Asn Asp Ile Pro Arg Leu Val Glu
 195 200 205
 Glu Leu Ala Val Asp Gln Val Thr Ile Ala Ile Pro Ser Leu Asn Gly
 210 215 220
 Lys Glu Arg Glu Lys Ile Val Glu Ile Cys Asn Thr Thr Gly Val Thr
 225 230 235 240
 Val Asn Asn Met Pro Ser Ile Glu Asp Ile Met Ala Gly Asn Met Ser
 245 250 255
 Val Ser Ala Phe Gln Glu Ile Asp Val Ala Asp Leu Leu Gly Arg Pro
 260 265 270
 Glu Val Val Leu Asp Gln Asp Glu Leu Asn Gln Phe Phe Gln Gly Lys
 275 280 285
 Thr Ile Leu Val Thr Gly Ala Gly Gly Ser Ile Gly Ser Glu Leu Cys
 290 295 300
 Arg Gln Ile Ala Lys Phe Thr Pro Lys Arg Leu Leu Leu Leu Gly His
 305 310 315 320
 Gly Glu Asn Ser Ile Tyr Leu Ile His Arg Glu Leu Leu Glu Lys Tyr
 325 330 335
 Gln Gly Lys Ile Glu Leu Val Pro Leu Ile Ala Asp Ile Gln Asp Arg
 340 345 350
 Glu Leu Ile Phe Ser Ile Met Ala Glu Tyr Gln Pro Asp Val Val Tyr
 355 360 365
 His Ala Ala Ala His Lys His Val Pro Leu Met Glu Tyr Asn Pro His
 370 375 380
 Glu Ala Val Lys Asn Asn Ile Phe Gly Thr Lys Asn Val Ala Glu Ala
 385 390 395 400
 Ala Lys Thr Ala Lys Val Ala Lys Phe Val Met Val Ser Thr Asp Lys
 405 410 415
 Ala Val Asn Pro Pro Asn Val Met Gly Ala Thr Lys Arg Val Ala Glu
 420 425 430
 Met Ile Val Thr Gly Leu Asn Glu Pro Gly Gln Thr Gln Phe Ala Ala
 435 440 445
 Val Arg Phe Gly Asn Val Leu Gly Ser Arg Gly Ser Val Val Pro Leu
 450 455 460
 Phe Lys Glu Gln Ile Arg Lys Gly Gly Pro Val Thr Val Thr Asp Phe
 465 470 475 480

Leu Pro Gln Leu Phe Asn Ile Phe Lys Gly Glu Met Ala Ile Val Gly
100 105 110

135

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

Lys Tyr Gly Ala Asn Asp Ile Arg Pro Gly Leu Thr Gly Trp Ala Gln
 130 135 140

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

Gly Tyr Tyr Val Gln Asn Met Ser Leu Gly Leu Asp Ile Lys Cys Phe
 165 170 175

Leu Gly Thr Phe Leu Ser Val Ala Arg Ser Glu Gly Val Val Glu Gly
 180 185 190

Gly Thr Gly Gln Lys Gly Lys Gly
 195 200

<210> 42

<211> 269

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS9G amino acid sequence"

<400> 42

Met Lys Phe Ser Val Leu Met Ser Val Tyr Glu Lys Glu Lys Pro Glu
 1 5 10 15

Phe Leu Arg Glu Ser Leu Glu Ser Ile Leu Val Asn Gln Thr Met Ile
 20 25 30

Pro Thr Glu Val Val Leu Val Glu Asp Gly Pro Leu Asn Gln Ser Leu
 35 40 45

Tyr Ser Ile Leu Glu Glu Phe Lys Ser Arg Phe Ser Phe Phe Lys Thr
 50 55 60

Ile Ala Leu Glu Lys Asn Ser Gly Leu Gly Ile Ala Leu Asn Glu Gly
 65 70 75 80

Leu Lys His Cys Asn Tyr Glu Trp Val Cys Thr Lys Trp Ile Leu Met
 85 90 95

Met Leu His Ile His Thr Arg Phe Glu Lys Gln Val Asn Phe Ile Lys
 100 105 110

Gln Asn Pro Thr Ile Asp Ile Glu Ile Asp Glu Phe Leu Asn Ser Thr
 115 120 125

Ser Glu Ile Val Ser His Lys Asn Val Pro Thr Gln His Asp Glu Ile
 130 135 140

136

Leu Lys Met Ala Arg Arg Glu Lys Ser Met Cys His Met Thr Val Met
 145 150 155 160
 Phe Lys Lys Lys Ser Val Glu Arg Ala Gly Gly Tyr Gln Thr Leu Pro
 165 170 175
 Tyr Val Glu Asp Tyr Phe Leu Trp Val Arg Met Ile Ala Ser Gly Ser
 180 185 190
 Lys Phe Ala Asn Ile Asp Glu Thr Leu Val Leu Ala Arg Val Gly Asn
 195 200 205
 Gly Met Phe Asn Arg Arg Gly Asn Arg Glu Gln Ile Asn Ser Trp Thr
 210 215 220
 Leu Leu Ile Glu Phe Met Leu Ala Gln Gly Ile Val Thr Pro Leu Asp
 225 230 235 240
 Val Phe Ile Asn Gln Ile Tyr Ile Arg Val Phe Val Tyr Met Pro Thr
 245 250 255
 Trp Ile Lys Lys Leu Ile Tyr Gly Lys Ile Leu Arg Lys
 260 265

<210> 43

<211> 143

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS9H amino acid sequence"

<400> 43

Met Ile Thr Val Leu Met Ala Thr Tyr Asn Gly Ser Pro Phe Ile Ile
 1 5 10 15
 Lys Gln Leu Asp Ser Ile Arg Asn Gln Ser Val Ser Ala Asp Lys Val
 20 25 30
 Ile Ile Trp Asp Asp Cys Ser Thr Asp Asp Thr Ile Lys Ile Ile Lys
 35 40 45
 Asp Tyr Ile Lys Lys Tyr Ser Leu Asp Ser Trp Val Val Ser Gln Asn
 50 55 60
 Lys Ser Asn Gln Gly His Tyr Gln Thr Phe Ile Asn Leu Thr Lys Leu
 65 70 75 80
 Val Gln Glu Gly Ile Val Phe Phe Ser Asp Gln Asp Asp Ile Trp Asp
 85 90 95
 Cys His Lys Ile Glu Thr Met Leu Pro Ile Phe Asp Arg Glu Asn Val
 100 105 110

137

Ser Met Val Phe Cys Lys Ser Arg Leu Ile Asp Glu Asn Gly Asn Ile
 115 120 125

Ile Ser Ser Pro Asp Thr Ser Asp Arg Ile Asn Thr Tyr Ser Leu
 130 135 140

<210> 44

<211> 3732

<212> DNA

<213> Streptococcus suis

<220>

<223> /note="CPS7 nucleotide sequence"

<400> 44

ctgcagcaca	taagcatggt	ccattgatgg	aatataatcc	acatgaagca	gtgaagaata	60
atatttttgg	aacgaagaat	gtggctgagg	cggctaaaac	tgcaaagggt	gccaaatttg	120
ttatggtttc	aacagataaa	gctgttaatc	cgccaaatgt	catgggagcg	actaaacgtg	180
ttgcagaaat	gattgtaaca	ggtttaaacg	agccagggtca	gactcaattt	gcggcagtc	240
gttttgggaa	tgttctaggt	agtcgtggaa	gtgttggtcc	gctattcaaa	gagcaaatta	300
gaaaagggtg	acctgttacg	gttaccgact	ttaggatgac	tcgttatctt	atgacgattc	360
ctgaggcaag	tcgtttgggt	atccaagctg	gacatttggt	aaaagggtga	gaaatctttg	420
tcttgatat	gggtgagcca	gtacaaatcc	tgggaattgg	aagaaaagtt	atcttggtta	480
gcggacatac	agaggaagaa	atcgggattg	tagaatctgg	aatcagacca	ggcgagaaac	540
tctacgagga	attgttatca	acagaagaac	gtgtcagcga	acagattcat	gaaaaaatat	600
ttgtgggtcg	cgttacaaat	aagcagtcgg	acattgtcaa	ttcattttat	aatggattac	660
tccaaaaaga	tagaaatgaa	ttaaaagata	tgttgattga	atttgcaaaa	caagaataag	720
aaagtaaaaa	atattttttac	tttccctagag	tttaaaccgat	gtttaagttc	taggaagggt	780
ggaattgctt	tcgtggaggt	gatagataga	aacctatata	tttgtagaag	aaaggatatt	840
aaactaaagg	tgaatcggaa	cataaagttt	agatagagtt	ggtatttaat	gccaaacagg	900
tgaatgcaac	ctctcgctcg	ttactaagca	ggagatagta	aagttgcttg	aaagagagtt	960
tgttaatcag	tataagtagg	ctaaagttag	aatatatatc	tattattatc	ggtaatgata	1020
ctattattga	gaattattgt	agtggggata	aaaataatct	ttggtgattt	tatcgtccga	1080
cttaaagggt	ggttaaaaaa	gtacttatat	tcttttagaa	ttgatgaaaa	atatggggga	1140
atataatatt	tataggagat	acgatgacta	gagtagagtt	gattactaga	gaatttttta	1200
agaagaatga	agcaaccagt	aaatatcttc	agaagataga	atcaagaaga	ggtgaattat	1260
ttattaaatt	ctttatggat	aagttacttg	cgcttatect	attattgcta	ttatccccag	1320
taatcattat	attagctatt	tggataaaa	tagatagtaa	ggggccaatt	ttttatcgcc	1380
aagaacgtgt	tacgagatat	ggtcgaattt	ttagaatatt	taagttttaga	acaatgattt	1440
ctgatgcgga	taaagtcgga	agtcctgtca	cagtcgggtca	agataatcgt	attacgaaag	1500
tcggtcacat	tatcagaaaa	tatcggctgg	acgaagtgcc	ccaacttttt	aatgttttta	1560
tgggggatat	gagctttgta	ggtgtaagac	cagaagtaca	aaaatatgta	aatcagtata	1620
ctgatgaaat	gtttgcgacg	ttactttttac	ctgcaggaat	tacttcacca	gcgagtattg	1680
catataagga	tgaagatatt	gttttagaag	aatattgttc	tcaaggctat	agtcctgatg	1740
aagcatatgt	tcaaaaagta	ttaccagaaa	aaatgaagta	caatttggaa	tatatcagaa	1800
actttggaat	tatttctgat	tttaaagtaa	tgattgatac	agtaattaaa	gtaataaaat	1860
aggagattaa	aatgacaaaa	agacaaaata	ttccattttc	accaccagat	attacccaag	1920
ctgaaattga	tgaagttatt	gacacactaa	aatctggttg	gattacaaca	ggaccaaaga	1980
caaaagagct	agaacgtcgg	ctatcagtat	ttacaggaac	caataaaaact	gtgtgtttta	2040
attctgctac	tgcaggattg	gaactagtct	tacgaattct	tgggtgttga	cccggagatg	2100
aagttattgt	tcctgctatg	acctatactg	cctcatgtag	tgtcattact	catgtaggag	2160
caactcctgt	gatggttgat	attcaaaaaa	acagctttga	gatggaatat	gatgcttttg	2220
aaaaagcgat	tactccgaaa	acaaaagtta	tcattcctgt	tgatctagct	ggtattcctt	2280

138

```

gtgattatga taagatttat accatcgtag aaaacaaacg ctctttgtat gttgcttctg 2340
ataataaatg gcagaaactt tttgggcgag ttattatcct atctgatagt gcacactcac 2400
taggtgctag ttataaggga aaaccagcgg gttccctagc agattttacc tcattttctt 2460
tccatgcagt taagaatttt acaactgctg aaggaggtag tgtgacatgg agatcacatc 2520
ctgatttgga tgacgaagag atgtataaag agtttcagat ttactctctt catggtcaga 2580
caaaggatgc attagctaag acacaattag ggtcatggga atatgacatt gttattcctg 2640
gttacaagtg taatatgaca gatattatgg caggtatcgg tcttggtgcaa ttagaacggt 2700
acccatcttt gttgaatcgt cgcagagaaa tcattgagaa atacaatgct ggctttgagg 2760
ggacttcgat taagccgttg gtacacctga cggagataa acaatcgtct atgcacttgt 2820
atatcacgca tctacaaggc tatacttttag aacaacgaaa tgaagtcatt caaaaaatgg 2880
ctgaagcagg tattgcgtgc aatgttcact acaaaccatt acctcttctc acagcctaca 2940
agaatcttgg ttttgaaatg aaagattttc cgaatgccta tcagtatttt gaaaatgaag 3000
ttacactgcc tcttcatacc aacttgagtg atgaagatgt ggagtatgtg atagaaatgt 3060
ttttaaaaaat tgtagtaga gattagttat tttggaagga gatatgggtg aaagagatat 3120
ggtggaagaa gacacgttgg tatctataat aatgccctcg tggaatacag ctaagtatat 3180
atctgaatca atccagtcag tgttggacca aacacaccaa aattgggaac ttataatcgt 3240
tgatgattgt tctaattgacg aaactgaaaa agttgtttcg catttcaaag attcaagaat 3300
aaagtttttt aaaaattcga ataatttagg ggcagctcta acacgaaata aggcactaag 3360
aaaagctaga ggtaggtgga ttgcgttctt ggattcagat gatttatggc acccgagtaa 3420
gctagaaaaa cagcttgaat ttatgaaaaa taatggatat tcatttactt atcacaattt 3480
tgaaaagatt gatgaatcta gtcagtcttt acgtgtcctg gtgtcaggac cagcaattgt 3540
gactagaaaa atgatgtaca attacggcta tccagggtgt ttgactttca tgtatgatgc 3600
agacaaaatg ggtttaattc agataaaaga tataaagaaa aataacgatt atgcgatatt 3660
acttcaattg tgtaagaagt atgactgtta tcttttaaat gaaagttag cttcgtatcg 3720
aattagaaaa aa 3732

```

<210> 45

<211> 238

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS7E amino acid sequence"

<400> 45

```

Ala Ala His Lys His Val Pro Leu Met Glu Tyr Asn Pro His Glu Ala
 1              5              10              15

Val Lys Asn Asn Ile Phe Gly Thr Lys Asn Val Ala Glu Ala Ala Lys
      20              25              30

Thr Ala Lys Val Ala Lys Phe Val Met Val Ser Thr Asp Lys Ala Val
      35              40              45

Asn Pro Pro Asn Val Met Gly Ala Thr Lys Arg Val Ala Glu Met Ile
      50              55              60

Val Thr Gly Leu Asn Glu Pro Gly Gln Thr Gln Phe Ala Ala Val Arg
      65              70              75              80

Phe Gly Asn Val Leu Gly Ser Arg Gly Ser Val Val Pro Leu Phe Lys
      85              90              95

```

139

Glu Gln Ile Arg Lys Gly Gly Pro Val Thr Val Thr Asp Phe Arg Met
 100 105 110
 Thr Arg Tyr Phe Met Thr Ile Pro Glu Ala Ser Arg Leu Val Ile Gln
 115 120 125
 Ala Gly His Leu Ala Lys Gly Gly Glu Ile Phe Val Leu Asp Met Gly
 130 135 140
 Glu Pro Val Gln Ile Leu Glu Leu Ala Arg Lys Val Ile Leu Leu Ser
 145 150 155 160
 Gly His Thr Glu Glu Glu Ile Gly Ile Val Glu Ser Gly Ile Arg Pro
 165 170 175
 Gly Glu Lys Leu Tyr Glu Glu Leu Leu Ser Thr Glu Glu Arg Val Ser
 180 185 190
 Glu Gln Ile His Glu Lys Ile Phe Val Gly Arg Val Thr Asn Lys Gln
 195 200 205
 Ser Asp Ile Val Asn Ser Phe Ile Asn Gly Leu Leu Gln Lys Asp Arg
 210 215 220
 Asn Glu Leu Lys Asp Met Leu Ile Glu Phe Ala Lys Gln Glu
 225 230 235

<210> 46

<211> 232

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS7F amino acid sequence"

<400> 46

Met Thr Arg Val Glu Leu Ile Thr Arg Glu Phe Phe Lys Lys Asn Glu
 1 5 10 15
 Ala Thr Ser Lys Tyr Phe Gln Lys Ile Glu Ser Arg Arg Gly Glu Leu
 20 25 30
 Phe Ile Lys Phe Phe Met Asp Lys Leu Leu Ala Leu Ile Leu Leu Leu
 35 40 45
 Leu Leu Ser Pro Val Ile Ile Ile Leu Ala Ile Trp Ile Lys Leu Asp
 50 55 60
 Ser Lys Gly Pro Ile Phe Tyr Arg Gln Glu Arg Val Thr Arg Tyr Gly
 65 70 75 80
 Arg Ile Phe Arg Ile Phe Lys Phe Arg Thr Met Ile Ser Asp Ala Asp
 85 90 95

140

Lys Val Gly Ser Leu Val Thr Val Gly Gln Asp Asn Arg Ile Thr Lys
 100 105 110
 Val Gly His Ile Ile Arg Lys Tyr Arg Leu Asp Glu Val Pro Gln Leu
 115 120 125
 Phe Asn Val Leu Met Gly Asp Met Ser Phe Val Gly Val Arg Pro Glu
 130 135 140
 Val Gln Lys Tyr Val Asn Gln Tyr Thr Asp Glu Met Phe Ala Thr Leu
 145 150 155 160
 Leu Leu Pro Ala Gly Ile Thr Ser Pro Ala Ser Ile Ala Tyr Lys Asp
 165 170 175
 Glu Asp Ile Val Leu Glu Glu Tyr Cys Ser Gln Gly Tyr Ser Pro Asp
 180 185 190
 Glu Ala Tyr Val Gln Lys Val Leu Pro Glu Lys Met Lys Tyr Asn Leu
 195 200 205
 Glu Tyr Ile Arg Asn Phe Gly Ile Ile Ser Asp Phe Lys Val Met Ile
 210 215 220
 Asp Thr Val Ile Lys Val Ile Lys
 225 230

<210> 47

<211> 404

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS7G amino acid sequence"

<400> 47

Met Thr Lys Arg Gln Asn Ile Pro Phe Ser Pro Pro Asp Ile Thr Gln
 1 5 10 15
 Ala Glu Ile Asp Glu Val Ile Asp Thr Leu Lys Ser Gly Trp Ile Thr
 20 25 30
 Thr Gly Pro Lys Thr Lys Glu Leu Glu Arg Arg Leu Ser Val Phe Thr
 35 40 45
 Gly Thr Asn Lys Thr Val Cys Leu Asn Ser Ala Thr Ala Gly Leu Glu
 50 55 60
 Leu Val Leu Arg Ile Leu Gly Val Gly Pro Gly Asp Glu Val Ile Val
 65 70 75 80
 Pro Ala Met Thr Tyr Thr Ala Ser Cys Ser Val Ile Thr His Val Gly
 85 90 95

141

Ala Thr Pro Val Met Val Asp Ile Gln Lys Asn Ser Phe Glu Met Glu
 100 105 110
 Tyr Asp Ala Leu Glu Lys Ala Ile Thr Pro Lys Thr Lys Val Ile Ile
 115 120 125
 Pro Val Asp Leu Ala Gly Ile Pro Cys Asp Tyr Asp Lys Ile Tyr Thr
 130 135 140
 Ile Val Glu Asn Lys Arg Ser Leu Tyr Val Ala Ser Asp Asn Lys Trp
 145 150 155 160
 Gln Lys Leu Phe Gly Arg Val Ile Ile Leu Ser Asp Ser Ala His Ser
 165 170 175
 Leu Gly Ala Ser Tyr Lys Gly Lys Pro Ala Gly Ser Leu Ala Asp Phe
 180 185 190
 Thr Ser Phe Ser Phe His Ala Val Lys Asn Phe Thr Thr Ala Glu Gly
 195 200 205
 Gly Ser Val Thr Trp Arg Ser His Pro Asp Leu Asp Asp Glu Glu Met
 210 215 220
 Tyr Lys Glu Phe Gln Ile Tyr Ser Leu His Gly Gln Thr Lys Asp Ala
 225 230 235 240
 Leu Ala Lys Thr Gln Leu Gly Ser Trp Glu Tyr Asp Ile Val Ile Pro
 245 250 255
 Gly Tyr Lys Cys Asn Met Thr Asp Ile Met Ala Gly Ile Gly Leu Val
 260 265 270
 Gln Leu Glu Arg Tyr Pro Ser Leu Leu Asn Arg Arg Arg Glu Ile Ile
 275 280 285
 Glu Lys Tyr Asn Ala Gly Phe Glu Gly Thr Ser Ile Lys Pro Leu Val
 290 295 300
 His Leu Thr Glu Asp Lys Gln Ser Ser Met His Leu Tyr Ile Thr His
 305 310 315 320
 Leu Gln Gly Tyr Thr Leu Glu Gln Arg Asn Glu Val Ile Gln Lys Met
 325 330 335
 Ala Glu Ala Gly Ile Ala Cys Asn Val His Tyr Lys Pro Leu Pro Leu
 340 345 350
 Leu Thr Ala Tyr Lys Asn Leu Gly Phe Glu Met Lys Asp Phe Pro Asn
 355 360 365
 Ala Tyr Gln Tyr Phe Glu Asn Glu Val Thr Leu Pro Leu His Thr Asn
 370 375 380

142

Leu Ser Asp Glu Asp Val Glu Tyr Val Ile Glu Met Phe Leu Lys Ile
 385 390 395 400

Val Ser Arg Asp

<210> 48

<211> 209

<212> PRT

<213> Streptococcus suis

<220>

<223> /note="CPS7H amino acid sequence"

<400> 48

Met Val Glu Arg Asp Met Val Glu Arg Asp Thr Leu Val Ser Ile Ile
 1 5 10 15

Met Pro Ser Trp Asn Thr Ala Lys Tyr Ile Ser Glu Ser Ile Gln Ser
 20 25 30

Val Leu Asp Gln Thr His Gln Asn Trp Glu Leu Ile Ile Val Asp Asp
 35 40 45

Cys Ser Asn Asp Glu Thr Glu Lys Val Val Ser His Phe Lys Asp Ser
 50 55 60

Arg Ile Lys Phe Phe Lys Asn Ser Asn Asn Leu Gly Ala Ala Leu Thr
 65 70 75 80

Arg Asn Lys Ala Leu Arg Lys Ala Arg Gly Arg Trp Ile Ala Phe Leu
 85 90 95

Asp Ser Asp Asp Leu Trp His Pro Ser Lys Leu Glu Lys Gln Leu Glu
 100 105 110

Phe Met Lys Asn Asn Gly Tyr Ser Phe Thr Tyr His Asn Phe Glu Lys
 115 120 125

Ile Asp Glu Ser Ser Gln Ser Leu Arg Val Leu Val Ser Gly Pro Ala
 130 135 140

Ile Val Thr Arg Lys Met Met Tyr Asn Tyr Gly Tyr Pro Gly Cys Leu
 145 150 155 160

Thr Phe Met Tyr Asp Ala Asp Lys Met Gly Leu Ile Gln Ile Lys Asp
 165 170 175

Ile Lys Lys Asn Asn Asp Tyr Ala Ile Leu Leu Gln Leu Cys Lys Lys
 180 185 190

Tyr Asp Cys Tyr Leu Leu Asn Glu Ser Leu Ala Ser Tyr Arg Ile Arg
 195 200 205

```
<210> 49
<211> 120
<212> PRT
<213> Streptococcus suis

<220>
<223> /note="First sequence of the allignment of CPS2J and CPS2K"

<400> 49
```

```
<210> 50
<211> 117
<212> PRT
<213> Streptococcus suis

<220>
<223> /note="Second sequence of the allignment of CPS2J and CPS2K"

<400> 50
```

Met Ile Asn Ile Ser Ile Ile Val Pro Ile Tyr Asn Val Glu Gln Tyr
1 5 10 15
Leu Ser Lys Cys Ile Asn Ser Ile Val Asn Gln Thr Tyr Lys His Ile
20 25 30

144

Glu Ile Leu Leu Val Asn Asp Gly Ser Thr Asp Asn Ser Glu Glu Ile
 35 40 45
 Cys Leu Ala Tyr Ala Lys Lys Asp Ser Arg Ile Arg Tyr Phe Lys Lys
 50 55 60
 Glu Asn Gly Gly Leu Ser Asp Ala Arg Asn Tyr Gly Ile Ser Arg Ala
 65 70 75 80
 Lys Gly Asp Tyr Leu Ala Phe Ile Asp Ser Asp Asp Phe Ile His Ser
 85 90 95
 Glu Phe Ile Gln Arg Leu His Glu Ala Ile Glu Arg Glu Asn Ala Leu
 100 105 110
 Val Ala Val Ala Gly
 115

<210> 51
 <211> 100
 <212> DNA
 <213> Streptococcus suis

<220>
 <223> /note="Allignment of CPS2G and CPS2H within CPS2M
 between CPS20 and CPS2P (first sequence)"

<400> 51

aagggcacct ctataaactc ccaaaattgc gaatttggag ttacgaaagc cttgttaaatt 60
 caacatttta aatttttagaa aattagtttt tagagctccc 100

<210> 52
 <211> 100
 <212> DNA
 <213> Streptococcus suis

<220>
 <223> /note="Allignment of CPS2G and CPS2H within CPS2M
 between CPS20 and CPS2P (second sequence)"

<400> 52

gggggcacct ctataaattc ccaaaattgc gaatttggag ttacgaaagc cttgttaaatt 60
 caacatctta aatttttagaa aattagtttt tagaggctccc 100

<210> 53
 <211> 101
 <212> DNA
 <213> Streptococcus suis

145

<220>

<223> /note="Allignment of CPS2G and CPS2H within CPS2M
between CPS2O and CPS2P (third sequence) "

<400> 53

aagggcacct ctataaactc ccaaaattgc gaatttggag ttacgaaagc cttgttaaatt 60
caaacatttt aaattttaga aaattagttt ttagaggtcc c 101

CLAIMS:

1. An isolated or recombinant nucleic acid comprising a gene of a capsular gene cluster of *Streptococcus suis*, wherein said capsular gene cluster of *Streptococcus suis* is as set forth in SEQ ID NO: 9, SEQ ID NO: 30, SEQ ID NO: 38, or SEQ ID NO: 44.
2. A nucleic acid according to claim 1 comprising a *Streptococcus suis* serotype-specific central region.
3. A nucleic acid capable of hybridising as a probe or primer under stringent conditions to a complement of the nucleic acid according to claim 1 or 2, wherein said hybridisation comprises:
 - hybridising said probe or primer to said complement at 65 °C,
 - washing twice with a solution of 40 mM sodium phosphate, pH 7.2, 1mM EDTA, 5% SDS for 30 minutes at 65 °C, and
 - washing twice with a solution of 40 mM sodium phosphate, pH 7.2, 1 mM EDTA, 1% SDS for 30 minutes at 65 °C.
4. An isolated or recombinant nucleic acid according to claim 1, comprising a gene of a capsular gene cluster of *Streptococcus suis* serotype 2, wherein said capsular gene cluster of *Streptococcus suis* serotype 2 is as set forth in SEQ ID NO: 9.
5. An isolated or recombinant nucleic acid according to claim 1, comprising a gene of a capsular gene cluster of *Streptococcus suis* serotype 1, wherein said capsular gene cluster of *Streptococcus suis* serotype 1 is as set forth in SEQ ID NO: 30.
6. An isolated or recombinant nucleic acid according to claim 1, comprising a gene of a capsular gene cluster of *Streptococcus suis* serotype 9, wherein said capsular gene cluster of *Streptococcus suis* serotype 9 is as set forth in SEQ ID NO: 38.
7. A nucleic acid probe or primer capable of hybridizing under stringent conditions to a nucleic acid according to any one of claims 1 to 6, allowing species or

serotype specific detection of *Streptococcus suis*, wherein said hybridisation comprises:

- hybridising said probe or primer to said nucleic acid at 65 °C,
- washing twice with a solution of 40 mM sodium phosphate, pH 7.2, 1mM EDTA, 5% SDS for 30 minutes at 65 °C, and
- washing twice with a solution of 40 mM sodium phosphate, pH 7.2, 1 mM EDTA, 1% SDS for 30 minutes at 65 °C.

8. A probe or primer according to claim 7 provided with at least one reporter molecule.

9. A diagnostic kit comprising a probe or primer according to claim 7 or 8 in conjunction with instructions for species or serotype specific detection of *Streptococcus suis*.

10. A protein encoded by a nucleic acid according to any one of claims 1, 2 or 4 to 6.

11. Expression products of a capsular gene cluster according to claim 1.

12. A *Streptococcus suis* mutant comprising a recombinant nucleic acid according to any one of claims 1, 2 or 4 to 6 comprising a modification of a capsular gene cluster.

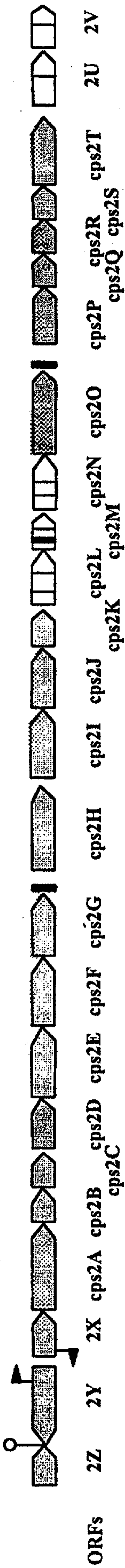
13. A micro-organism, comprising, as a heterologous sequence, a capsular gene cluster of *Streptococcus suis*, wherein said capsular gene cluster of *Streptococcus suis* is as set forth in SEQ ID NO: 9, SEQ ID NO: 30, SEQ ID NO: 38, or SEQ ID NO: 44.

14. A micro-organism according to claim 13 which is a lactic acid bacterium.

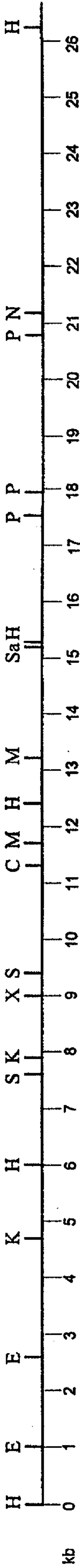
15. A vaccine comprising a mutant according to claim 12 or a micro-organism according to claim 13 or 14.

16. A vaccine according to claim 15 comprising a streptococcus mutant deficient in capsular expression.
17. A vaccine according to claim 15 or 16, wherein said mutant expresses a non-*Streptococcus* protein.
18. A vaccine according to claim 17, wherein said non-*Streptococcus* protein is from a pathogen.
19. Use of a vaccine according to any one of claims 15 to 18 for controlling or eradicating a *Streptococcus suis* disease.
20. Use of a vaccine according to any one of claims 15 to 18 for detecting a *Streptococcus suis* by testing a sample collected from at least one subject for the presence of an encapsulated *Streptococcus suis* strain with said vaccine.
21. Use of a vaccine according to any one of claims 15 to 18 for detecting a *Streptococcus suis* by testing a sample collected from at least one subject for the presence of capsule-specific antibodies directed against a *Streptococcus suis* strain with said vaccine.
22. Use of a vaccine according to any one of claims 15 to 18 for detecting a *Streptococcus suis* by testing a sample collected from at least one subject for the presence of an encapsulated *Streptococcus suis* strain or for the presence of capsule-specific antibodies directed against a *Streptococcus suis* strain with said vaccine.
23. An isolated or recombinant nucleic acid comprising a gene of a capsular gene cluster of *Streptococcus suis* serotype 7, wherein said capsular gene cluster of *Streptococcus suis* serotype 7 is as set forth in SEQ ID NO: 44.

A



B



C

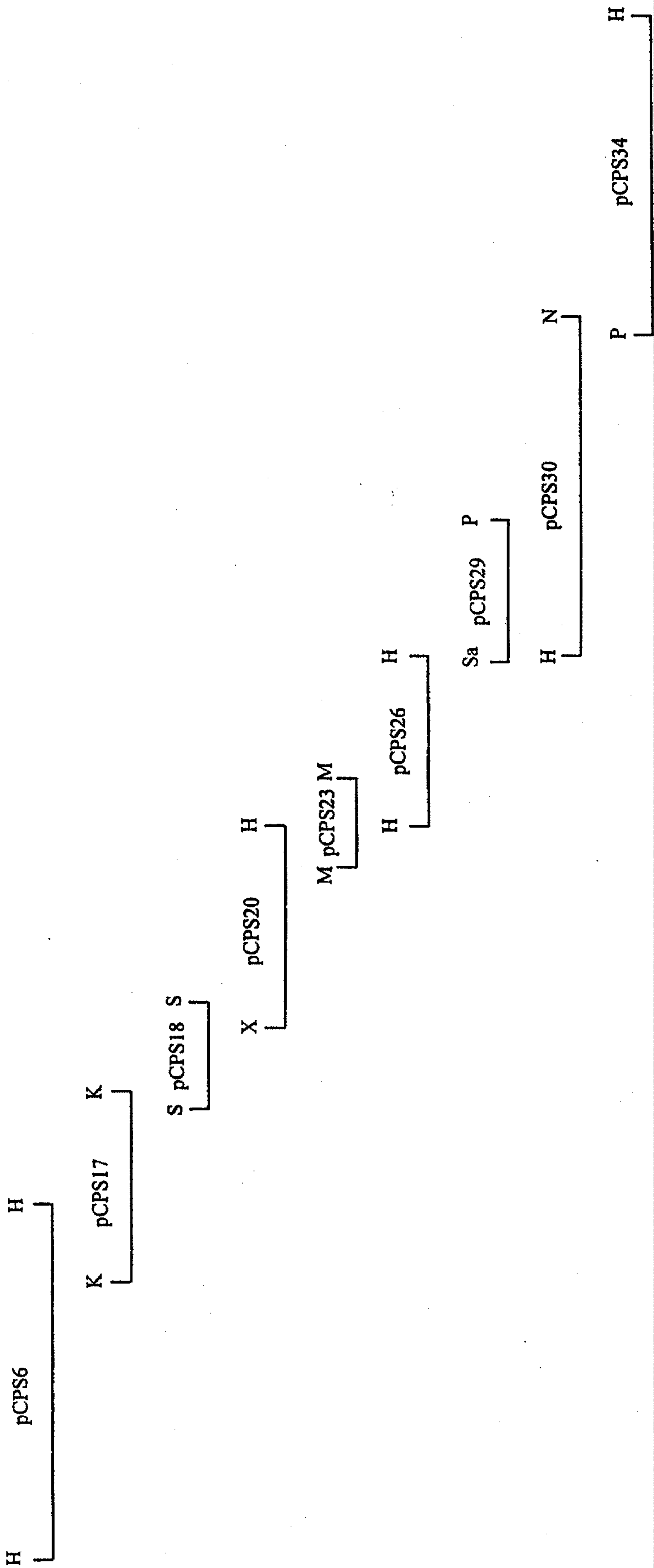


Fig. 1

WO 00/05378

2/59

PCT/NL99/00460

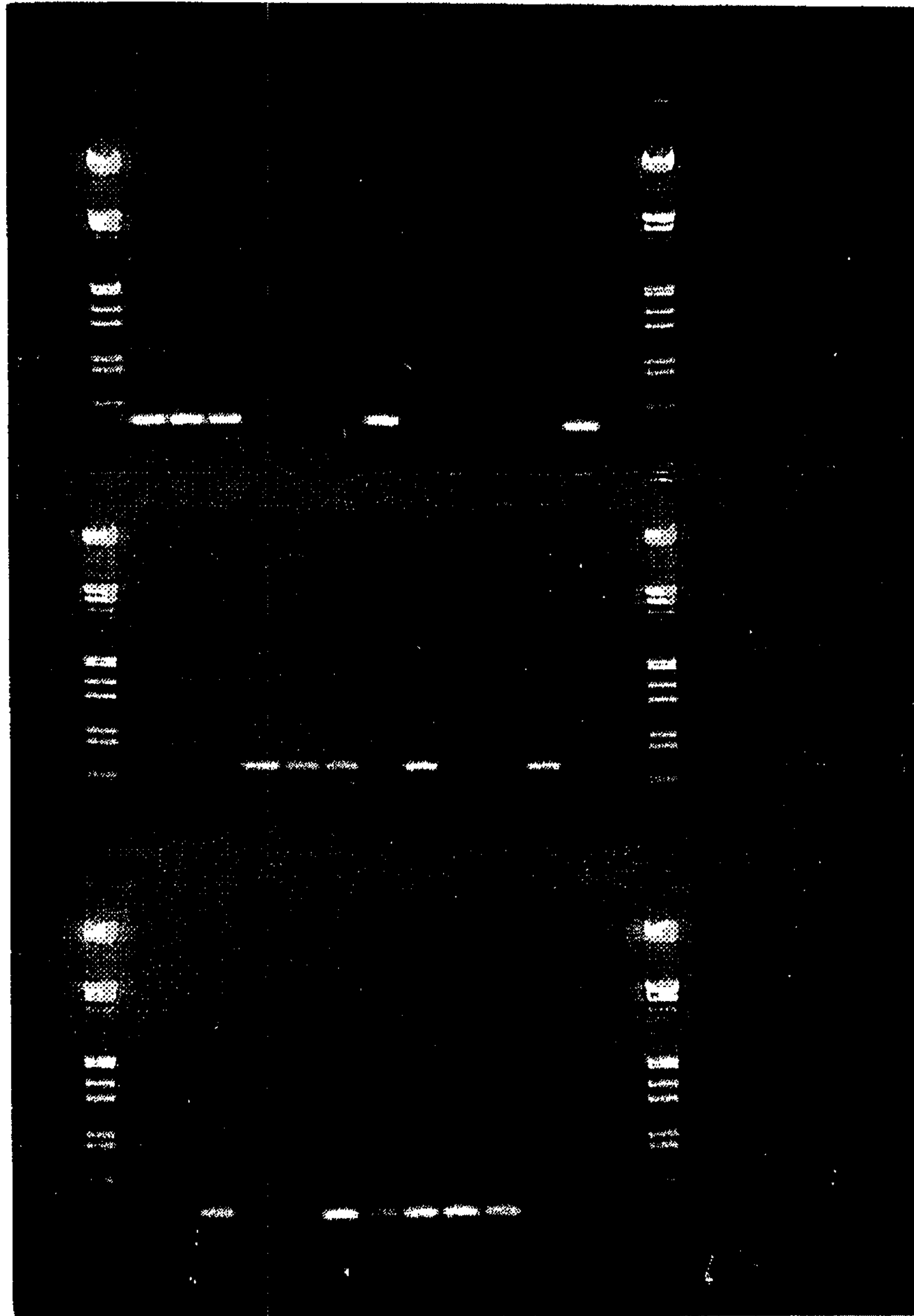


Fig. 2

WO 00/05378

PCT/NL99/00460

3/59

```

AAGCTTGGAT ATTGATCACA TGATGGAGGT GATGGAAGCA TCTAAGTCTG CAGCGGGGTC
GGCGTGCCCA AGTCCGCAGG CTTATCAGGC AGCTTTTGAG GGAGCTGAGA
ACATTATCGT TGTGACGATT ACAGGTGGGC TATCGGGTAG TTTTAATGCG GCACGTGTAG
CTAGGGATAT GTATATCGAA GAGCATCCGA ATGTCAATAT CCATTTGATA
GATAGTTTGT CAGCCAGTGG GGAAATGGAT TTAAGTGTAC ACCAAATCAA TCGCTTAATT
AGTGCAGGAT TAGATTTTCC ACAAGTAGTA GAAGCGATAA CTCACTATCG
GGAACACAGT AAGCTCCTCT TTGTTTTAGC GAAAGTTGAT AATCTTGTTA AGAATGGAAG
ACTGAGCAAA TTGGTAGGCA CTGTCGTTGG TCTTCTCAAT ATCCGTATGG
TTGGTGAGGC AAGTGCTGAA GGAAAATTAG AGTTGCTTCA AAAGGCGCGT GGTCATAAGA
AATCTGTGAC AGCAGCCTTT GAAGAAATGA AAAAAGCAGG CTATGATGGT
GGTCGAATTG TTATGGCCCA CCGCAACAAT GCTAAGTTCT TCCAACAATT CTCAGAGTTG
GTAAAAGCAA GTTTTCCAAC GGCTGTTATT GACGAAGTTG CAACATCAGG
TCTATGCAGT TTTTATGCTG AAGAAGGTGG ACTTTTGATG GGCTACGAAG TGAAAGCGTG
ATTCACAGAG TAATAATTTT GGGCTGTAAT TTCCGCTATA GAATAATCCC
CCTCTTCTTC TAAGTTCGAG GGGGATTGTT TGTATGAGAC TATTGGATTT CATTCATTCA
AATATCTTAC GAATTGCTCC AGTTTATCTG CAAAATCTTG TTCAAAGAAG
ATCTGTAAGA AATCAGCTTT CTGTCCGCTG AAATAATAAC ATTTTCCAAA CATGTGTTGG
ATGCTAGGAG AAAGAATCCC CTTGCTTAGC TGAAAGGTCA CGCTCCCCTT
TGGAATTCGA TACGGGATGT TTAAAGCGTA TTTCTCTAGA CAGTCTTTTA TTTTATTCCA
TTGAGCGTGA TAAATGTGAT GAAGATGCTG TGTGTTCCGC GCAAACATAC
CGTTATCAAT GTAGAGCGAG AGAGCTTTTT GCATGATAAG ATTGGTATCG TAGTCGATTA
GACTCTTATG TTTGATGAAG ATATCACGTA GCTGATTAGG AAGGCTGATT
GCACCGATTC GGAGGGCAGG AAAGAGTGTC GGTGTAAAAG ATTTTATATA GATGACGCGA
TTATCTGTAT CAAGATAGTG TAAAGGTAGG CTATGACTAG AGTCGAAATC
TGCTAAATAG TCATCCTCAA TGATGTAGAC ATCGTATTGC TTTGCTAATT TTACGATGGC
TGTTTTTTGTT GCTATATCAT AGGTGGAACC GAGAGGGTTG TGCAAGCGAG
GAATTGTGTA GAAAAACTTA ATTTTCCAG TTTGGAAGAT ACTTTCCAAT TCTTCTAGGT
CAATTCCATC TAAATTCCGT TCAATTGTTT GATAGGGGAT TCCTTGATGT
CGAATGAGCT CTATCATTCG TGAATAGGTA GGGTTCTCTA TCAAGATTTC CGTTTTTCCA
GCCAAGGTTT CCATTTGTGT GAGAATATAT AGAGCTTGTT GACTACCAGC
TGTGATAACC AGCTGGTCTT TTTTGTATA GACATGATAG TCCATTAACA GACTTTGAAC
GGAGGAAATC AATTCTGCCA ATCCCTCTTG CTGGTGATAG TAGTTGAATA
GGTAATTTTC CCGCCCAATA AGACTTTCTT TTAGACAAAT CCGAAAATCT TCATAGGTAA
TTCTTGAAAG TCTGTAGGAT TGAGCTCTAC AGGTATGGTC TTGGAAATCT
CTATCCTCTA AGATATAATA ACCGCTTTTT TCGACAGCGT AGATCTTATT TTGGTATTTT
AATTCCAACA TAGCCTTTTG GACAGTGTCT TTGCTACAAT GATATTGCTC
GCGGAGTTGA CGGATAGAAG GTAATTTCTC TCCACGTTTG AATCGATGTT CCTCTATTCC
AGTCAAAATA TCTTGATGA TAACTTGATA TTTTTCATC TAGGTCCCCT
TTTTTATAGA CTATGTTACT AGCTAGTATA TAGAAAAAAT TGAAGAAAGA CAATATATGA
ATAATGGGGT TGAGGTTTCA GAATTAAGCT ACTCTATGGT ATAATTAAGT
GATGAAAATA ATTATACCTA ATGCAAAAGA AGTAAATACA AATCTAGAGA ATGCCTCGTT
TTATCTCCTG TCTGATCGAA GCAAGCCGGT GCTGGATGCC ATAAGTCAAT
TTGATGTAAA AAAGATGGCT GCCTTTTATA AATTGAATGA AGCAAAGGCT GAGTTAGAAG
CTGACCGTTG GTATCGAATC AGGACAGGTC AAGCAAAAAC CTATCCAGCC
TGGCAGTTAT ATGATGGTCT CATGTATCGT TATATGGATA GGCGAGGTAT AGATTCGAAA
GAAGAAAATT ATTTACGTGA CCACGTTTCGT GTAGCGACAG CCTTATACGG
ATTGATTCAT CCTTTTGAAT TCATTTCACC TCACCGCTTA GATTTTCAAG GGAGCTTAAA
GATAGGCAAT CAGTCTTTGA AACAGTACTG GCGACCGTAT TATGACCAAG
AAGTTGGTGA TGATGAACTG ATTCTCTCAC TGGCTTCGTC AGAATTTGAG CAGGTGTTTT
CTCCCCAGAT TCAGAAAAGA TTAGTTAAAA TTCTTTTCAT GGAAGAAAAA
GCAGGTGAGC TAAAAGTTCA CTCGACTATA TCAAAAAAAG GCAGAGGAAG ATTGCTGTCC
TGGTTGGCTA AGAACAATAT TCAGGAATTA TCGGACATTC AAGATTTTAA
GGTGGATGGC TTTGAATATT GTACTTCCGA ATCAACGGCA AACCAACTTA CCTTCATACG
ATCAATAAAA ATGTGAAATT ATGAAAAAGA TAACGTTTTT CAGCGCTAAA
AAGGGTAGAA AAATATTAAT TTCTATGATA TAATGGATGC GTTATAGGTA AAAGTCTAGG
AAGGTTGTTT ATGAAAAAGA GAAGCGGACG AAGTAAGTCG TCCAAGTTCA
AATTGGTAAA TTTTGCGCTT TTGGGACTTT ATTCCATTAC TCTATGTTTG TTCTTAGTGA
CCATGTATCG CTATAACATC CTAGATTTCC GGTATTTAAA CTATATTGTG
ACGCTTTTGC TAGTAGGAGT GGCAGTATTG GCTGGATTAT TGATGTGGCG TAAGAAAGCG
CGCATATTTA CAGCGCTCTT ACTTGTTTTT TCACTGGTCA TCACGTCTGT

```

Fig. 3

DNA Serotype 2

WO 00/05378

PCT/NL99/00460

4/59

TGGGATCTAT	GGAATGCAAG	AAGTTGTAAA	ATTTTCAACA	CGACTAAATT	CAAATTTCGAC
ATTTTCAGAA	TATGAAATGA	GTATCCTTGT	CCCAGCAAAT	AGTGATATTA	
CGGACGTTTCG	TCAGCTTACT	AGTATCCTTG	CTCCAGCCGA	ATACGACCAA	GATAACATCA
CCGCTTTTATT	GGATGACATA	TCCAAAATGG	AATCTACTCA	ACTAGCAACT	
AGCCCCGGGA	CTTCTTACCT	GACAGCATAT	CAATCTATGT	TGAATGGCGA	GAGTCAAGCG
ATGGTGTTCA	ACGGAGTTTT	TACCAATATT	TTAGAAAATG	AAGATCCAGG	
CTTTTCTTCA	AAAGTGAAAA	AAATATATAG	TTTCAAAGTG	ACTCAGACTG	TTGAAACAGC
TACTAAGCAG	GTGAGTGGAG	ATAGCTTTAA	TATCTATATT	AGTGGTATTG	
ATGCTTATGG	ACCGATTTCT	ACGGTCTCTC	GTTTCAGATGT	CAATATCATT	ATGACTGTCA
ATCGTGCGAC	ACATAAGATT	TTATTGACAA	CTACTCCACG	AGATTCATAC	
GTTGCTTTTCG	CAGATGGCGG	GCAAAATCAA	TACGATAAAC	TAACACATGC	TGGTATTTAC
GGTGTCAATG	CTTCTGTGCA	CACCTTAGAA	AATTTTTTATG	GGATTGACAT	
TAGCAATTAT	GTGCGGTTGA	ACTTCATTTT	CTTCCTTCAA	TTAATCGACT	TGGTGGGTGG
AATTGATGTA	TATAACGATC	AAGAATTTAC	AAGTTTACAT	GGGAATTATC	
ATTTCCCTGT	TGGACAAGTT	CATTTAAACT	CAGACCAAGC	ATTAGGCTTC	GTTCGAGAGC
GCTACTCTTT	AACAGGGGGT	GACAATGACC	GTGGTAAAAA	CCAGGAAAAA	
GTGATTGCTG	CCTTGATTAA	AAAGATGAGT	ACGCCAGAGA	ATCTAAAAAA	TTACCAGGCA
ATCCTATCTG	GATTGGAAGG	CTCAATTCAA	ACGGATTTGA	GCTTAGAAAC	
GATTATGAGT	TTAGTGAATA	CCCAACTAGA	ATCAGGAACA	CAATTTACAG	TAGAGTCACA
AGCATTGACA	GGAACAGGAC	GCTCAGACTT	ATCTTCTTAT	GCGATGCCTG	
GATCACAAC	TTATATGATG	GAAATTAACC	AAGATAGTCT	GGAGCAATCA	AAGGCAGCGA
TTCAGTCCGT	ACTTGTTGAA	AAATAAAGAT	TTTAGGAGAA	AATATGAACA	
ATCAAGAAGT	AAATGCAATC	GAAATCGATG	TTTTATTCTT	ACTAAAAACA	ATTTGGAGAA
AGAAATTTTT	AATTCTCTTA	ACTGCAGTGT	TGACTGCGGG	GTTGGCATT	
GTCTACAGTA	GTTTTTTAGT	GACACCTCAA	TATGACTCCA	CTACCCGTAT	CTATGTAGTG
AGTCAAAATG	TTGAAGCCGG	TGCGGGCTTG	ACTAACCAAG	AGTTACAAGC	
GGGTACCTAT	TTGGCAAAAG	ACTATCGGGA	AATTATCCTA	TCACAAGATG	TATTGACACA
AGTAGCAACG	GAATTGAATC	TGAAAGAGAG	TTTGAAAGAA	AAAATATCAG	
TTTCTATTCC	TGTTGATACT	CGTATCGTTT	CTATTTCTGT	GCGTGATGCG	GATCCAAATG
AAGCGGCACG	TATTGCAAAT	AGCCTTCGCA	CCTTTGCAGT	GCAAAAGGTT	
GTTGAGGTCA	CCAAGGTAAG	CGATGTGACG	ACACTTGAAG	AAGCAGTCCC	AGCGGAAGAA
CCAACCACTC	CAAATACAAA	ACGAAATATC	TTGCTTGGTT	TATTAGCTGG	
AGGTATCTTG	GCAACAGGTC	TTGTACTGGT	TATGGAGGTT	TTGGATGACC	GTGTAAAACG
TCCTCAGGAC	ATCGAAGAGG	TAATGGGATT	GACATTGCTA	GGTATAGTAC	
CAGATTTCGAA	GAAATTAAAA	TAGGAGAACA	ATATGGCGAT	GTTAGAAATT	GCACGTACAA
AAAGAGAGGG	AGTAAATAAA	ACCGAGGAGT	ATTTCAATGC	TATCCGTACC	
AATATTCAGC	TTAGCGGAGC	AGATATTAAG	GTTGTTGGTA	TTACCTCTGT	TAAATCGAAT
GAAGGTAAGA	GTACAACGTC	GGCTAGTCTC	GCTATTGCCT	ATGCTCGTTC	
AGGTTATAAG	ACCGTCTTGG	TGGATGCAGA	TATCCGAAAT	TCAGTCATGC	CTGGTTTCTT
CAAGCCAATT	ACAAAGATTA	CAGGTTTGAC	GGATTACCTA	GCAGGGACAA	
CAGACTTGTC	TCAAGGATTA	TGCGATACAG	ATATTCCAAA	CTTGACCGTA	ATTGAGTCAG
GAAAGGTTTC	TCCCAACCCT	ACTGCCCTTT	TACAAAGTAA	GAATTTTGAA	
AATCTACTTG	CGACTCTTCG	TCGCTATTAT	GATTATGTTA	TCGTTGACTG	TCCACCATTA
GGACTGGTAA	TTGATGCAGC	TATCATTGCA	CAAAAATGTG	ATGCGATGGT	
TGCAGTAGTA	GAAGCAGGCA	ATGTTAAGTG	CTCATCTTTG	AAAAAAGTAA	AAGAGCAGTT
GGAACAAACA	GGCACACCGT	TCTTAGGCGT	TATCTTGAAC	AAATATGATA	
TTGCCACTGA	GAAGTATAGT	GAATACGGAA	ATTACGGCAA	AAAAGCCTAA	TTTCTCAGAT
AACATAAGTT	TGATAAGTAG	GTATTAATAT	GATTGATATC	CATTCGCATA	
TCATATTTGG	TGTGGATGAC	GGTCCCAAAA	CTATTGAAGA	GAGCCTGAGT	TTGATAAGCG
AAGCTTATCG	TCAAGGTGTT	CGCTATATCG	TAGCGACATC	TCATAGACGA	
AAAGGGATGT	TTGAAACACC	AGAAAAAATC	ATCATGATTA	ACTTTCTTCA	ACTTAAAGAG
GCAGTAGCAG	AAGTTTATCC	TGAAATACGA	TTGTGCTATG	GTGCTGAATT	
GTATTATAGT	AAAGATATCT	TAAGCAAAC	TGAAAAAAG	AAAGTACCAA	CACTTAATGG
CTCGTGCTAT	ATTCTCTTGG	AGTTCAGTAC	GGATACTCCT	TGGAAAGAGA	
TTCAAGAAGC	AGTGAACGAA	ATGACGCTAC	TTGGGCTAAC	TCCCGTACTT	GCCCATATAG
AGCGTTATGA	TGCTCTGGCA	TTTCAGTCAG	AGAGAGTAGA	AAAGCTAATT	
GACAAGGGAT	GCTACACTCA	GGTAAATAGT	AACCATGTGT	TGAAGCCTGC	TTTAATTGGC
GAACGAGCAA	AAGAATTTAA	AAAACGTACT	CGATATTTTT	TAGAGCAGGA	
TTTAGTACAT	TGTGTTGCTA	GCGATATGCA	TAATTTATAT	AGTAGACCTC	CGTTTATGAG
GGAGGCGTAT	CAGCTTGTA	AAAAGAGTA	TGGTGAGGAT	AGAGCGAAGG	

Fig. 3 cont.

WO 00/05378

PCT/NL99/00460

5/59

CTTTGTTCAA GAAAAATCCT TTGTTGATAT TGAAAAATCA AGTACAGTAA CCTCATAGAA
 ATAGTGGAGG AGCTATGAAT ATTGAAATAG GATATCGCCA AACGAAATTG
 GCATTGTTTG ATATGATAGC AGTTACGATT TCTGCAATCT TAACAAGTCA TATACCAAAT
 GCTGATTTAA ATCGTTCTGG AATTTTTATC ATAATGATGG TTCATTATTT
 TGCATTTTTT ATATCTCGTA TGCCGGTTGA ATTTGAGTAT AGAGGTAATC TGATAGAGTT
 TGAAAAAACA TTTAACTATA GTATAATATT TGTAATTTTT CTTATGGCAG
 TTTTATTTAT GTTAGAGAAT AATTTTCGCAC TTTCAAGACG TGGTGCCGTG TATTTACAT
 TAATAAACTT CGTTTTGGTA TACCTATTTA ACGTAATTAT TAAGCAGTTT
 AAGGATAGCT TTCTATTTTC GACAACCTAT CAAAAAAGA CGATTCTAAT TACAACGGCT
 GAACTATGGG AAAATATGCA AGTTTTATTT GAATCAGATA TACTATTTCA
 AAAAAATCTT GTTGCAATTGG TAATTTTAGG TACAGAAATA GATAAAATTA ATTTACCATT
 ACCGCTCTAT TATTCTGTTG AAGAAGCTAT AGGGTTTTCA ACAAGGGAAG
 TGGTCGACTA CGTCTTTATA AATTTACCAA GTGAATATTT TGACTTAAAG CAATTAGTTT
 CAGACTTTGA GTTGTTAGGT ATTGATGTAG GCGTTGATAT TAATTCATTC
 GGTTTTACTG TGTTGAAGAA TAAAAAATC CAAATGCTAG GTGACCATAG CATCGTCACT
 TTTTCCACAA ATTTTTTATA GCCTAGTCAC ATCTGGATGA AACGACTTTT
 AGATATACTT GGAGCAGTAG TCGGGTTAAT TATTAGTGGT ATAGTTTCTA TTTTGTTAAT
 TCCAATTATT CGTAGAGATG GTGGGCCAGC CATTTTTGCT CAGAAACGAG
 TTGGACAGAA TGGACGCATA TTTACATTCT ACAAGTTTCG TTCGATGTTT GTTGATGCCG
 AGGTACGTAA GAAAGAATTA ATGGCTCAAA ACCAGATGCA AGGTGGGATG
 TTCAAAATGG ACAACGATCC TAGAATTACT CCAATTGGAC ACTTCATACG AAAACAAGT
 TTAGATGAGT TACCACAATT TTATAATGTT CTAATTGGAG ATATGAGTCT
 AGTCGGTACC CGTCCGCCTA CAGTTGATGA ATTTGAAAAA TATACTCCTA GTCAAAAGAG
 AAGATTGAGT TTTAAACCAG GGATTACAGG TCTTTGGCAA GTGAGCGGAA
 GAAGTGATAT CACAGATTTT AATGAAGTCG TTAGGCTGGA CCTAACATAC ATTGATAATT
 GGACCATCTG GTCAGACATT AAGATTTTAT TGAAGACAGT GAAAGTTGTA
 TTGTTGAGAG AGGGAGGTCA GTAAGACTCC TTTAAACAA AGAATAGTAG TAGGGGATAT
 GAGAACAGTT TATATTATTG GTTCAAAAGG AATACCAGCA AAGTATGGTG
 GTTTCGAGAC TTTCTAGAGAA AAATTAACCTG AGTATCAGAA AGATAAATCA ATTAATTATT
 TTGTTGCATG TACAAGAGAA AATTCAGCAA AATCAGATAT TACAGGAGAA
 GTTTTTGAAC ATAATGGAGC AACATGTTTT AATATTGATG TGCCAAATAT TGGTTCAGCA
 AAAGCCATTC TTTATGATAT TATGGCTCTC AAGAAATCTA TTGAAATTGC
 CAAAGATAGA AATGATACCT CTCCAATTTT CTACATTCTT GCTTGTCTGA TTGGTCCTTT
 CATTTATCTT TTTAAGAAGC AGATTGAATC AATTGGAGGT CAACTTTTCG
 TAAACCCAGA CGGTCATGAA TGGCTACGTG AAAAGTGGAG TTATCCCGTC CGACAGTATT
 GGAAATTTTC TGAGAGTTTG ATGTTAAAAT ACGCTGATTT ACTAATTTGT
 GATAGCAAAA ATATTGAAAA ATATATTCAT GAAGATTATC GAAAATATGC TCCTGAAACA
 TCTTATATTG CTTATGGAAC AGACTTAGAT AAATCACGCC TTTCTCCGAC
 AGATAGTGTA GTACGTGAGT GGTATAAGGA GAAGGAAATT TCAGAAAATG ATTACTATTT
 GGTGTTGGA CGATTTGTGC CTGAAAATAA CTATGAAGTA ATGATTCGAG
 AGTTTATGAA ATCATATTCA AGAAAAGATT TTGTTTTGAT AACGAATGTA GAGCATAATT
 CCTTTTATGA GAAATTGAAA AAAGAAACAG GGTTCGATAA AGATAAGCGT
 ATAAAGTTTG TTGGAACAGT CTATAATCAG GAGCTGTAA AATATATTCT TGAAAATGCA
 TTTGCTTATT TTCATGGTCA CGAGGTTGGA GGAACGAACC CATCTTTACT
 TGAAGCACTT TCTTCTACTA AACTAAATCT TCTTCTAGAT GTGGGCTTTA ATAGAGAAGT
 AGGGGAAGAA GGAGCGAAAT ACTGGAATAA AGATAATCTT CACAGAGTTA
 TTGACAGTTG TGAGCAATTA TCACAAGAAC AAATTAATGA TATGGATAGT TTATCAACAA
 AACAGTCAA AGAAAGATTT TCTTGGGATT TTATTGTTGA TGAGTATGAG
 AAGTTGTTTA AAGGATAAGT TATGAAAAAG ATTCTATATC TCCATGCTGG AGCAGAATTA
 TATGGGGCAG ATAAGGTTCT CTTGGAACCT ATAAAAGGCT TAGATAAGAA
 TGAATTTGAA GCGCATGTTA TCCTACCTAA TGATGGAGTC CTAGTGCCAG CATTAAGAGA
 AGTTGGTGCG CAAGTTGAAG TTATTAATA TCCAATTCTA CGTAGGAAAT
 ATTTTAATCC AAAAGGGATT TTTGACTACT TCATATCATA TCATCACTAT TCTAAACAGA
 TTGCTCAATA TGCCATAGAA AATAAGGTTG ACATAATTCA CAATAATACT
 ACCGCTGTCT TAGAAGGCAT TTATCTGAAG CGAAAACCTA AATTACCTTT GTTGTGGCAT
 GTTCATGAGA TTATTGTCAA ACCTAAATTC ATCTCTGATT CGATCAATTT
 TTTAATGGGG CGTTTTGCTG ATAAGATTGT GACAGTTTCA CAGGCTGTGG CAAACCATAT
 AAAACAATCA CCTCATATCA AAGATGACCA AATCAGTGTA ATCTACAATG
 GGGTAGATAA TAAAGTGTTT TATCAGTCCG ATGCTCGGTC TGTTGAGAG AGATTTGACA
 TTGACGAAGA GGCTCTTGTC ATTGGTATGG TCGGTCGAGT CAATGCGTGG

Fig. 3 cont.

WO 00/05378

PCT/NL99/00460

6/59

AAAGGACAAG	GAGATTTTTT	AGAAGCAGTT	GCTCCTATAC	TCGAACAGAA	TCCAAAAGCT
ATCGCCTTTA	TAGCAGGAAG	TGCTTTTGAA	GGAGAAGAGT	GGCGAGTAGT	
AGAATTAGAA	AAGAAGATTT	CTCAATTAAA	GGTCTCTTCT	CAAGTCAGAC	GAATGGATTA
TTATGCAAAT	ACCACTGAAT	TATATAATAT	GTTTGATATT	TTTGTACTTC	
CAAGTACTAA	TCCAGACCCT	CTACCAACGG	TTGTACTAAA	AGCAATGGCA	TGCGGTAAAC
CTGTTGTCGG	TTACCGACAT	GGTGGTGTTC	GTGAGATGGT	GAAAGAAGGT	
GTTAACGGTT	TCTTAGTCAC	TCCGAACCTCA	CCGTAAATTT	TATCAAAAGT	AATTCTTCAG
TTATCGGAAA	ATATAAATCT	CAGAAAAAAA	ATTGGTAATA	ATTCTATAGA	
ACGTCAAAAA	GAACATTTTT	CGTTAAAAAG	CTATGTAAAA	AATTTTTTCG	AAGTCTACAC
CTCCCTCAAA	GTATACTGAT	TGGCTGAAGT	GAATGCTTTA	GTATAGCGAT	
TTATCGTATT	CTCATTTCGAT	AAAACAAATG	TTCAGAAACA	GTTATAAGTT	ATTTCTAAAG
GGCACCTCTA	TAAACTCCCA	AAATTGCGAA	TTTGGAGTTA	CGAAAGCCTT	
GTTAAATCAA	CATTTTAAAT	TTTAGAAAAT	TAGTTTTTTAG	AGCTCCCCTA	AAATAGAAGA
TAACAGAAGG	GAGCCTTCAA	AAACTTCATT	TTTAATTGGA	TTGTAGAAAA	
ACTGTTAAAT	CAATATTTAG	ATTTTTAGGA	GTTCAGTTTT	TGGGGGGAGA	GCTTAATAAT
CTATGCACTA	TATTTTCGAAA	AATATATGGT	GTAAATCAG	AACTGATGGT	
CGTGCGAAAA	AAGAGAATGA	GGAATTTATG	AAAATTATTT	CTTTTACAAT	GGTTAATAAC
GAAAGTGAGA	TAATAGAGTC	ATTTATACGG	TATAATTATA	ACTTTATTGA	
CGAGATGGTC	ATTATTGATA	ATGGTTGTAC	AGATAACACG	ATGCAAATTA	TTTTTAATTT
GATTAAAGAG	GGATATAAAA	TATCCGTATA	TGATGAGTCT	TTAGAGGCAT	
ATAATCAGTA	TCGACTTGAT	AATAAATATC	TAACGAAAT	AATTGCTGAA	AAAAATCCAG
ATTTGATAAT	ACCTTTGGAT	GCGGATGAAT	TTTTAACAGC	CGATTCAAAT	
CCACGGAAAC	TTTTGGAACA	ACTGGACTTA	GAAAAGATAC	ATTATGTGAA	TTGGCAATGG
TTTGTTATGA	CTAAAAAAGA	TGATATTAAT	GATTCGTTTA	TACCACGTAG	
AATGCAATAT	TGTTTTGAAA	AACCTGTTTG	GCATCATTCT	GATGGTAAAC	CAGTTACTAA
ATGTATAATT	TCCGCTAAGT	ATTACAAAAA	AATGAATTTA	AAGCTATCGA	
TGGGACATCA	CACTGTTTTT	GGTAACCCAA	ATGTAAGGAT	AGAACATCAT	AATGATTTGA
AATTTGCACA	TTATCGAGCT	ATTAGCCAAG	AGCAATTAAT	TTATAAAACA	
ATTTGTTACA	CTATTCGCGA	TATTGCTACT	ATGGAGAACA	ATATCGAAAC	AGCTCAAAGA
ACAAATCAGA	TGGCGCTCAT	TGAATCTGGC	GTGGATATGT	GGGAAACGGC	
GAGAGAAGCC	TCTTATTCAG	GTTATGATTG	TAATGTTATA	CATGCACCAA	TTGATTTAAG
TTTTTGTAAG	GAAAATATTG	TAATAAAATA	TAACGAACATA	TCCAGAGAAA	
CAGTAGCAGA	ACGCGTGATG	AAAACGGGAA	GAGAAATGGC	TGTTTCGTGCA	TATAATGTGG
AGCGAAACAA	AAAAGAAAAG	AAATTTCTAA	AACCTATTAT	ATTTGTATTA	
GATGGGTTAA	AAGGAGATGA	GTATATTTCAT	CCCAATCCAT	CAAATCATTT	GACGATCTTA
ACTGAAATGT	ATAACGTCAG	AGGCTTACTT	ACCGATAATC	ACCAAATTAA	
ATTTCTCAAA	GTTAATTATA	GATTAATTAT	AACTCCAGAT	TTTGCTAAGT	TTTTACCGCA
TGAATTTATT	GTTGTACCAG	ATACCTTGGA	TATAGAGCAA	GTTAAAAGCC	
AGTATGTTGG	TACAGGTGTA	GACTTGTCAG	AGATTATTTT	TTTAAAAGAG	TATCGAAAAG
AGATAGGCTT	TATTGGTAAT	TTGTATGCGC	TTTTAGGATT	TGTTCCGAAT	
ATGCTCAATA	GAATTTATCT	ATATATTCAG	AGAAACGGTA	TTGCAAACAC	TATTATAAAA
ATCAAGTCGA	GATTGTGAGA	GTTGTTTACT	TTTATTTGTA	ATTTTAAAAG	
TAATGCAGGC	AGATAGGAGA	AAAACGTTTG	GAAAATGAG	AATAAGAATT	AATAATTTGT
TTTTTGTTGC	CATAGCGTTT	ATGGGCATAA	TTATTAGTAA	TTGCAAGTT	
GTTCTAGCGA	TAGGCAAAGC	TTCTGTGATT	CAGTATCTAT	CTTATTTAGT	TTTGATTTTA
TGTATAGTTA	ATGATTTATT	AAAAAATAAC	AAACATATTG	TAGTTTATAA	
ATTAGGGTAT	TTGTTTCTTA	TTATATTTTT	ATTTACTATC	GGAATATGTC	AGCAAATTCT
TCCTATAACA	ACTAAAATAT	ATTTATCAAT	TTCAATGATG	ATTATTTTCAG	
TTTTAGCAAC	GTTGCCAATA	AGTTTGATAA	AAGATATTGA	TGATTTTAGA	CGGATTTCAA
ATCATTTGTT	ATTCGCTCTT	TTTATAACTT	CGATATTAGG	AATAAAGATG	
GGGGCAACGA	TGTTACCGGG	GCGAGTAGAA	GGTATCGGTT	TTAGTCAGGG	TTTTAATGGA
GGATTGACGC	ATAAGAACTT	TTTTGGAATA	ACTATTTTAA	TGGGGTTCGT	
ATTAACCTAC	TTGGCGTATA	AGTATGGTTC	CTATAAAGA	ACGGATCGTT	TTATTTTAGG
ATTAGAATTG	TTTTTGATTC	TTATTTCAAA	CACACGCTCA	GTTTATTTAA	
TACTATTGCT	TTTTCTATTT	CTTGTTAATC	TTGACAAAT	CAAAATAGAA	CAAAGACAAT
GGAGTACGCT	TAAATATATT	TCCATGCTAT	TTTGTGCTAT	TTTTTTATAC	
TATTTCTTTG	GTTTTTTAAT	AACACATAGT	GATTCTTACG	CTCATCGCGT	TAATGGTCTT
ATTAATTTTT	TTGAGTATTA	TAGAAATGAT	TGGTTCCATC	TAATGTTTGG	
TGCAGCGGAT	TTGGCATATG	GGGATTTAAC	TTTAGACTAT	GCTATAAGGG	TTAGACGCGT
TTTAGGTTGG	AATGGAACGC	TTGAAATGCC	CTTACTGAGT	ATTATGTTAA	

Fig. 3. cont.

WO 00/05378

PCT/NL99/00460

7/59

```

AAAATGGTTT TATCGGTCTG GTAGGGTATG GGATTGTTTT ATATAAACTT TATCGTAATG
TAAGAATATT AAAACAGAT AATATAAAA CAATAGGAAA GTCTGTATTT
ATCATTGTAG TCCTATCTGC AACAGTAGAA AATTATATTG TAAATTTAAG TTTTGTATTT
ATGCCAATAT GTTTTTGTTT ATTAAATTCT ATATCTACTA TGGAAATCAAC
TATTAACAAA CAACTGCAAA CATAAATTGG CAGGAATAGA GTTTTGAGTT GCTATTAATT
TGGTAGAGCA TATGTTCTAT AGGTGGCAAG ATAAAGATAG TATTTTTTAC
ATGATGATTT TTATGATAGC AAAGCAAGTT ACGGCATAAA AGGAATTAGA GGATGGAAAA
AGTCAGCATT ATTGTACCTA TTTTAAATAC GGAAAAGTAC TTAAGAGAGT
GTTTAGATAG CATTATTTCC CAATCGTATA CTAATCTAGA GATTCTTTTG ATAGATGACG
GTTCTTCAGA TTCATCAACG GATATATGTT TGGAAATACGC AGAGCAAGAT
GGTAGAATAA AACTTTTCCG GTTACCAAAT GGTGGTGTGT CAAACGCAAG GAATTACGGT
ATCAAAAATA GCACAGCAAA TTATATTATG TTTGTAGATT CTGATGATAT
TGTTGACGGC AACATTGTTG AGTCCTTATA CACCTGTTTA AAAGAGAATG ATAGTGATTT
GTCGGGAGGG TTACTTGCTA CTTTTGATGG AAATTATCAA GAATCTGAGC
TGCAAAAGTG TCAAATTGAT TTGGAAGAGA TAAAGAGAGT GCGAGACTTA GGAAATGAAA
ATTTTCCCAA TCATTATATG AGCGGTATCT TTAATAGCCC TTGTTGCAAA
CTTTATAAGA ATATATATAT AAACCAAGGT TTTGACACTG AACAGTGTTT AGGAGAGGAC
TTATTATTTA ATCTAAATTA TTTAAAGAAT ATAAAAAAG TCCGCTATGT
TAACAGAAAT CTTTATTTTG CCAGAAGAAG TTTACAAAGT ACTACAAATA CGTTTAAATA
TGATGTTTTT ATTCAATTAG AAAATTTAGA AGAAAAAAGT TTTGATTTGT
TTGTTAAAT ATTGTTGGA CAATATGAAT TTTCTGTTTT TAAAGAGACG CTACAGTGGC
ATATTATTTA TTATAGCTTA TTAATGTTCA AAAATGGAGA TGAATCGCTT
CCAAAGAAAT TGCATATATT TAAGTATTTA TACAATAGGC ATTCTTTAGA TACTCTAAGT
ATTAAACGAA CGTCCTCTGT TTTTAAAGA ATATGTAAAT TAATTGTTGC
TAATAATTTG TTTAAATTT TTTTAAATAC TTTAATTAGG GAAGAAAAAA ATAATGATTA
ACATTTCTAT CATCGTCCCA ATTTACAATG TTGAACAATA TCTATCCAAG
TGTATAAATA GCATTGTAAA TCAGACCTAC AAACATATAG AGATTCTTCT GGTGAATGAC
GGTAGTACGG ATAATTCGGA AGAAATTTGT TTAGCATATG CGAAGAAAGA
TAGTCGCATT CGTTATTTTA AAAAAGAGAA CGGCGGGCTA TCAGATGCCC GTAATTATGG
CATAAGTCGC GCCAAGGGTG ACTACTTAGC TTTTATAGAC TCAGATGATT
TTATTCATTC GGAGTTCATC CAACGTTTAC ACGAAGCAAT TGAGAGAGAG AATGCCCTTG
TGGCAGTTGC TGGTTATGAT AGGGTAGATG CTTGCGGGCA TTTCTTAACA
GCAGAGCCGC TTCCTACAAA TCAGGCTGTT CTGAGCGGCA GGAATGTTTG TAAAAAGCTG
CTAGAGGCGG ATGGTCATCG CTTTGTGGTG GCCTGGAATA AACTCTATAA
AAAAGAACTA TTTGAAGATT TTCGATTTGA AAAGGGTAAG ATTCATGAAG ATGAATACTT
CACTTATCGC TTGCTCTATG AGTTAGAAAA AGTTGCAATA GTTAAGGAGT
GCTTGTACTA TTATGTTGAC CGAGAAAATA GTATCATAAC TTCTAGTATG ACTGACCATC
GCTTCCATTG CCTACTGGAA TTTCAAAATG AACGAATGGA CTTCTATGAA
AGTAGAGGAG ATAAAGAGCT CTTACTAGAG TGTTATCGTT CATTTTTAGC CTTTGCTGTT
TTGTTTTTAG GCAAATATAA TCATTGGTTG AGCAAACAGC AAAAGAAGCT
TCTCCAAACG CTATTTAGAA TTGTATATAA ACAATTGAAG CAAAATAAGC GACTTGCTTT
ACTAATGAAT GCTTATTATT TGGTAGGGTG TCTTCATCTT AATTTTAGTG
TCTTTCTGAA AACGGGGAAA GATAAAATTC AAGAAAGATT GAGAAGAAGT GAAAGTAGTA
CTCGGTAAGA ATGTTGTAAT AAATGGTTGA AAGAAAAGGG GATTAAAATG
AATCCAACAA ATAGTAGAAT AGCACTCTTT GATACGATTA AATGTATCAT GGTACTTTGT
GTTATTTTTA CACATCTGGA TTGGTCTGTT GAGCAGCGTC AATGGTTTAT
CTTTCCGTAT TTCGTTGACA TGGCTGTTCC AATTTTTCTG TTGCTTTCTG CCTATTTTCG
AACGAATAAG TGGAATACAA AACAAGAGAC GCTAAAGCTC AAGTTCAGCA
GTGGTATAAA AGAAAGTATA AACATGCTTT GTCTCTATGC TATCGTGATG GCTGTTAATG
TTTTATTGAG CTATTCGAGA ACCATCTGAT AGGAGTAAAG CCTTTTTCAG
GTTCTTCATC GCTCCGTTCA TTTGTCCTGT GGCTACTTTC TGGAGAATCG GGTCCAGGGA
GTTGGGAGTT ACTATGTTCC GTTGTTGATT CAGGTAGTTT TTTTATTACC
AATTTTGTAT GTTCTTTTCG AGAAAAATAA ATGGTTGGGC TTGCTTACTT GTTTTTTAGT
AACTTTTCA GTGGATGCCA TATTTGCTAA CATGGCTGAA CACGGCATAT
ATATATAGAC TAATATCACT TCGTTATCTT TTTGTTCTAG GGCTTGGTTT TTTCTTTCAA
AGCAGGATGT GCGTTCCAAG GTAGATACTT TCATTGCGAC CCTATTTGGG
ATTATTGGAG CAATTCTGAT TTTTGTGAAT CATTCTATAG AGCCCTTCTC CTGGTTTTAT
GGTTGGAAGT CTACTTCCTT TCTATGCGTC CCATTTGCGT ATGCTATGCT
ATTTTTTATG ATAAAGTATG GACAGAAGAT TCCAGCAATA CTGTTGTCAA AATTGGGAGT
TGCTTCTTAT CATATCTACT TGACCCAGAT GCTGTATTTT TCAGTAGTCG

```

Fig. 3 cont.

WO 00/05378

PCT/NL99/00460

8/59

CACCATTTTT	AGCAGTGCAA	TTTAAGGTAT	CTTCGTTGAA	TTTGTGGAAC	GGCTTGTTTA
CCTTTCTAAT	TTGCCTGTTT	GGTGGCTATA	TTTTCTACAA	AGTGGATCTG	
TTTATGAGAG	TACGTGGAAA	ACGATAATGA	CTCATTTTCAG	ATTAGCAGAT	GCCATTTTCGT
TTATTAGCAG	ATTTCGCATGT	TAATATTCCG	ACAAAGAAAT	TCAAATAGGT	
TGACGAGAGA	GGAGTGGTAT	CTGTTTCTAA	ACCCAGATAT	CCCCCTTTAT	TTTCAAAGCT
ATATTTATTA	ACTGAACAAG	GAGAATTTTT	AAGAGAACTG	TTTGTTTAAT	
CCCAGCACGA	TCTGGTTCGA	AAGGCTTACC	GAATAAAAAC	ATGCTATTTT	TGGACGGGAA
ACCCATGATT	TTTCACACGA	TTGATGTGGC	AATTGAATCA	GGTTGTTTTG	
AGAAAGAAGA	CATCTATGTC	AGTACGGATT	CAGAAATGTA	TAAGGGGGGC	ACCTCTATAA
ATTCCCAAAA	TTGCGAATTT	GGAGTTACGA	AAGCCTTGTT	AAATCAACAT	
CTTAAATTTT	AGAAAATTAG	TTTTTAGAGG	TCCCCAAGGG	GATTTGCGAG	ACAAGAGGCA
TCAATGTATT	GTTAAGACCC	AAAGAACTAT	CTACTTATCA	TACTCCATCG	
AATGAAGTCA	GTACGCACTT	TTTTACGAAT	CTGGATTTTA	TGAAGATTGT	ATATTTGTTC
TTCTGCAAGT	CACCTCACCG	TTACGGACTG	GCGAACAGAT	AAAAGAAGCC	
ATGAATATGT	ACTTACAGGG	GGACTCAGAA	AATGTTTTGC	ATTTCAATGA	TGAAGGGCAA
GAAAGAGTGA	ATCAGTACAT	TATCGAAGCT	GTACAGGGGT	TATAAAAAGG	
GGTTACTTAT	CCTTAAAGTC	TGTATGTAGA	AGGAGAAAAA	TTGAGACGAA	TTTATATTTG
CCATACGATG	TATCAGATCC	TGATTTCTTT	GTTAAAGATG	GACGTTGAGA	
GAGATAGTTT	GATGTCCGTT	GATATCATCG	GGCATTTTCC	AGATGTCAGG	GAGCAACTGC
AGCAGCATGT	TCATCTAATC	GAGGGAGACG	GAGCGTTCAT	TTGATCTATA	
TTCTTTGATA	GCTAGATCAA	AAACAAAAGA	ACGCCTTTCC	TTGTTACAGA	GCTATGACGA
GGTGATCATT	TTTCAAGATC	ACCGTCAAGT	CGGTCATTTT	TTAAATAAAC	
ATCGGATTCC	CTATTCTCTT	TTGGAGGATG	GTTATAATTT	TTTCAAGGAT	AAAAGAGTGT
GCGATTTGGA	GTCAATTCAA	TCATCTGTCT	GGAAAAGACT	CTTTTATCAA	
TGGTATTTTA	AACCAACATA	TTTGATTGGT	TCAAGTCTCT	ATTGTCAATC	CATTGAGGTC
AATGATCTGT	CGCTCGTACA	ATTTGACTAG	GCTTATAAAC	CCTTTGTAGA	
AGTTCCGAGA	AAGCAATTAT	TTGATCAAGC	ATCGCCAGAG	AAGGTGCAAG	CGCTGCTGCA
GATATTTGGA	GCAAGGGCGA	TAGTAGCGGA	TGAAGAGTCT	TCTCAAAAAC	
GATTGCTATT	ATTGACCCAG	CCCTTGCTCT	GGGATTATCA	TGTGACCGAA	GAGAGTTGTT
GGAGATTTAT	GTAGCAGGTC	TTGCCCTTAA	TCGGGAAGAC	TATACAATCT	
ACATAAAACC	GCACCCACGA	GATGGGGTTG	ATTATTCAAT	TCTGGGTAAG	GCTGTGGTGC
TTCTGCCTCA	AGGTATTCCG	TTTGAGTTGT	TCGAAATGGC	AGGTAATATC	
CGTTTTGATA	TCGGTATGAC	CTATAGTTCT	TCTGCTTTAG	ATTTTTTAAA	TTGTTTTGAA
GAGAAAGTGT	ATTTAAAGGA	CACTTTTCTT	CTTCTTTCAA	AAAATGATAT	
TTTGCGTGAG	GGGATAGAAT	AGGAGGATTC	ATGTCTAAAA	AATCAATAGT	TGTCTCAGGT
CTCGTCTATA	CGATTGGAAC	CATCCTCGTT	CAGGGATTAG	CCTTCATTAC	
CCTCCCCATC	TATACTCGTG	TCATTTCTCA	GGAAGTATAT	GGGCAGTTTA	GCTTGTATAA
TTCGTGGGTG	GGGCTAGTTG	GTCTCTTTAT	CGGTCTACAG	TTAGGTGGGG	
CTTTTGCCCC	GGGATGGGTA	CACCTCCGCG	AGAAATTTGA	TGATTTTCGT	TCCACCTTGA
TGGTCTCTTC	TATCGCTTTC	TTTTTACCAA	TTTTTGGGCT	ATCTTTTCTC	
CTCAGTCAGC	CCCTATCGCT	CCTATTTGGT	TTGCCTGATT	GGGTCGTTCC	GCTTTACTTT
TTGCAAAGTT	TTATGAGTGT	TGTGCAAGGA	TTTTTTACGA	CCTATTTAGT	
GCAGCGGCAG	CAGTCCATGT	GGACTTTACT	CCTATCGGTA	CTGAGCGCTG	TTATCAACAC
TGCTTTATCT	TTATTTCTCA	TCTTTTTCGAT	GGAGAATGAT	TTCATCGCTC	
GTGTAATGGC	AAACTCGGCA	ACGACTGGTG	TTTTTGCTTG	TGTGTCCTTG	TTGTTTTTCT
ATAAGAAGAT	TGGGCTTCAT	TTTCGAAAGG	ACTATCTTCG	GTATGGTTTA	
AGTATATCGA	TTCTCTTTAT	TTTTCATGGA	TTAGGTCATA	ATGTACTCAA	TCAATTTGAC
AGAATCATGC	TCGGCAAGAT	GCTAACACTG	TCAGATGTAG	CCCTATACAG	
TTTCGGCTAC	ACACTTGCGT	CTATCTTACA	AATTGTGTTT	TCGAGCTTGA	ATACGGTATG
GTGTCCGTGG	TATTTTGAGA	AAAAGAGAGG	TGCAGATAAA	GATTTGCTCA	
GTTATGTCCG	TTACTATCTG	GCGATTGGCC	TGTTTGTGAC	TTTTGGATTT	CTAACAATTT
ACCCTGAATT	AGCGATGTTG	TTAGGTGGAT	CTGAGTATCG	TTTCAGTATG	
GGATTTATTC	CCATGATTAT	TGTCGGGGTG	TTCTTTGTAT	TTCTTTATAG	TTTTCCAGCC
AATATCCAGT	TTTATAGTGG	AAATACAAAG	TTTTTGCCAA	TTGGTACTTT	
TATAGCAGGT	GTAATAAATA	TTTCCGTCCA	CTTTGTTTTG	ATACCGACAA	AGAATTTATG
GTGCTGCTTT	GCAACGACTG	CTTCTATCTT	GTTGTTGCTA	GTCTTGCAAT	
ATTTTGTTGC	TAAGAAAAAG	TATGCTTACG	ATGAAGTTGC	GATTTCAACA	TTTGTTAAGG
TAATTGCTCT	TGTTGTCTGC	TATACAGGCT	TGATGACAGT	ATTTGTCCGT	
TCAATCTGGA	TTCTGTTGGT	ACTAGGAATA	GCGGTTCTAG	TCGTTTATGC	CTACATTTTT
AGAAAGGAAT	TAACAGTTGC	CCTCAATACA	TTCAGGGAAA	AACGGTCTAA	

Fig. 3 cont.

WO 00/05378

PCT/NL99/00460

9/59

ATAAGGGCAC	CTCTATAAAC	TCCCAAATTT	GCGAATTTGG	AGTTACGAAA	GCCTTGTTAA
ATCAAACATT	TTAAATTTTA	GAAAATTAGT	TTTTAGAGGT	CCCCATATAA	
AAACGTCCCA	AATGAGAGGT	GCTCATAAGA	ATTGACCATC	ACTGCCATCT	ACCCAAAGTT
CAAGTATTCT	CTACCATGAA	AATTGTGCTA	TAATCAAGTA	TAAAGAAGGG	
AATGTTTCTT	AAAGGACGTA	TGCGCCTCTG	CTTATGCCAG	AAGTCATGAG	GTAAATCTCC
CTAAAAATTG	GGTAGAAAAG	CAGATTAAAC	TTCCACCAAT	CTATTGAAGA	
TCGTGTTGAA	GAGCAGGCTT	TAGAAGCAAC	AAGCCCTGAG	ACTATTGCAA	AGAAATCTAG
GGCTATTTTT	TCTAATCGGC	TATCAGAAGT	GAAGTAGCGA	TCTTTATTAG	
TGTTCTTTTA	CTACTTAAGG	AAAACCAAGC	TGCTCCCTCA	AGACTTTATG	GGAGCGATTT
ACAGTCATTT	TTAGAAAGGA	AATAAAATGG	TTTATATTAT	TGCAGAAATT	
GGTTGTAATC	ACAACGGTGA	TGTTTATCTA	GCACGGAAAA	TGGTAGAAGT	TGCCGTTGAT
TGTGGTGTGG	ATGCCGTTAA	ATTTTCAACA	TTTAAGGCAG	ATTTGTTGAT	
TTCAAATATC	GCACCAAAGG	CCGAATACCA	AAAAATTACA	ACAGGAGAGT	CAGATTCTCA
GCTCGAAATG	ACTCGTCGTT	TGGAATTGAG	CTTTGAAGAG	TATCTTGATT	
TGCGTGATTA	CTGTCTTGAA	AAGGGAGTTG	ATGTGTTTTT	GACACCTTTT	GATGAGGAAT
CATTGGACTT	CTTGATTAGC	ACAGATATGC	CCGTTTATAA	GATTCCATCT	
GGTGAGATTA	CCAATCTTCC	CTATTTGGAA	AAAATTGGTC	GTCAAGCTAA	GAAAGTTATT
CTTTCAACTG	GTATGGCTGT	TATGGATGAA	ATTCATCAAG	CGGTGAAGAT	
TTTGCAGGAA	AATGGAACGA	CCGATATTTT	GATTTTGCAT	TGTACAACCG	AGTATCCAAC
CCCTTACCCT	GCTTTGAATT	TGAATGTCTT	GCATACCTTG	AAAAAAGAAT	
TTCCAAACTT	AACAATTGGC	TATTCAGACC	ATAGTGTGGG	TTTCAAGATA	CCCATCGCTG
CTGCAGCAAT	GGGAGCTGAA	TTGATTGAAA	AGCACTTTAC	TCTGGACAAT	
GAAATGGAAG	GACCAGATCA	TAAAGCGAGT	GCTACTCCTG	ATATCTTAGC	AGCCTTGGA
AAAGGAGTGA	GGATAGTGGA	ACAATCTCTT	GGTAAATTTG	AAAAAGAGCC	
AGAAGAAGTT	GAAGTACGAA	ATAAAATTGT	AGCTAGAAAA	TCTATTGTTG	CCAAAAAGC
AATTGCTAAA	GGCGAAGTCT	TTACAGAAGA	AAACATCACT	GTCAAAAGAC	
CAGGAAATGG	AATTTTCGCCA	ATGGAATGGT	ACAAAGTCTT	GGGGCAGGTG	AGTGAGCAGG
ATTTTGAGGA	AGACCAAAT	ATTTGCCATA	GTGCTTTTGA	AAATCAAATG	
TAAGCGGAGT	AAGGATGAAA	AAAATTTGTT	TTGTGACAGG	CTCTCGTGCC	GAATATGGGA
TTATGCGTCG	CTTATTGAGC	TATCTACAGG	ATGATCCAGA	AATGGAGCTG	
GATCTTGTA	TGACAGCCAT	GCATCTAGAA	GAAAAATATG	GGATGACGGT	CAAAGACATC
GAAGCGGACA	AGCGTAGGAT	TGTCAAGCGG	ATTCCATTGC	ATTTGACGGA	
TACGTCTAAG	CAGACAATCG	TCAAATCTTT	AGCGACCTTG	ACAGAGCAAC	TCACGGTTCT
TTTTGAAGAA	GTCCAGTATG	ACTTGGTGTT	GATTCTGGGG	GATCGCTATG	
AGATGCTACC	AGTTGCCAAT	GCTGCGTTGC	TTTATAATAT	TCCTATTTGC	CATATTCATG
GTGGTGAAAA	AACCATGGGA	AATTTTGATG	AGTCGATTCT	CCATGCCATT	
ACCAAGATGA	GTCACCTTCA	TCTGACATCA	ACGGATGAAT	TTAGAAATCG	TGTCATTCAA
CTAGGAGAAA	ATCCAACCAT	GTAATGAACA	TGGGAGCTAT	GGGTGTTGAA	
AATGTTTTTA	AACAAGACTT	TTTGACAAGA	GAAGAGTTGG	CGATGGAAT	TGGAATTGAT
TTTGCCGAGG	ATTACTATGT	TGTACTCTTT	CACCCTGTTA	CCTTGGAGGA	
TAACACAGCC	GAAGAACAAA	CGCAGGCCTT	ATTAGATGCT	CTAAAAGAAG	ATGGTAGCCA
GTGTTTGATA	ATTGGATCCA	ATTCGGATAC	ACATGCCGAT	AAGATAATGG	
AATTGATGCA	TGAATTTGTA	AAACAAGACT	CTGATTCTTA	CATCTTTACT	TCGCTTCCAA
CTCGTTATTA	CCATTCCTTG	GTCAAGCATT	CACAAGGTTT	AATAGGGAAT	
TCTTCGTCAG	GTTTGATTGA	AGTGCCCTCA	TTACAGGTTT	CGACCTTAAA	TATTGGAAAT
CGCCAATTTG	GACGTTTGTC	AGGACCGAGT	GTGGTACATG	TTGGAACCTC	
TAAGGAAGCG	ATTGTTGGTG	GTTTGGGGCA	ATTACGTGAT	GTGATAGATT	TTACCAATCC
ATTTGAACAA	CCTGATTCTG	CTTTACAAGG	TTATCGAGCT	ATCAAGGAAT	
TTTTATCTGT	ACAGGCCTCA	ACCATGAAAG	AGTTTTATGA	TAGATAGGGG	AGAAAGTTTG
ATGAAAAAAG	TAGCCTTTCT	AGGAGCGGGT	ACCTTTTTCAG	ATGGTGTCCT	
TCCTTGGTTG	GATAGAACTC	GATATGAACT	CATTGGATAT	TTTGAAGATA	AACCGATCAG
TGACTATCGT	GGCTATCCTG	TATTTGGTCC	CTTGCAAGAT	GTCCTAACCT	
ATTTGGATGA	TGGAAAAGTA	GATGCTGTCT	TCGTCACTAT	AGGTGACAAT	GTCAAGCGCA
AGGAAATCTT	TGACTTGCTT	GCCAAAGATC	ATTATGATGC	TTTGTTCAAC	
ATCATTAGCG	AGCAAGCCAA	TATTTTTTCC	CCAGATAGTA	TCAAGGGACG	AGGGGTTTTT
ATAGGTTTTT	CAAGTTTTGT	AGGAGCCGAT	TCCTATGTCT	ATGACAATTG	
TATCATCAAT	ACGGGTGCCA	TTGTGGAACA	TCATACCACG	GTGGAGGCCC	ATTGTAACAT
TACTCCAGGA	GTGACCATAA	ATGGCTTGTT	CCGTATCGGA	GAAAGCACTT	
ATATTGGAAG	TGGTTCAACA	GTGATTCAAT	GTATCGAGAT	TGCACCTTAT	ACAACATTGG
GGGCAGGGAC	AGTTGTTTTG	AAATCGTTGA	CGGAGTCAGG	GACCTATGTT	

Fig. 3 cont.

WO 00/05378

PCT/NL99/00460

10/59

GGTGTACCTG CTAGAAAGAT TAAATAGGTG AATTGATGGA ACCAATTTGT CTGATTCCTG
CTCGGTCAGG ATCAAAAGGT TTACCAAATA AAAACATGTT ATTTT TAGAT
GGTGTACCGA TGATTTTCCA TACCATTCTGA GCTGCGATTG AGTCTGGATG TTTTAAGAAA
GAAAATATAT ATGTCAGTAC TGATTCAGAG GTTTACAAGG AAATTTGTGA
AACAACTGGG GTTCAAGTCC TCATGCGTCC AGCTGACTTG GCGACAGATT TTACAACCTC
TTTTCAACTG AACGAACATT TTTTACAAGA TTTTCTGAT GACCAAGTAT
TTGTTCTCCT GCAAGTTACG TCCCCATTAA GATCGGGAAA ACATGTCAAG GAGGCGATGG
AGTTATATGG GAAAGGTCAA GCTGACCACG TTGTTAGCTT TACCAAAGTC
GATAAGTCTC CAACATTGTT TTCAACTTTA GACGAAAACG GATTCGCTAA GGATATTGCA
GGATTAGGTG GCAGTTATCG TCGTCAAGAT GAGAAAACAC TCTACTATCC
TAATGGAGCG ATTTATATTT CTTCTAAGCA GGCTTATTTA GCGGATAAAA CTTATTTTTC
TGAAAAACA GCGGCCTATG TGATGACGAA GGAAGATTCTG ATTGATGTAG
ATGATCACTT TGATTTTACT GGTGTTATTG GTCGAATTTA CTTTGATTAC CAGCGTCGTG
AGCAACAAAA CAAACCATT TATAAAAGAG AGTTAAAGCG TTTATGTGAG
CAACGAGTCC ATGATAGTCT TGTGATTGGC GATAGTCGTC TGTTAGCCTT GTTACTGGAT
GGTTTCGATA ATATCAGCAT CGGTGGGATG ACAGCTTCGA CAGCACTTGA
AAACCAAGGT CTCTTTTGG CTTACTCCGAT AAAGAAAGTT TTGCTTTCTC TTGGTGTGAA
TGATTTGATT ACTGACTATC CCTTGCCATAT GATTGAGGAT ACTATTCGCC
AGCTGATGGA AAGTCTTGTT TCCAAAGCAG AGCAGGTTTT TGTGACGACG ATTGCCTACA
CGCTGTTTCG TGATAGCGTT TCCAATGAAG AAATTGTGCA GCTGAATGAC
GTTATTGTTT AGTCAGCAAG TGAAGTGGG ATTTTCAGTGA TTGATCTAAA TGAAGTTGTT
GAAAAAGAGG CGATGCTTGA CTATCAGTAT ACCAATGATG GATTGCATTT
CAATCAGATT GGACAAGAGC GTGTGAATCA GCTGATTTTG ACAAGTTTGA CAAGATAATT
TGGTGATAGA AGCTATTTCA GTGGCTAGAC TATGTTGGTA TGTGTTTTAG
AGCCCAGGAA TAACATCTGT AGAGGATGCT AGCCTTGAGA ATTGACAACC ATTTAGTTGT
TTTAATTATA TAAGGGGACC TCTAAAACT CCCTAAATTT CCCAAAAATG
AGATAATAGA ATAAAAAGTA ATGAGGAGAG CTGTCATGCA TTTATTCACA GACGATGAAA
AAATCTTGTC AAAACTATCA GAGAAAGGCA ATCCCTTAGA ACGTTTGGAT
GCCGTTATGG ATTGGAATAT CTTTCTTCCA TTGTTGTCAG AGTTATTCAG TCGTAAAGAT
AAAGTCATCA GTCGTGGCGG TCGTCCTCAC CTAGACTATC TCATGATGTT
CAAAGCGCTC TTGCTTCAAC GTCTTCATAA CCTATCTGAC GATGCCATGG AATATCAACT
GCTGGATCGT ATATCTTTTC GTCGTTTTGT TGGTTGTCAT GAAGACACTG
TTCCCGATGC GAAAACTATC TGGCTCTATC GTGAGAAATT AACCAAGTCA GGTGCGTGAA
AGGAGTTGTT CGATTTGTTC TATGCCCATC TCACAGATGA AGGGGTGATT
GCCCATTCAG GTCAGATTGT GGATGCTACC TTTGTGCAAT GCCCTAAACA ACGCAATTCA
CGTGAGGACA ATCAGAAAAT CAAAACCTTAT CGAAAATTAT GAGGTCACAA
CAGCTAGTGT ACACGACTCC AATGTCCTAG CTCCTCTTTG TGATGCCAAT GAAGCGGTTT
TTGATGACAG TGCTTATGTT GGAAAATCAG TACCAGAAGG TTGTCGCCAC
CACACGATTC GTCGTGCTTT TAGAAATAAA CCGTTGACTG AGACTGATAA GGTCATTAAT
CGACATATTA CCAAAGTCCG TTGTCGCGTT GAGCATGGTT TTGGCTTCAT
TGAAACTAAC ATGAAAGGTA ACATCTGTCT AGCAATTGGG AAGGCACGAG CTGAAACCAA
TGTGACCTTA ACCAACCTGC TCTACAATAT CTGTCGTTTT GAGCAAATCA
AACGACTGGG ATTACCATCC GTGGGCTTAG TGCGCCCAA AAATAGGAAA ATAAGCAAAA
AGAGGCTGGG CAAAACTAG TTTCTCACAA TAAAAAACG GCTCTTTGTC
AACTGTAGTG GGTAGACGAA AAGCTAACAC CTAGAGAGGA CGAAATTCGT TCTCTCATTT
TTGATGTTTA AAGCGTAACC GCCTAATAAC AAGGTATCTA TCCAATCACA
CATTCCTCCA TTATATAGTT AAATGAAACA AAAACAGTAC ATCTATGATA TAATGTATTT
ATGGCATATT CATTAGATTT TCGTAAAAA GTTCTCGCAT ACTGTGAGAA
AACCAGGAGT ATTACTGAAG CATCAGCTAT TTTCCAAGTT TCACGTAACA CTATCTATCA
ATGGCTAAAA TTAAAAGAGA AAACCGGCGA GCTTCATCAC CAAGTTAAAG
GAACCAAGCC AAGAAAAGTG GATAGAGATA AATTAAAGAA TTATCTTGAA ACTCATCCAG
ATGCTTATTT GACTGAAATA GCTTCTGAAT TTGACTGTCA TCCAACAGCT
ATTCATTACC CCCTCAAAGC TATGGGATAT ACTCGAAAAA AAAGAGCTGT ACCTACTATG
AACAGACCC TGAAAAAGTA GAACTGTTCC TTAAAGAATT GAATAACTTA
AGCCACTTGA CTCCTGTTTA TATTGACGAG ACAGGGTTTG AGACATATTT TCATCGAAAA
TATGGTCGCT CTTTGAAAGG TCAGTTGATA AAAGGTAAGG TCTCTGGAAG
AAGATACCAG CGGATATCTT TAGTAGCAGG TCTCATAAAT GGTGCGCTTA TAGCCCCGAT
GACATACAAA GATACTATGA CGAGTGGCTT TTTCGAAGCT T

Fig. 3 cont.

WO 00/05378

PCT/NL99/00460

11/59

SLDIDHMEVMEASKSAAGSACPSPQAYQAAFEAGAENIIVVTITGGLSGSFNAARVARDM
YIEEHPNVNIHLIDSLASGEMDLLVHQINRLISAGLDFPQVVEAITHYREHSKLLFVLA
KVDNLVKNGRLSKLVGTVVGLLNIRMVGEASAEGKLELLQKARGHKKSVTAAFEEMKKAG
YDGGRIVMAHRNNAKFFQQFSELVKASFPTAVIDEVATSGLC SFYAEEGGLLMGYEVKA

Fig. 3 cont.

ORF2Z

WO 00/05378

PCT/NL99/00460

12/59

MKKYQVIIQDILTGIEEHRFKRGEKLPISIRQLREQYHCSKDTVQKAMLELKYQNKIYAVE
KSGYYILED R D F Q D H T C R A Q S Y R L S R I T Y E D F R I C L K E S L I G R E N Y L F N Y Y H Q Q E G L A E L
I S S V Q S L L M D Y H V Y T K K D Q L V I T A G S Q Q A L Y I L T Q M E T L A G K T E I L I E N P T Y S R M I E L I R
H Q G I P Y Q T I E R N L D G I D L E E L E S I F Q T G K I K F F Y T I P R L H N P L G S T Y D I A T K T A I V K L A K
Q Y D V Y I I E D D Y L A D F D S S H S L P L H Y L D T D N R V I Y I K S F T P T L F P A L R I G A I S L P N Q L R D I
F I K H K S L I D Y D T N L I M Q K A L S L Y I D N G M F A R N T Q H L H H I Y H A Q W N K I K D C L E K Y A L N I P Y
R I P K G S V T F Q L S K G I L S P S I Q H M F G K C Y Y F S G Q K A D F L Q I F F E Q D F A D K L E Q F V R Y L N E

Fig. 3 cont.

ORF2Y

WO 00/05378

PCT/NL99/00460

13/59

MKIIIPNAKEVNTNLENASFYLLSDRSKPVLDAISQFDVKKMAAFYKLNEAKAELEADRW
YRIRTGQAKTYPAWQLYDGLMYRYMDRRGIDSKEENYL RDHVRVATALYGLIHPFEFISP
HRLDFQGS LKIGNQSLKQYWRPYDQEVGDDELILSLASSEFEQVFSPQIQKRLVKILFM
EEKAGQLKVHSTISKKGRGRLLSWLAKNNIQELSDIQDFKVDGFEYCTSESTANQLTFXR
SIKM

Fig. 3 cont.

ORF2X

WO 00/05378

PCT/NL99/00460

14/59

MKKRSGRSKSSKFKLVNFALLGLYSITLCLFLVTMYRYNILDFRYLNIVTLLLVGVAVL
AGLLMWRKKARIFTALLVFSLVITSVGIYGMQEVVKFSTRLNSNSTFSEYEMSILVPAN
SDITDVRQLTSILAPAEYDQDNITALDDISKMESTQLATSPGTSYLTAYQSMLNGESQA
MVFNQVFTNILENEDPGFSSKVKKIYSEKVTQTVETATKQVSGDSFNIYISGIDAYGPIS
TVSRSDVNIIMTVNRATHKILLTTTPRDSYVAFADGGQNQYDKLTHAGIYGVNASVHTLE
NFYGIDISNYVRLNFISFLQLIDLVGIDVYNDQEFTSLHGNYHFPVGQVHLNSDQALGF
VRERYSLTGGDNDRGKNQEKVIAALIKKMSTPENLKNYQAILSGLEGSIQTDLSLETIMS
LVNTQLESGTQFTVESQALTGTGRSDLSSYAMPGSQLYMMEINQDSLEQSKAAIQSVLVE
K

Fig. 3 cont.

CPS2A

WO 00/05378

PCT/NL99/00460

15/59

MNNQEVNAIEIDVLFLLKTIWRKKFLILLTAVLTAGLAFVYSSFLVTPQYDSTTRIYVVS
QNVEAGAGLTNQELQAGTYLAKDYREIILSQDVLTOVATELNLKESLKEKISVSIIPVDTR
IVSISVRDADPNEAARIANSLRTFAVQKVVEVTKVSDVTTLEEAVPAEEPTTPNTRNIL
LGLLAGGILATGLVLVMEVLDDRVKRPQDIEEVMGLTLLGIVPDSKKLK

Fig. 3 cont.

CPS2B

WO 00/05378

PCT/NL99/00460

16/59

MAMLEIARTKREGVNKTEEYFNAIRTNIQLSGADIKVVGITSVKSNEGKSTTAASLAIAY
ARSGYKTVLVDADIRNSVMPGFFKPITKITGLTDYLAGTTDLSQGLCDTDIPNLTVIESG
KVSPNPTALLQSKNFENLLATLRRYYDYVIVDCPPLGLVIDAAIIAQKCDAMVAVVEAGN
VKCSSLKKVKEQLEQTGTPFLGVILNKYDIATEKYSEYGNYGKKA

Fig. 3 cont.

CPS2C

WO 00/05378

PCT/NL99/00460

17/59

MIDIHSHIIFGVDDGPKTIEESLSLISEAYRQGVRYIVATSHRRKGMFETPEKIIMINFL
QLKEAVAEVYPEIRLCYGAELYYSKDILSKLEKKKVPTLNGSCYILLEFSTDTPWKEIQE
AVNEMTLLGLTPVLAHIERYDALAFQSERVEKLIDKGCYTQVNSNHVLKPALIGERAKEF
KKRTRYFLEQDLVHCVASDMHNLYSRPPFMREAYQLVKKEYGEDRAKALFKKNPLLILKN
QVQ

Fig. 3 cont.

CPS2D

WO 00/05378

PCT/NL99/00460

18/59

MNIEIGYRQTKLALFDMIAVTISAILTSHIPNADLNRSGIFIIMMVHYFAFFISRMPVEF
EYRGNLIEFEKTFNYSIIFVIFLMAVSFMLENNFALSRRGAVYFTLINEFVLVYLEFNVIK
QFKDSFLFSTTYQKKTILITTAELWENMQVLFESDILFQKNLVALVILGTEIDKINLPLP
LYYSVEEAIGFSTREVVDYVFINLPSEYFDLKQLVSDFELLGIDVGVDINSEFGFTVLKNK
KIQMLGDHSIVTFSTNFYKPSHIWMKRLLDILGAVVGIIISGIVSILLIPIIRRDGGPAI
FAQKRVGQNGRIFTFYKFRSMFVDAEVRKKELMAQNQMGGMFKMDNDPRITPIGHFIRK
TSLDELPQFYNVLI GDMSLVGTRPPTVDEFEKYTPSQKRRLSFKPGITGLWQVSGRSDIT
DFNEVVRLDLTYIDNWTIWS DIKILLKTVKVLLREGGQ

Fig. 3 cont.

CPS2E

WO 00/05378

PCT/NL99/00460

19/59

MRTVYIIGSKGIPAKYGGFETFVEKLT EYQKDKSINYFVACTRENSAKSDITGEVFEHNG
ATCFNIDVPNIGSAKAILYDIMALKKSIEIAKDRNDTSPIFYILACRIGPFIYLFKKQIE
SIGGQLFVNPDGHEWLREKWSYPVRQYWKFSESLMLKYADLLICDSKNIEKYIHEDYRKY
APETSYIAYGTDLDKSRLSPTDSVVREWYKEKEISENDYYLVVGRFVPENNYEVMIREFM
KSYSRKDFVLI TNVEHNSFYEKLKKETGFDKDKRIKFVGT VYNQELLKYIRENAFAYFHG
HEVGGTNPSLLEALSSTKLNLLLDVGFNREVGEEGAKYWNKDNLHRVIDSCEQLSQEQIN
DMDSLSTKQVKERFSWDFIVDEYEKLFKG

Fig. 3 cont.

CPS2F

WO 00/05378

PCT/NL99/00460

20/59

MKKILYLHAGAELYGADKVLELILKGLDKNEFEAHVILPNDGVLVPALREVGAQVEVINY
PILRRKYFNPKGIFDYFISYHHYSKQIAQYAIENKVDIIHNNTTAVLEGIYLRKRLKLPL
LWHVHEIIVKPKFISDSINFLMGRFADKIVTVSQAVANHIKQSPHIKDDQISVIYNGVDN
KVIFYQSDARSVRERFDIDEEALVIGMVGRVNAWKGGDFLEAVAPIEQNPKAIAFIAGS
AFEGEEWRVVELEKKISQLKVSSQVXRMDYYANTTELYNMFDFVLPSTNPDPLPTVVLK
AMACGKPVVGYRHGGVCEMVKEGVNGFLVTPNSPLNLSKVILQLSENINLRKKIGNNSIE
RQKEHFSLSYVKNFSKVYTSKVVY

Fig. 3 cont.

CPS2G

WO 00/05378

PCT/NL99/00460

21/59

MKIISFTMVNNESEIIESFIRYNYNFIDEMVIIDNGCTDNTMQIIFNLIKEGYKISVYDE
SLEAYNQYRLDNKYLTKIIAEKNPDLIIPLDADEFLTADSNPRKLEQLDLEKIHYVNWQ
WFVMTKKDDINDSFIPRRMQYCFEKPVWHSDGKPVTKCIISAKYYKKMNLKLSMGHHTV
FGNPNVRIEHHNDLKFAHYRAISQEQLIYKTICYTIRDIA TMENNIETAQRTNQMALIES
GVDMWETAREASYSGYDCNVIHAPIDLSFCKENIVIKYNELSR ETVAERVMKTGREMAVR
AYNVERKQKEKKFLKPIIFVLDGLKGDEYIHPNPSNHLTILTEMYNVRGLLTDNHQIKFL
KVNYRLIITPDEAKFLPHEFIVVPD TXDIEQVKSQYVGTGVDLSKIISLKEYRKEIGFIG
NLYALLGFVPNMLNRIYLYIQRNGIANTI IKIKSRL.

Fig. 3 cont.

CPS2H

WO 00/05378

PCT/NL99/00460

22/59

MQADRRKTFGKMRIRINNLFVVAIAFMGIIISNSQVLAIGKASVIQYLSYLVILLCIVN
DLLKNNKHIVVYKLGYLFLIIFLETIGICQOILPITTKIYLSISMMIISVLATLPISLIK
DIDDFRISNHLLEFALFITSILGIKMGATMFTGAVEGIGFSQGFNGGLTHKNFFGITILM
GFVLTYLAYKYGSYKRTDRFILGLELFLILISNTRSVYLILLLEFLVLNLDKIKIEQRQW
STLKYISMLFCAIFLYYFFGFLITHSDSYAHRVNGLINFFEYRNDWFHLMFGAADLAYG
DLTLDYAIRVRRVLGWNGTLEMPLLSIMLKNGFIGLVGYGIVLYKLYRNVRLKTDNIKT
IGKSVFIIVLSATVENYIVNLSFVFMPICFCLLNSISTMESTINKQLQT

Fig. 3 cont.

CPS2I

WO 00/05378

PCT/NL99/00460

23/59

MEKVSIIIVPIFNTEKYLRECLDSIISQSYTNLEILLIDDGSSDSSTDICLEYAEQDGRIK
LFRLPNGGVSARNYGIKNSTANYIMFVDSDDIVDGNIVESLYTCLKENDSDLSGGLLAT
FDGNYQESLQKCQIDLEEIKEVRDLGNENFPNHYMSGIFNSPCKLYKNIYINQGFDE
QWLGEDLLFNLNLYLKNIKKVRYVNRNLYFARRSLQSTTNTFKYDVFIQLENLEEKTFDLF
VKIFGGQYEFVFKETLQWHIIYYSLMFKNNGDESLPKKLHIFKYLYNRHSLDTLSIKRT
SSVFKRICKLIVANNLFKIFLNTLIREEKND

Fig. 3 cont.

CPS2J

WO 00/05378

PCT/NL99/00460

24/59

MINISIIVPI YNVEQYLSKC INSIVNQTYK HIEILLVNDG STDNSEEICL AYAKKDSRIR
YFKKENGGLS DARNYGISRA KGDYLA FIDS DDFIHSEFIQ RLHEAIEREN
ALVAVAGYDR VDASGHFELTA EPLPTNQAVL SGRNVCKKLL EADGHRFVVA WNKLYKKELF
EDFRFEKGKI HEDEYFTYRL LYELEKVAIV KECLYYYVDR ENSIITSSMT
DHRFHCLLEF QNERMDFYES RGDKELLLEC YRSFLAFAVL FLGKYNHWLS KQOKKLLQTL
FRIVYKQLKQ NKRLALLMNA YYLVGCLHLN FSVFLKTGKD KIQERLRSE
SSTR

Fig. 3 cont.

CPS2K

WO 00/05378

PCT/NL99/00460

25/59

MSKKSIVVSG LvyTIGTiLV QGLAFITLPI YTRVISQEVY GQFSLYNSWV GLVGLFIGLQ
LGGAFGPGWV HFREKFDDFV STLMVSSIAF FLPIFGLSFL LSQPLSLLFG
LPDWVPLIF LQSLMIVVQG FFTTYLVQRQ QSMWTLPLSV LSAVINTALS LFLTFPMEND
FIARVMANPA TTGVLACVSX WFSQKKNGLH FRKDYLRYGL SISIPLIFHG
LGHNVLNQFD RIMLGKMLTL SDVALYSFGY TLASILQIVE SSLNTVWCPW YFEKKRGADK
DLLSYVRYYL AIGLFVTFGF LTIYPELAM LGGSEYRFSM GFIPMIIVGV
FEVFLYSFPA NIQFYSGNTK FLPIGTFIAG VLNISVHFVL IPTKNLWCCF ATTASYLLLL
VLHYFVAKKK YAYDEVAIST FVKVIALVVV YTGLMTVFGV SIWIRWSLGI
AVLVVYAYIF RKELTVALNT FREKRSK

Fig. 3 cont.

CPS20

WO 00/05378

PCT/NL99/00460

26/59

MVYIIAEIGC NHNGDVHLAR KMVEVAVDCG VDAVKFQTFK ADLLISKYAP KAEYQKITTG
ESDSQLEMTR RLELSFEEYL DLRDYCLEKG VDVFSTPFDE ESLDFLISTD
MPVYKIPSGE ITNLPYLEKI GRQAKKVILS TGMVMDIEH QAVKILQENG TTDISILHCT
TEYPTYPAL NLNVLHTLKK EFPNLTIGYS DHSVGSEVPI AAAAMGAELI
EKHFTLDNEM EGPDHKASAT PDILAALVKG VRIVEQSLGK FEKEPEEEVEV RNKIVARKSI
VAKKAIKGE VFTEENITVK RPGNGISPME WYKVLGQVSE QDFEEDQNIC
HSAFENQM

Fig. 3 cont.

CPS2P

WO 00/05378

PCT/NL99/00460

27/59

MKKICFVTGS RAEYGIMRRL LSYLQDDPEM ELDLVVTAMH LEEKYGMTVK DIEADKRRIV
KRIPLHLTDT SKQTIVKSLA TLTEQLTVLF EEVQYDLVLI LGDRYEMLPV
ANAALLYNIP ICHIHGGEKT MGNFDESIRH AITKMSHLHL TSTDEFNRNV IQLGENPTMY

Fig. 3 cont.

CPS2Q

WO 00/05378

PCT/NL99/00460

28/59

MELGIDFAED YYVLFHPVT LEDNTAEEQT QALLDALKED GSQCLIIGSN SDTHADKIME
LMHEFVKQDS DSYIFTSLPT RYYHSLVKHS QGLIGNSSSG LIEVPSLQVP
TLNIGNRQFG RLSGPSVVHV GTSKEAIVGG LGQLRDVIDF TNPFEQPDSA LQGYRAIKEF
LSVQASTMKE FYDR

Fig. 3 cont.

CPS2R

WO 00/05378

PCT/NL99/00460

29/59

MKKVAFLGAG TFS DGVLPWL DRTRYELIGY FEDKPISDYR GYPVFGPLQD VLTYLDDGKV
DAVFVTIGDN VKRKEIFDLL AKDHYDALFN IISEQANIFS PDSIKGRGVF
IGFSSEFVGAD SYVYDNCIIN TGAIVEHHTT VEAHCNITPG VTINGLCRIG ESTYIGSGST
VIQCIEIAPY TTLGAGTVVL KSLTESGTYV GVPARKIK

Fig. 3 cont.

CPS2S

WO 00/05378

PCT/NL99/00460

30/59

MEPICLIPAR SGSKGLPNKN MLFLDGVPMI FHTIRAAIES GCFKKENIYV STDSEVYKEI
CETTGVOVLM RPADLATDFT TSFQLNEHFL QDFSDDQVFV LLQVTSPLRS
GKHVKEAMEL YGKGQADHVV SFTKVDKSPT LFSTLDENG F AKDIAGLGGS YRRQDEKTLY
YPNGAIYISS KQAYLADKTY FSEKTAAYVM TKEDSIDVDD HFDFTGVIGR
IYFDYQRREQ QNKPFYKREL KRLCEQRVHD SLVIGDSRLL ALLLDGFDNI SIGGMTASTA
LENQGLFLAT PIKKVLLSLG VNDLITDYPL HMIEDTIRQL MESLVSKAEQ
VFVTTIAYTL FRDSVSNEEI VQLNDVIVQS ASELGISVID LNEVVEKEAM LDYQYTNDGL
HFNQIGQERV NQLILTSLTR

Fig. 3 cont.

CPS2T

WO 00/05378		31/59		PCT/NL99/00460	
ATCGCCAAAC	GAAATTGGCA	TTATTTGATA	TGATAGCAGT	TGCAATTTCT	GCAATCTTAA CAAGTCATAT
ACCAAATGCT	GATTTAAATC	GTTCTGGAAT	TTTTATCATA		
ATGATGGTTC	ATTATTTTGC	ATTTTTTATA	TCTCGTATGC	CAGTTGAATT	TGAGTATAGA GGTAATCTGA
TAGAGTTTGA	AAAAACATTT	AACTATAGTA	TAATATTTGC		
AATTTTTCTT	ACGGCAGTAT	CATTTTTTGT	GGAGAATAAT	TTCGCACTTT	CAAGACGTGG TGCCGTGTAT
TTCACATTAA	TAAACTTCGT	TTTGGTATAC	CTATTTAACG		
TAATTATTAA	GCAGTTTAAG	GATAGCTTTC	TATTTTCGAC	AATCTATCAA	AAAAAGACGA TTCTAATTAC
AACGGCTGAA	CGATGGGAAA	ATATGCAAGT	TTTATTTGAA		
TCACATAAAC	AAATTCAAAA	AAATCTTGTT	GCATTGGTAG	TTTTAGGTAC	AGAAATAGAT AAAATTAATT
TATCATTACC	GCTCTATTAT	TCTGTGGAAG	AAGCTATAGA		
GTTTTCAACA	AGGGAAGTGG	TCGACCACGT	CTTTATAAAT	CTACCAAGTG	AGTTTTTAGA CGTAAAGCAA
TTCGTTTCAG	ATTTTGAGTT	GTTAGGTATT	GATGTAAGCG		
TTGATATTAA	TTCATTCGGT	TTTACTGCGT	TGAAAAACAA	AAAAATCCAA	CTGCTAGGTG ACCATAGCAT
TGTAACTTTT	TCCACAAATT	TTTATAAGCC	TAGTCATATC		
ATGATGAAAC	GACTTTTGGG	TATACTCGGA	GCGGTAGTCG	GGTTAATTAT	TTGTGGTATA GTTTCTATTT
TGTTAGTTCC	AATTATTCGT	AGAGATGGTG	GACCGGCTAT		
TTTTGCTCAG	AAACGAGTTG	GACAGAATGG	ACGCATATTT	ACATTCTACA	AGTTTCGATC GATGTATGTT
GATGCTGAGG	AGCGCAAAAA	AGACTTGCTC	AGCCAAAACC		
AGATGCAAGG	GTGGGTATGT	TTTAAAATGG	GAAAAACGAT	CCTAGAATTA	CTCCAATTGG ACATTTTATA
CGCAAAAACA	AGTTTAGACG	AGTTACCACA	GTTTTATAAT		
GTTTTAATTG	GCGATATGAG	TCTAGTTGGT	ACACGTCCAC	CTACAGTTGA	TGAATTTGAA AAATATACTC
CTGGTCAAAA	GAGACGATTG	AGTTTTAAAC	CAGGGATTAC		
AGGTCTCTGG	CAGGTTAGTG	GTCGTAGTAA	TATCACAGAC	TTCGACGACG	TAGTTCGGTT GGACTTAGCA
TACATTGATA	ATTGGACTAT	CTGGTCAGAT	ATTAAAATTT		
TATTAAAGAC	AGTGAAAGTT	GTATTGTTGA	GAGAGGGAAG	TAAGTAAAAG	TATATGAAAG TTTGTTTGGT
CGGTTCTTCA	GGGGGACATT	TGACTCACTT	GTATTTGTTA		
AAACCGTTTT	GGAAGGAAGA	AGAACGTTTT	TGGGTAAACAT	TTGATAAAGA	GGATGCAAGA AGTCTTTTGA
AGAATGAAAA	AATGTATCCA	TGTTACTTTC	CAACAAATCG		
CAATCTCATT	AATTTAGTGA	AAAATACTTT	CTTAGCTTTC	AAAATTTTAC	GTGATGAGAA ACCAGATGTT
ATTATTTTAT	CTGGTGCGGC	CGTTGCTGTC	CCCTTCTTTT		
ACATCGGAAA	ACTATTTGGA	GCAAAGACGA	TTTATATTGA	AGTATTTGAT	CGAGTTAATA AATCTACATT
AACTGGAAAA	CTAGTTTATC	CCGTAACAGA	TATTTTTTAT		
GTTTCAAGTGG	AAGAAATGAA	GAAGGTATAT	CCTAAATCTA	TTAACTTGGG	GAGTATTTTT TAATGATTTT
TGTAACAGTA	GGAACATCAT	AACAACAGTT	TAATCGATTG		
ATAAAAGAGA	TTGATTTATT	GAAAAAAAAT	GGAAGTATAA	CCGACGAAAT	ATTTATTCAA ACAGGATATT
CTGACTATAT	TCCAGAATAT	TGCAAGTATA	AAAAATTTCT		
CAGTTACAAA	GAAATGGAAC	AATATATTAA	CAAATCAGAA	GTAGTTATTT	GCCACGGAGG CCCCCTACT
TTTATGAATT	CATTATCCAA	AGGAAAAAAA	CAATTATTGT		
TTCCTAGACA	AAAAAAGTAT	GGTGAACATG	TAAATGATCA	TCAAGTAGAG	TTTGTAAGAA GAATTTTACA
AGATAATAAT	ATTTTATTTA	TAGAAAATAT	AGATGATTTG		
TTTGAAAAAA	TTATTGAAGT	TTCTAAGCAA	ACTAACTTTA	CATCAAATAA	TAATTTTTTT TGTGAAAGAT
TAAACAAAT	AGTTGAAAAA	TTTAATGAGG	ATCAAGAAAA		
TGAATAATAA	AAAAGATGCA	TATTTGATAA	TGGCTTATCA	TAATTTTTCT	CAGATTTTAC TGGAGAGGGA
TACAGATATT	ATCATCTTCT	CTCAGGAGAA	TGCACACCAT		
TAGTTCCTTC	AGAATACCTG	TATAATTATT	TTAAATATTC	TCAGGATTTA	TATGTTGAAT TTACAAAAGA
TGAGCAAAAA	TATAAAGAAA	ATAGGATATA	TGAACGAGTT		
AAATGTTACA	GATTATTTCC	TAATATATCA	GAAAAAACTA	TTGATAATGT	ACTGTTTAGA ATTTTATTAA
GAATGTATCG	AGCTTTTGAA	TACTATTTAC	AAAGATTGTT		
GTTTATTGAT	AGAATAAAAA	ACATGGTCTA	AGAATAAGAT	TTGGTTCTAA	TTGGGTTTCG CTTCCACATG
ATTTTGTGGC	AATTCCTTTA	TCAAATGAAA	ACGAAACAGC		
TTATTTATTT	AAGTAATCTA	AATGTCCAGA	TGAACATTTT	ATACAGACAA	TTATAGAAAA ATATGAATTT
TCAAATAGAT	TATCTAAATA	TGGAAATTTA	AGATATATAA		
AGTGGAAAAA	ATCAACATCT	TCTCCTATTG	TCTTTACAGA	TGATTCTATT	GATGAATTGC TAAATGCAAG
AAATTTAGGT	TTTTTATTTG	CTAGAAAGTT	AAAAATAGAA		
AATAATCTA	AATTTAAAGA	AATTATTACT	AAAAAATAAA	ATAGTTGATT	TTGTGAGAGT AATGTATGTT
TAAATTATTT	AAATATGACC	CGGAATATTT	TATTTTAAAG		
TACTTCTGGT	TGATTATTTT	TATTCAGAG	CAAAAGTATG	TATTTTTATT	AATTTTTATG AATTTAATTT
TATTTTCATAT	AAAATTTTTG	AAAACATAAGC	TAATATTAAA		
AAATGAAATT	TTATTGTTTT	TATTATGGTC	TATATTATGT	TTTGTTTCAG	TAGTCACAAG TATGTTTGT
GAAATAAATT	TTGAAAGATT	ATTTGCAGAT	TTTACTGCTC		
CCATAATTTG	GATTATTGCA	ATAATGTATT	ATAATTTGTA	TTCAATTTATA	AATATTGATT ATAAAAAATT
AAAAAATAGT	ATCTTTTTTTA	GTTTTTTTAGT	TTTATTAGGT		
ATATCTGCAT	TGTATATTAT	TCAAATGGG	AAAGATATTG	TATTTTTAGA	CAGACACCTT ATAGGACTAG
ACTATCTTAT	AACAGGCGTC	AAAACAAGGT	TGGTTGGCTT		
TATGAACAT	CCTACGTAA	ATACCACTAC	AATTATAGTT	TCAATTCGGT	TAATCTTTGC ACTTATAAAA
AATAAAATGC	AACAATTTTT	TTTCTTGTGT	CTTGCTTTTA		

Fig. 4

WO 00/05378

PCT/NL99/00460

32/59

TACCGATCTA TTTAAGTGGA TCGAGAATTG GTAGTTTATC GCTAGCAATA TTAATTATAT GCTTGTTATG
GAGATATATA GGTGGAAAAT TTGCTTGGAT AAAAAAGCTA
ATAGTAATAT TTGTAATACT ACTTATTATT TTAAATACTG AATTGCTTTA CCATGAAATT TTGGCTGTTT
ATAATTCTAG AGAATCAAGT AACGAAGCTA GATTTATTAT
TTATCAAGGA AGTATTGATA AAGTATTAGA AAACAATATT TTATTTGGAT ATGGAATATC CGAATATTCA
GTTACGGGAA CTTGGCTCGG AAGTCATTCA GGCTATATAT
CATTTTTTTA TAAATCAGGA ATAGTTGGGT TGATTTTACT GATGTTTTCT TTTTTTTATG TTATAAAAAA
AAGTTATGGA GTTAATGGGG AAACAGCACT ATTTTATTTT
ACATCATTAG CCATATTTTT CATATATGAA ACAATAGATC CGATTATTAT TATATTAGTA CTATTCTTTT
CTTCAATAGG TATTTGGAAT AATATAAATT TAAAAAGGA
TATGGAGACA AAAAATGAAT GATTTAATTT CAGTTATTGT ACCAATTTAT AATGTCCAAG ATTATCTTGA
TAAATGTATT AACAGTATTA TTAACCAAAC ATATACTAAT
TTAGAGGTTA TTCTCGTAAA TGATGGAAGT ACTGATGATT CTGAGAAAAT TTGCTTAAAC TATATGAAGA
ACGATGGAAG AATTAAATAT TACAAGAAAA TTAATGGCGG
TCTAGCAGAT GCTCGAAATT TCGGACTAGA ACATGCAACA GGTAAATATA TTGCTTTTGT CGATTCTGAT
GACTATATAG AAGTTGCAAT GTTCGAGAGA ATGCATGATA
ATATAACTGA GTATAATGCC GATATAGCAG AGATAGATTT TTGTTTAGTA GACGAAAACG GGTATACAAA
GAAAAAAGA AATAGTAATT TTCATGTCTT AACGAGAGAA
GAGACTGTAA AAGAATTTTT GTCAGGATCT AATATAGAAA ATAATGTTTG GTGCAAGCTT TATTCACGAG
ATATTATAAA AGATATAAAA TTCCAAATTA ATAATAGAAG
TATTGGTGAG GATTTGCTTT TTAATTTGGA GGTCTTGAAC AATGTAACAC GTGTAGTAGT TGATACTAGA
GAATATTATT ATAATTATGT CATTCGTAAC AGTTCGCTTA
TTAATCAGAA ATTCTCTATA AATAATATTG ATTTAGTCAC AAGATTGGAG AATTACCCCT TTAAGTTAAA
AAGAGAGTTT AGTCATTATT TTGATGCAAA AGTTATTAAA
GAGAAGGTTA AATGTTTAAA CAAAATGTAT TCAACAGATT GTTTGGATAA TGAGTTCTTG CCAATATTAG
AGTCTTATCG AAAAGAAATA CGTAGATATC CATTTATTAA
AGCGAAAAGA TATTTATCAA GAAAGCATTT AGTTACGTTG TATTTGATGA AATTTTCGCC TAAACTATAT
GTAATGTTAT ATAAGAAATT TCAAAAGCAG TAGAGGTAAA
AATGGATAAA ATTAGTGTTA TTGTTCCAGT TTATAATGTA GATAAATATT TAAGTAGTTG TATAGAAAGC
ATTATTAATC AAAATTATAA AAATATAGAA ATATTATTGA
TAGATGATGG CTCTGTAGAT GATTCTGCTA AAATATGCAA GGAATATGCA GAAAAAGATA AAAGAGTAAA
AATTTTTTTC ACTAATCATA GTGGAGTATC AAATGCTAGA
AATCATGGAA TAAAGCGGAG TACAGCTGAA TATATTATGT TTGTTGACTC TGATGATGTT GTTGATAGTA
GATTAGTAGA AAAATTATAT TTTAATATTA TAAAAAGTAG
AAGTGATTTA TCTGGTTGTT TGTACGCTAC TTTTTCAGAA AATATAAATA ATTTTGAAGT GAATAATCCA
AATATTGATT TTGAAGCAAT TAATACCGTG CAGGACATGG
GAGAAAAAAA TTTTATGAAT TTGTATATAA ATAATATTTT TTCTACTCCT GTTTGTAAAC TATATAAGAA
AAGATACATA ACAGATCTTT TTCAAGAGAA TCAATGGTTA
GGAGAAGATT TACTTTTTAA TCTGCATTAT TTAAGAATA TAGATAGAGT TAGTTATTTG ACTGAACATC
TTTATTTTTA TAGGAGAGGT AACTAAGTA CAGTAAATTC
TTTTAAAGAA GGTGTGTTTT TGCAATTGGA AAATTTGCAA AAACAAGTGA TAGTATTGTT TAAGCAAATA
TATGGTGAGG ATTTTGACGT ATCAATTGTT AAAGATACTA
TACGTTGGCA AGTATTTTAT TATAGCTTAC TAATGTTTAA ATACGGAAAA CAGTCTATTT TTGACAAATT
TTTAATTTTT AGAAATCTTT ATAAAAAATA TTATTTTAAAC
TTGTTAAAAG TATCTAACAA AAATTCTTTG TCTAAAAATT TTTGTATAAG AATTGTTTCG AACAAAGTTT
TTAAAAAAT ATTATGGTTA TAATAGGAAG ATATCATGGA
TACTATTAGT AAAATTTCTA TAATTGTACC TATATATAAT GTAGAAAAAT ATTTATCTAA ATGTATAGAT
AGCATTGTAA ATCAGACCTA CAAACATATA GAGATTCTTC
TGGTGAATGA CGGTAGTACG GATAATTCGG AAGAAATTTG TTTAGCATAT GCGAAGAAAG ATAGTCGCAT
TCGTTATTTT AAAAAAGAGA ACGGCGGGCT ATCAGATGCC
CGTAATTATG GCATAAGTCG CGCCAAGGGT GACTACTTAG CTTTTATAGA CTCAGATGAT TTTATTCATT
CGGAGTTCAT CCAACGTTTA CACGAAGCAA TTGAGAGAGA
GAATGCCCTT GTGGCAGTTG CTGGTTATGA TAGGGTAGAT GCTTCGGGGC ATTTCTTAAC AGCAGAGCCG
CTTCCTACAA ATCAGGCTGT TCTGAGCGGC AGGAATGTTT
GTAAAAAGCT GCTAGAGGCG GATGGTCATC GCTTTGTGGT GGCCTGTAAT AAACCTCTATA AAAAAGAACT
ATTTGAAGAT TTTGATTG AAAAGGGTAA GATTCATGAA
GATGAATACT TCACTTATCG CTTGCTCTAT GAGTTAGAAA AAGTTGCAAT AGTTAAGGAG TGCTTGTAAT
ATTATGTTGA CCGAGAAAAT AGTATCACAA CTTCTAGCAT
GACTGACCAT CGCTTCCATT GCCTACTGGA ATTTCAAAT GAACGAATGG ACTTCTATGA AAGTAGAGGA
GATAAAGAGC TCTTACTAGA GTGTTATCGT TCATTTTTAG
CCTTTGCTGT TTTGTTTTTA GGCAATATA ATCATTGGTT GAGCAAACAG CAAAAGAAGC TT

Fig. 4 cont.

WO 00/05378

PCT/NL99/00460

33/59

ROTKLALFDM	IAVAISAILT	SHIPNADLNR	SGIFIIMMVH	YFAFFISRMP	VEFEYRGNLI
EFEKTFNYSI	IFAIFLTAVS	FLENNFALS	RRGAVYFTLI	NFVLVYLFNV	
IIKQFKDSFL	FSTIYQKKT	LITTAERWEN	MOVLFESHKQ	IQKNLVALVV	LGTEIDKINL
SLPLYYSVEE	AIEFSTREVV	DHVFINLPSE	FLDVKQFVSD	FELLGIDVSV	
DINSFGFTAL	KNKKIQLLGD	HSIVTFSTNF	YKPSHIMMKR	LLDILGAVVG	LIICGIVSIL
LVPIIRRDGG	PAIFAQKRVG	QNGRIFTFYK	FRSMYVDAEE	RKKDLLSQNQ	
MOGWVCFKMG	KTILELLQLD	ISYAKTSLDE	LPQFYNVLIG	DMSLVGTRPP	TVDEFEKYTP
GQKRRLSFKP	GITGLWQVSG	RSNITDFDDV	VRLDLAYIDN	WTIWSDIKIL	
LKTVKVLLR	EGSK				

Fig. 4 cont.

CPS1E

WO 00/05378

PCT/NL99/00460

34/59

MKVCLVGSSG GHLTHLYLLK PFWKEEERFW VTFDKEDARS LLKNEKMYP C YFPTNRNLIN
LVKNTFLAFK ILRDEKPDVI ISSGA AVVP FFYIGKLFGA KTIYIEVFDR
V NKSTLTGKL VYPVTDIFIV QWEEMKKVYP KSINLGSIF

Fig. 4 cont.

CPS1F

WO 00/05378

35/59

PCT/NL99/00460

MIFVTVGTHE QQFNRLIKEI DLLKKNCSIT DEIFIQTGYS DYIPEYCKYK KFLSYKEMEQ
YINKSEVVIC HGGPATFMNS LSKGKKQLLF PRQKKYGEHV NDHQVEFVRR
ILODNNILFI ENIDDLFEKI IEVSKQTNFT SNNNFFCERL KQIVEKFED QENE

Fig. 4 cont.

CPS1G

WO 00/05378

PCT/NL99/00460

36/59

MFKLFKYDPE YFIFKYFWLI IFIPEQKYVF LLIFMNLILF HIKFLKTRLI LKNEILLELL
WSILCFVSVV TSMFVEINFE RLEADFTAPI IWIIAIMYYN LYSFINIDYK
KLKNSIFFSF LVLLGISALY IIQNGKDIVF LDRHLIGLDY LITGVKTRLV GFMNYPTLNT
TTIIVSIPLI FALIKNKMQQ FFFLCCLAFIP IYLSGSRIGS LSPLAILIIC
LLWRYIGGKF AWIKKLIVIF VILLIILNTE LLYHEILAVY NSRESSNEAR FIIYQGSIDK
VLENNILFGY GISEYSVTGT WLGSHTSGYIS FFYKSGIVGL ILLMFSFFYV
IKKSYGVNGE TALFYFTSLA IFFIYETIDP IIIILVLFFS SIGIWNNINF KKDMETKNE

Fig. 4 cont.

CPS1H

WO 00/05378

PCT/NL99/00460

37/59

MNDLISVIVP IYNVQDYLDK CINSIINQTY TNLEVILVND GSTDDSEKIC LNYMKNDGRI
KYYKKINGGL ADARNFGLH ATGKYIAFVD SDDYIEVAMF ERMHDNITEY
NADIAEIDFC LVDENGYTKK KRNSNFHVL T REETVKEFLS GSNIENNVWC KLYSRDIIKD
IKFQINNRSI GEDLLFNLEV LNNVTRVVVD TREYYNYVI RNSSLINQKF
SINNIDLVR LENYPFKLKR EFSHYFDKV IKEKVKCLNK MYSTDCLDNE FLPILESYRK
EIRRYPFIKA KRYLSRKHLV TLYLMKFSPK LYVMLYKKFQ KQ

Fig. 4 cont.

CPS11

WO 00/05378

PCT/NL99/00460

38/59

MDKISVIVPV YNVDKYLSSC IESIINQNYK NIEILLIDDG SVDDSAKICK EYEKDKRVKI
FFTNHSGVSN ARNHGIKRST AEYIMEVDSD DVVDSRLVEK LYFNIIKSRS
DLSGCLYATF SENINNFEVN NPNIDFEAIN TVQDMGEKNF MNLXXNNIFS TPVCXLYQKR
YITDLFQENQ WLGEDLLFNL HYLKNIDRVS YLTEHLYFYR RGILSTVNSF
KEGVFLQLEN LQKQVIVLFK QIYGEDFDVS IVKDTIRWQV FYYSLLMFKY GKQSIFDKFL
IFRNLYKKYY FNLLKVSNNK SLSKNFCIRI VSNKVFKKIL WL

Fig. 4 cont.

CPS1J

WO 00/05378

PCT/NL99/00460

39/59

MDTISKISII VPIYNVEKYL SKCIDSIVNQ TYKHIEILLV NDGSTDNSEE ICLAYAKKDS
RIRYFKKENG GLSDARNYGI SRAKGDYLAF IDSDDFIHSE FIQRLHEAIE
RENALVAVAG YDRVDAAGHF LTAEPLPTNQ AVLSGRNVCK KLEADGHRF VVACNKLYKK
ELFEDFRFEK GKIHEDEYFT YRLLYELEKV AIVKECLYYY VDRENSITTS
SMTDHRFHCL LEFQNERMDF YESRGDKELL LECYRSFLAF AVLFLGKYNH WLSKQOKK

Fig. 4 cont.

CPS1K

WO 00/05378

PCT/NL99/00460

40/59

AAGCTTATCG	TCAAGGTGTT	CGCTATATCG	TGGCGACATC	TCATAGACGA	AAAGGGATGT
TTGAAACACC	AGAAAAAGTT	ATCATGACTA	ACTTTCTTCA	ATTTAAAGAC	
GCAGTAGCAG	AAGTTTATCC	TGAAATACGA	TTGTGCTATG	GTGCTGAATT	GTATTATAGT
AAAGATATAT	TAAGCAAAC	TGAAAAAAG	AAAGTACCCA	CACTTAATGG	
CTCGCGCTAT	ATTCTTTTGG	AGTTCAGTAG	TGATACTCCT	TGGAAAGAGA	TTCAAGAAGC
AGTGAACGAA	GTGACGCTAC	TTGGGCTAAC	TCCCGTACTT	GCCCATATAG	
AACGATATGA	CGCCCTAGCG	TTTCATGCAG	AGAGAGTAGA	AGAGTTAATT	GACAAGGGAT
GCTATACTCA	GGTAAATAGT	AATCATGTGC	TGAAGCCAC	TTTAATTGGT	
GATCGAGCAA	AAGAATTTAA	AAAACGTA	CGGTATTTTT	TAGAGCAGGA	TTTAGTACAT
TGTGTTGCTA	GCGATATGCA	TAATTTATCT	AGTAGACCTC	CGTTTATGAG	
GGAGGCTTAT	AAGTTGCTAA	CAGAGGAATT	TGGCAAAGAT	AAAGCGAAAG	CGTTGCTAAA
AAAGAATCCT	CTTATGCTAT	TAAAAAACCA	GGCGATTTAA	ACTGGTTACT	
CTAGATTGTG	GAGAGAAAAA	TGGATTTAGG	AACTGTTACT	GATAAACTGT	TAGAACGCAA
CAGTAAACGA	TTGATACTCG	TGTGCATGGA	TACGTGTCTT	CTTATAGTTT	
CCATGATTTT	GAGCAGACTG	TTTTTGGATG	TTATTATTGA	CATACCAGAT	GAACGCTTCA
TTCTTGCACT	TTTATTCGTA	TCAATTTTAT	ATTTGATTCT	ATCGTTTGA	
TTAAAAGTCT	TTTCATTAAT	TACGCGTTAC	ACAGGGTATC	AGAGTTATGT	AAAAATAGGA
CTTAGTTTAA	TATCTGCGCA	TTCATTGTTT	TTAATTATCT	CAATGGTGTT	
GTGGCAGGCT	TTTAGTTATC	GTTTCATCTT	AGTATCCTTA	TTTTTGTCGT	ATGTAATGCT
CATTACTCCG	AGGATTGTTT	GGAAAGTCTT	ACATGAGACG	AGAAAAAATG	
CTATCCGTAA	GAAGGATAGC	CCACTAAGAA	TCTTAGTAGT	AGGTGCTGGA	GATGGTGGTA
ATATTTTTAT	CAATACTGTC	AAAGATCGAA	AATTGAATTT	TGAAATTGTC	
GGTATCGTTG	ATCGTGATCC	AAATAAACTT	GGAACATTTA	TCCGTACGGC	TAAAGTTTAA
GGAAACCGTA	ATGATATTCC	ACGACTGGTA	GAGGAATTAG	CTGTTGACCA	
AGTGACGATT	GCCATCCCTT	CTTTAAATGG	TAAGGAGCGA	GAGAAGATTG	TTGAAATCTG
TAACACTACA	GGAGTGACCG	TCAATAATAT	GCCGAGTATT	GAAGACATTA	
TGGCGGGGAA	CATGTCTGTC	AGTGCCTTTC	AGGAAATTGA	CGTAGCAGAC	CTTCTTGGTC
GACCAGAGGT	TGTTTTGGAT	CAGGATGAAT	TGAATCAGTT	TTTCCAAGGG	
AAAACAATCC	TTGTCACAGG	AGCAGGTGGC	TCTATCGGTT	CAGAGCTATG	TCGTCAAATT
GCTAAGTTTA	CGCCTAAACG	CTTGTTGTTG	CTTGACATG	GAGAAAATTC	
AATCTATCTC	ATTCATCGAG	AGTTACTGGA	AAAGTACCAA	GGTAAGATTG	AGTTGGTCCC
TCTCATTGCA	GATATTCAAG	ATAGAGAATT	GATTTTTAGC	ATAATGGCTG	
AATATCAACC	CGATGTTGTT	TATCATGCTG	CAGCACATAA	GCATGTTTCT	TTGATGGAAT
ATAATCCACA	TGAAGCAGTG	AAGAATAATA	TTTTTGGAAC	GAAGAATGTG	
GCTGAGGCGG	CTAAAAC	AAAGGTTGCC	AAATTTGTTA	TGGTTTCAAC	AGATAAAGCT
GTTAATCCAC	CAAATGTCAT	GGGAGCGACT	AAACGTGTTG	CAGAAATGAT	
TGTTACAGGT	TTAAACGAGC	CAGGTCAGAC	TCAATTTGCG	GCAGTCCGGT	TTGGGAATGT
TCTAGGTAGT	CGTGGAAGTG	TTGTTCCGCT	ATTCAAAGAG	CAAATTAGAA	
AAGGTGGACC	TGTTACGGTT	ACCGACTTTA	GGATGACTCG	TTATTTTCATG	ACGATTCCTG
AGGCAAGTCG	TTTGTTTATC	CAAGCTGGAC	ATTTGGCAAA	AGGTGGAGAA	
ATATTTGTCT	TGGATATGGG	CGAGCCAGTA	CAAATCCTGG	AATTGGCAAG	AAAAGTTATC
TTGTTAAGTG	GACACACAGA	GGAAGAAATC	GGGATTGTAG	AATCTGGAAT	
CAGACCAGGC	GAGAACTCT	ACGAGGAATT	ATTATCAACA	GAAGAACGTG	TCAGCGAACA
GATTCATGAA	AAAATATTTG	TGGGTCGCGT	TACAAATAAG	CAGTCGGACA	
TTGTCAATTC	ATTTATCAAT	GGATTACTCC	AAAAAGATAG	AAATGAATTA	AAAAATATGT
TGATTGAATT	TGCAAAACAA	GAATAAGAAA	GTAAAAAATA	TTTTTACTTT	
CCTAGAGTTT	AAACGATGTT	TAAGTTCTAG	GAAGGTTAGA	ATACCTAATT	AACAACAATA
TTACTATTTA	TTAAGAGTCA	GATAATAGCA	ACTAAGTGCT	ACAAACTATC	
TTTATAATAA	GTATATTTGG	TCAAAAGGGA	GATGTGAAAT	GTATCCAATT	TGTAAACGTA
TTTTAGCAAT	TATTATCTCA	GGGATTGCTA	TTGTTGTTCT	GAGTCCAATT	
TTATTATTGA	TTGCATTGGC	AATTAAATTA	GATTCTAAAG	GTCCGGTATT	ATTTAAACAA
AAGCGGGTTG	GTAAAAACAA	GTCATACTTT	ATGATTTATA	AATTCGGTTC	
TATGTACGTT	GACGCACCAA	GTGATATGCC	GACTCATCTA	TTAAAGGATC	CTAAGGCGAT
GATTACCAAG	GTGGGCGCGT	TTCTCAGAAA	AACAAGTTTA	GATGAACTGC	
CACAGCTTTT	TAATATTTTT	AAAGGTGAAA	TGGCGATTGT	TGGTCCACGC	CCAGCCTTAT
GGAATCAATA	TGACTTAATT	GAAGAGCGAG	ATAAATATGG	TGCAAAATGAT	
ATTCGTCTCG	GACTAACCGG	TTGGGCTCAA	ATTAATGGTC	GTGATGAATT	GGAAATTGAT
GAAAAGTCAA	AATTAGATGG	ATATTATGTT	CAAAATATGA	GTCTAGGTTT	
GGATATTAAA	TGTTTCTTAG	GTACATTCCT	CAGTGTAGCC	AGAAGCGAAG	GTGTTGTTGA
AGGTGGAACA	GGGCAGAAAG	GAAAAGGATG	AAATTTTCAG	TATTAATGTC	
GGTCTATGAG	AAAGAAAAAC	CAGAGTTTCT	TAGGGAATCT	TTGGAAAGCA	TCCTTGTCAA
TCAAACAATG	ATTCCAACGG	AGGTTGTCTT	GGTAGAGGAT	GGGCCACTCA	
ATCAGAGCTT	ATATAGTATT	TTAGAAGAAT	TTAAAAGTCG	ATTTTCATTT	TTTAAACGA
TAGCCTTGGA	AAAGAATTCG	GGTTTAGGAA	TTGCACTGAA	TGAAGGTTTG	
AAACATTGTA	ATTATGAGTG	GGTTTGCACG	AAATGGATTC	TGATGATGTT	GCATATACAT
ACACGTTTTG	AAAAGCAAGT	TAACTTTATA	AAACAAAACC	CGACTATAGA	

DNA Serotype 9

WO 00/05378

PCT/NL99/00460

41/59

TATTGAGATA GATGAGTTCT TAAATTCTAC TAGTGAAATA GTTTCTCATA AAAATGTTCC
AACCCAGCAC GATGAAATAT TAAAGATGGC AAGGCGGGAG AAATCCATGT
GCCACATGAC TGTAATGTTT AAAAAGAAAA GTGTCGAGAG AGCAGGGGGG TATCAAACAC
TTCCGTACGT AGAAGATTAT TTCCTTTGGG TGCGCATGAT TGCTTCAGGA
TCGAAATTTG CAAACATTGA TGAAACACTA GTTCTTGCAC GTGTTGGAAA TGGGATGTTC
AATAGGAGGG GGAACAGAGA ACAAATTAAC AGTTGGACAT TACTAATTGA
ATTTATGTTA GCTCAAGGAA TTGTTACACC ACTAGATGTA TTTATTAATC AAATTTACAT
TAGGGTCTTT GTTTATATGC CAACTTGGAT AAAGAACTC ATTTATGGAA
AAATCTTAAG GAAATAGTAT GATTACAGTA TTGATGGCTA CATATAATGG AAGCCCATTT
ATAATAAAAC AGTTAGATTC AATTCGAAAT CAAAGTGTAT CAGCAGACAA
AGTTATTATT TGGGATGATT GCTCGACAGA TGATACAATA AAAATAATAA AAGATTATAT
AAAAAAATAT TCTTTGGATT CATGGGTGT CTCTCAAAT AAATCTAATC
AGGGGCATTA TCAAACATTT ATAAATTTGA CAAAGTTAGT TCAGGAAGGA ATAGTCTTTT
TTTCAGATCA AGATGATATT TGGGACTGTC ATAAAATTGA GACAATGCTT
CCAATCTTTG ACAGAGAAAA TGTATCAATG GTGTTTTGCA AATCCAGATT GATTGATGAA
AACGGAAATA TTATCAGTAG CCCAGATACT TCGGATAGAA TCAATACGTA
CTCTCTAGA

Fig. 5 cont.

WO 00/05378

PCT/NL99/00460

42/59

AYRQGVRYIV ATSHRRKGMF ETPEKVIMTN FLQFKDAVAE VYPEIRLCYG AELYYSKDIL
SKLEKKKVPT LNSRYILLE FSSDTPWKEI QEAVNEVTLL GLTPVLAHIE
RYDALAFHAE RVEELIDKGC YTQVNSNHVL KPTLIGDRAK EFKKRTRYFL EQDLVHCVAS
DMHNLSSRPP FMREAYKLLT EEF GKDKAKA LLKKNPLMLL KNQAI

Fig. 5 cont.

CPS9D

WO 00/05378

PCT/NL99/00460

43/59

MDLGTVTDKL	LERNKRLLIL	VCMDTCLLIV	SMILSRFLD	VIIDIPDERF	ILAVLFVSIL
YLILSFRLKV	FSLITRYTGY	QSYVKIGLSL	ISAHSLFLII	SMVLWQAFSY	
RFILVSLFLS	YVMLITPRIV	WKVLHETRKN	AIRKKDSPLR	ILVVGAGDGG	NIFINTVKDR
KLNFEIVGIV	DRDPNKLGTG	IRTAKVLGNR	NDIPRLVEEL	AVDQVTIAIP	
SLNGKEREKI	VEICNTTGVT	VNNMPSIEDI	MAGNMSVSAF	QEIDVADLLG	RPEVLDQDE
LNQFFQGKTI	LVTGAGGSIG	SELCRQIAKF	TPKRLLLLGH	GENSIYLIHR	
ELLEKYQGKI	ELVPLIADIQ	DRELIFSIMA	EYQPDVVYHA	AAHKHVPLME	YNPHEAVKNN
IFGTKNVAEA	AKTAKVAKFV	MVSTDKAVNP	PNVMGATKRV	AEMIVTGLNE	
PGQTQFAAVR	FGNVLGSRGS	VVPLFKEQIR	KGGPVTVTDF	RMTRYFMTIP	EASRLVIQAG
HLAKGGEIFV	LDMGEPVQIL	ELARKVILLS	GHTEEEIGIV	ESGIRPGEKL	
YEELLSTEER	VSEQIHEKIF	VGRVTNKQSD	IVNSFINGLL	QKDRNELKNM	LIEFAKQE

Fig. 5 cont.

CPS9E

WO 00/05378

44/59

PCT/NL99/00460

MYPICKRILA IIISGIAIVV LSPILLIAL AIKLDKGPV LFKQKRVGKN KSYFMIYKFR
SMYVDAPSDM PTHLLKDPKA MITKVGAFRL KTSLELPL FNIKFGEMAI
VGPRPALWNQ YDLIEERDKY GANDIRPGLT GWAQINGRDE LEIDEKSKLD GYYVQNMSLG
LDIKCFLGTF LSVARSEGVV EGGTGQKGKG

Fig. 5 cont.

CPS9F

WO 00/05378

PCT/NL99/00460

45/59

MKFSVLMSVY EKEKPEFLRE SLESILVNQT MIPTEVVLVE DGPLNQSLYS ILEEFKSRFS
FFKTIALEKN SGLGIALNEG LKHCNYEWVC TKWILMMLHI HTRFEKQVNF
IKQNPTIDIE IDEFLNSTSE IVSHKNVPTQ HDEILKMARR EKSMCHMTVM FKKKSVERAG
GYQTLPLYVED YFLWVRMIAS GSKFANIDET LVLARVGNGM FNRRGNREQI
NSWTLLIEFM LAQGIVTPLD VFINQIYIRV FVYMPTWIKK LIYGKILRK

Fig. 5 cont.

CPS9G

WO 00/05378

PCT/NL99/00460

46/59
MITVLMATYN GSPFIKQLD SIRQSVSAD KVIIWDDCST DDTIKIHKDY IKKYSLDSWV
VSQNKSNQGH YQTFINLTKL VQEGIVFFSD QDDIWDCHKI ETMLPIFDRE
NVSMVFCKSR LIDENGNIIS SPDTSDRINT YSL

Fig. 5 cont.

CPS9H

WO 00/05378

47/59

PCT/NL99/00460

CTGCAGCACA TAAGCATGTT CCATTGATGG AATATAATCC ACATGAAGCA GTGAAGAATA
 ATATTTTTTGG AACGAAGAAT GTGGCTGAGG CGGCTAAAAC TGCAAAGGTT
 GCCAAATTTG TTATGGTTTC AACAGATAAA GCTGTTAATC CGCCAAATGT CATGGGAGCG
 ACTAAACGTG TTGCAGAAAT GATTGTAACA GGTTTAAACG AGCCAGGTCA
 GACTCAATTT GCGGCAGTCC GTTTTGGGAA TGTTCAGGT AGTCGTGGAA GTGTTGTTCC
 GCTATTCAAA GAGCAAATTA GAAAAGGTGG ACCTGTTACG GTTACCGACT
 TTAGGATGAC TCGTTATTTT ATGACGATTC CTGAGGCAAG TCGTTTGGTT ATCCAAGCTG
 GACATTTGGC AAAAGGTGGA GAAATCTTTG TCTTGATAT GGGTGAGCCA
 GTACAAATCC TGGAAATTGGC AAGAAAAGTT ATCTTGTTAA GCGGACATAC AGAGGAAGAA
 ATCGGGATTG TAGAATCTGG AATCAGACCA GCGGAGAAAC TCTACGAGGA
 ATTGTTATCA ACAGAAGAAC GTGTCAGCGA ACAGATTCAT GAAAAAATAT TTGTGGGTCTG
 CGTTACAAAT AAGCAGTCGG ACATTGTCAA TTCATTTATC AATGGATTAC
 TCCAAAAGA TAGAAATGAA TTAAGAGATA TGTGATTGA ATTTGCAAAA CAAGAATAAG
 AAAGTAAAAA ATATTTTTTAC TTTCCTAGAG TTTAAACGAT GTTTAAGTTC
 TAGGAAGGTT GGAATTGCTT TCGTGGAGGT GATAGATAGA AACCTATATA TTTGTAGAAG
 AAAGGATATT AAATAAAGG TGAATCGGAA CATAAAGTTT AGATAGAGTT
 GGTATTTAAT GCCAACACAGG TGAATGCAAC CTCTCGCTCG TTACTAAGCA GGAGATAGTA
 AAGTTGCTTG AAAGAGAGTT TGTTAATCAG TATAAGTAGG CTAAAGTGAG
 AATATATATC TATTATTATC GGTAATGATA CTATTATTGA GAATTATTGT AGTGGGGATA
 AAAATAATTT TTGGTGATTT TATCGTCCGA CTAAAGGTG GGTAAAAAA
 GTACTTATAT TCTTTTAGAA TTGATGAAAA ATATGGGGGA ATATAATATT TATAGGAGAT
 ACGATGACTA GAGTAGAGTT GATTACTAGA GAATTTTTTA AGAAGAATGA
 AGCAACCAGT AAATATTTTC AGAAGATAGA ATCAAGAAGA GGTGAATTAT TTATTAAATT
 CTTTATGGAT AAGTTACTTG CGCTTATCCT ATTATTGCTA TTATCCCCAG
 TAATCATTAT ATTAGCTATT TGGATAAAAT TAGATAGTAA GGGGCCAATT TTTTATCGCC
 AAGAACGTGT TACGAGATAT GGTGCAATTT TTAGAATATT TAAGTTTAGA
 ACAATGATTT CTGATGCGGA TAAAGTCGGA AGTCTTGTCA CAGTCGGTCA AGATAATCGT
 ATTACGAAAG TCGGTCACAT TATCAGAAAA TATCGGCTGG ACGAAGTGCC
 CCAACTTTTT AATGTTTTTAA TGGGGGATAT GAGCTTTGTA GGTGTAAGAC CAGAAGTACA
 AAAATATGTA AATCAGTATA CTGATGAAAT GTTTGCGACG TTACTTTTAC
 CTGCAGGAAT TACTTCACCA GCGAGTATTG CATATAAGGA TGAAGATATT GTTTTAGAAG
 AATATTGTTT TCAAGGCTAT AGTCCTGATG AAGCATATGT TCAAAAAGTA
 TTACCAGAAA AAATGAAGTA CAATTTGGAA TATATCAGAA ACTTTGGAAT TATTTCTGAT
 TTTAAAGTAA TGATTGATAC AGTAATTAAA GTAATAAAAT AGGAGATTAA
 AATGACAAAA AGACAAAATA TTCCATTTTC ACCACCAGAT ATTACCCAAG CTGAAATTGA
 TGAAGTTATT GACACACTAA AATCTGGTTG GATTACAACA GGACCAAAGA
 CAAAAGAGCT AGAACGTCGG CTATCAGTAT TTACAGGAAC CAATAAACT GTGTGTTTAA
 ATTCTGCTAC TGCAGGATTG GAACTAGTCT TACGAATTCT TGGTGTTGGA
 CCCGGAGATG AAGTTATTGT TCCTGCTATG ACCTATACTG CCTCATGTAG TGTCATTACT
 CATGTAGGAG CAACTCCTGT GATGGTTGAT ATTCAAAAA ACAGCTTTGA
 GATGGAATAT GATGCTTTTG AAAAAGCGAT TACTCCGAAA ACAAAGTTA TCATTCCTGT
 TGATCTAGCT GGTATTCCTT GTGATTATGA TAAGATTTAT ACCATCGTAG
 AAAACAAACG CTCTTTGTAT GTTGCTTCTG ATAATAAATG GCAGAACTT TTTGGGCGAG
 TTATTATCCT ATCTGATAGT GCACACTCAC TAGGTGCTAG TTATAAGGGA
 AAACCAGCGG GTTCCCTAGC AGATTTTACC TCATTTTCTT TCCATGCAGT TAAGAATTTT
 ACAACTGCTG AAGGAGGTAG TGTGACATGG AGATCACATC CTGATTTGGA
 TGACGAAGAG ATGTATAAAG AGTTTCAGAT TTACTCTCTT CATGGTCAGA CAAAGGATGC
 ATTAGCTAAG ACACAATTAG GGTCATGGGA ATATGACATT GTTATTCCTG
 GTTACAAGTG TAATATGACA GATATTATGG CAGGTATCGG TCTTGTCGAA TTAGAACGTT
 ACCCATCTTT GTTGAATCGT CGCAGAGAAA TCATTGAGAA ATACAATGCT
 GGCTTTGAGG GGAATTCGAT TAAGCCGTTG GTACACCTGA CGGAAGATAA ACAATCGTCT
 ATGCACTTGT ATATCACGCA TCTACAAGGC TATACTTTAG AACACGAAA
 TGAAGTCATT CAAAAAATGG CTGAAGCAGG TATTGCGTGC AATGTTCACT ACAAACCATT
 ACCTCTTCTC ACAGCCTACA AGAATCTTGG TTTTGAAATG AAAGATTTTC
 CGAATGCCTA TCAGTATTTT GAAAATGAAG TTACTGTC TCTTCATACC AACTTGAGTG
 ATGAAGATGT GGAGTATGTG ATAGAAATGT TTTTAAAAAT TGTTAGTAGA
 GATTAGTTAT TTTGGAAGGA GATATGGTGG AAAGAGATAT GGTGGAAAGA GACACGTTGG
 TATCTATAAT AATGCCCTCG TGGAATACAG CTAAGTATAT ATCTGAATCA
 ATCCAGTCAG TGTTGGACCA AACACACCAA AATTGGGAAC TTATAATCGT TGATGATTGT
 TCTAATGACG AAATGAAAA AGTTGTTTCG CATTTCAAAG ATTCAAGAAT

DNA Serotype 7

Fig. 6

WO 00/05378

PCT/NL99/00460

48/59

AAAGTTTTTT AAAAATTCGA ATAATTTAGG GGCAGCTCTA ACACGAAATA AGGCACTAAG
AAAAGCTAGA GGTAGGTGGA TTGCGTTCTT GGATTCAGAT GATTTATGGC
ACCCGAGTAA GCTAGAAAAA CAGCTTGAAT TTATGAAAAA TAATGGATAT TCATTTACTT
ATCACAATTT TGAAAAGATT GATGAATCTA GTCAGTCTTT ACGTGTCTG
GTGTCAGGAC CAGCAATTGT GACTAGAAAA ATGATGTACA ATTACGGCTA TCCAGGGTGT
TTGACTTTCA TGTATGATGC AGACAAAATG GGTTTAATTC AGATAAAAGA
TATAAAGAAA AATAACGATT ATGCGATATT ACTTCAATTG TGTAAGAAGT ATGACTGTTA
TCTTTTAAAT GAAAGTTTAG CTTCGTATCG AATTAGAAAA AA

Fig. 6 cont.

WO 00/05378

49/59

PCT/NL99/00460

AAHKHVPLME YNPHEAVKNN IFGKNVAEA AKTAKVAKFV MVSTDKAVNP PNVMGATKRV
AEMIVTGLNE PGQTQFAAVR FGNVLGSRGS VVPLFKEQIR KGGPVTVTDF
RMTRYFMTIP EASRLVIQAG HLAGGGEIFV LDMGEPVQIL ELARKVILLS GHTEEEIGIV
ESGIRPGEKL YEELLSTEER VSEQIHEKIF VGRVTNKQSD IVNSFINGLL
OKDRNELKDM LIEFAKQE

Fig. 6 cont.

CPS7E

WO 00/05378

50/59

PCT/NL99/00460

MTRVELITRE FFKKNEATSK YFQKIESRRG ELFIKFFMDK LLALILLLLL SPVIIILAIW
IKLDSKGPIF YRQERVTRYG RIFRIFKFRT MISDADKVGS LVTVGQDNRI
TKVGHIIRKY RLDEVPQLFN VLMGDMSFVG VRPEVQKYVN QYTDEMFATL LLPAGITSPA
SIAYKDEDIV LEEYCSQGYS PDEAYVQKVL PEKMKNLEY IRNFGIISDF
KVMIDTVIKV IK

Fig. 6 cont.

CPS7F

WO 00/05378

51/59

PCT/NL99/00460

MTKRQNIPIFS PPDITQAEID EVIDTLKSGW ITTGPKTKEL ERRLSVFTGT NKTVCCLNSAT
AGLELVLRIL GVGPGDEVIV PAMTYTASCS VITHVGATPV MVDIQKNSFE
MEYDALEKAI TPCKTKVIIPV DLAGIPCDYD KIYTIVENKR SLYVASDNKW QKLFGRVIIL
SDSAHSLGAS YKGKPAGSLA DFTSFSFHAV KNFTTAEGGS VTWRSHPDLD
DEEMYKEFQI YSLHGQTKDA LAKTQLGSWE YDIVIPGYKC NMTDIMAGIG LVQLERYPSL
LNRRREIEK YNAGFEGTSI KPLVHLTEDK QSSMHLYITH LQGYTLEQRN
EVIQKMAEAG IACNVHYKPL PLLTAYKNLG FEMKDFPNAY QYFENEVTLP LHTNLSDEDV
EYVIEMFLKI VSRD

Fig. 6 cont.

CPS7G

WO 00/05378

PCT/NL99/00460

52/59

MVERDMVERD TLVSIIMPSW NTAKYISESI QSVLDQTHQN WELIIVDDCS NDETEKVVSH
FKDSRIKFFK NSNNLGAALT RNKALRKARG RWIAFLDSDD LWHPSKLEKQ
LEFMKNNGYS FTYHNFEEKID ESSQSLRVLV SGPAIVTRKM MYNYGYPGCL TEMYDADKMG
LIQIKDIKKN NDYAILLQLC KKYDCYLLNE SLASYRIRK

Fig. 6 cont.

CPS7H

				*					
Cps2J	MEKVSIIIVPI	FNTEKYLRER	LDSIISQSYT	NLEILLIDDG	SSDSSTDICL	EYAEQDGRIK		60	
Cps2K	MINISIIVPI	YNVEQYLSC	INSIVNQTYK	HIEILLVNDG	STDNSEEICL	AYAKKDSRIR		60	
				*					
Cps2J	LFRLPNGGV	SARNYGIKNS	TANYIMFVDS	DDIVDGNIVE	SLYTCLKEND	SDLGGGLLAT		120	
Cps2K	YFKKENGLS	DARNYGISRA	KGDYLA FIDS	DDFIHSEFIQ	RL_HEAIERE	NAL__VAVAG		117	

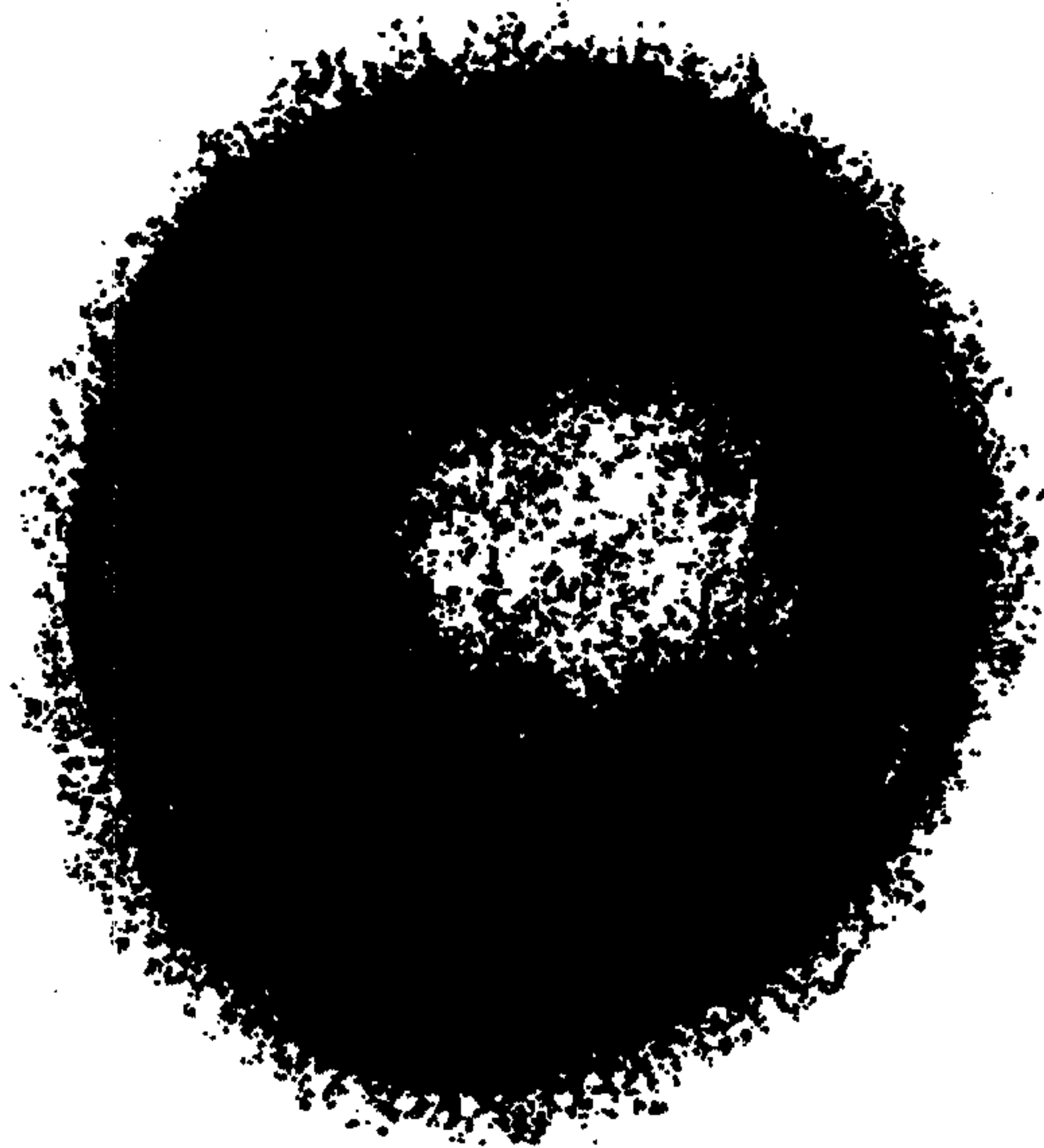
Fig. 7

54/59

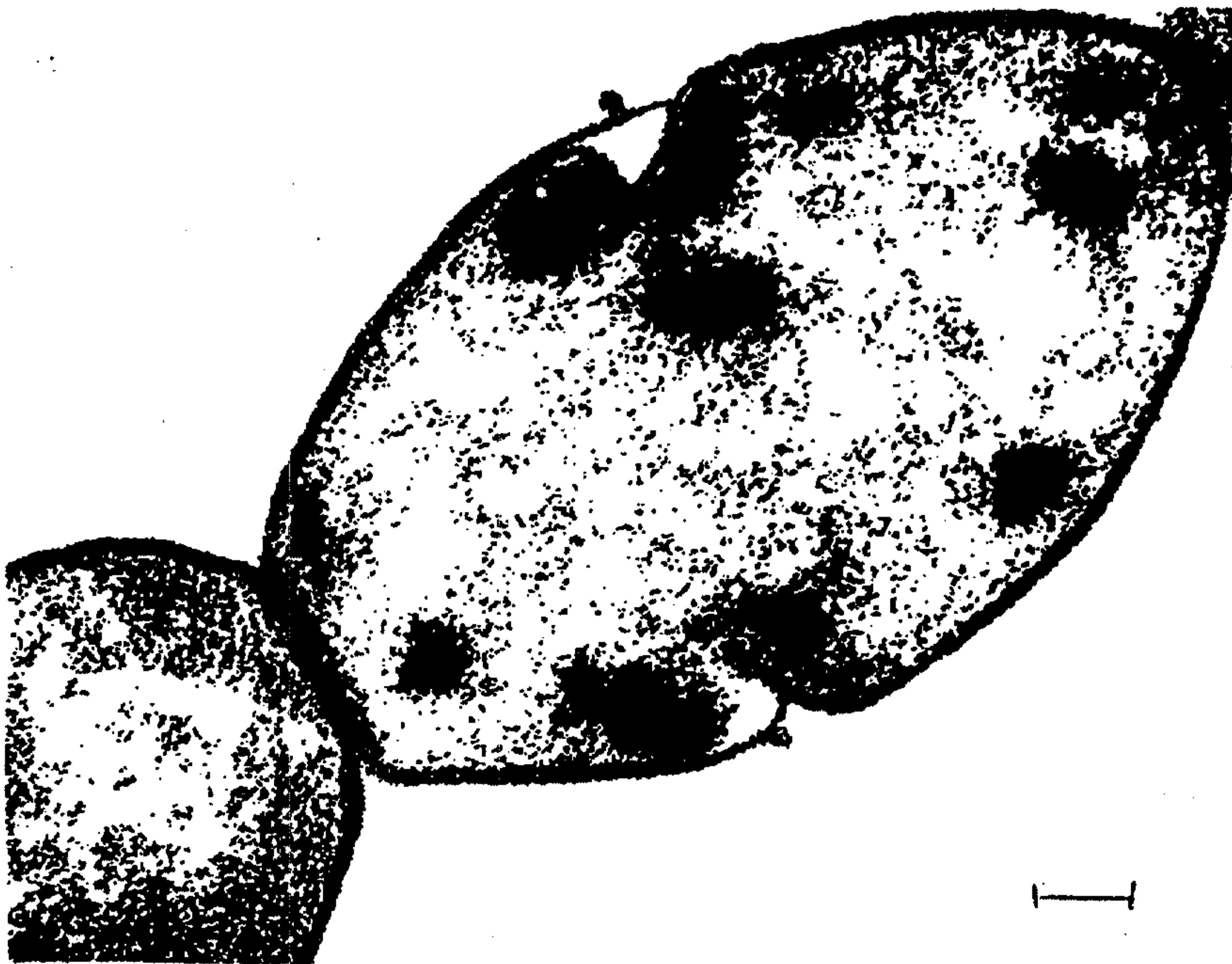
WO 00/05378

PCT/NL99/00460

A



B



C

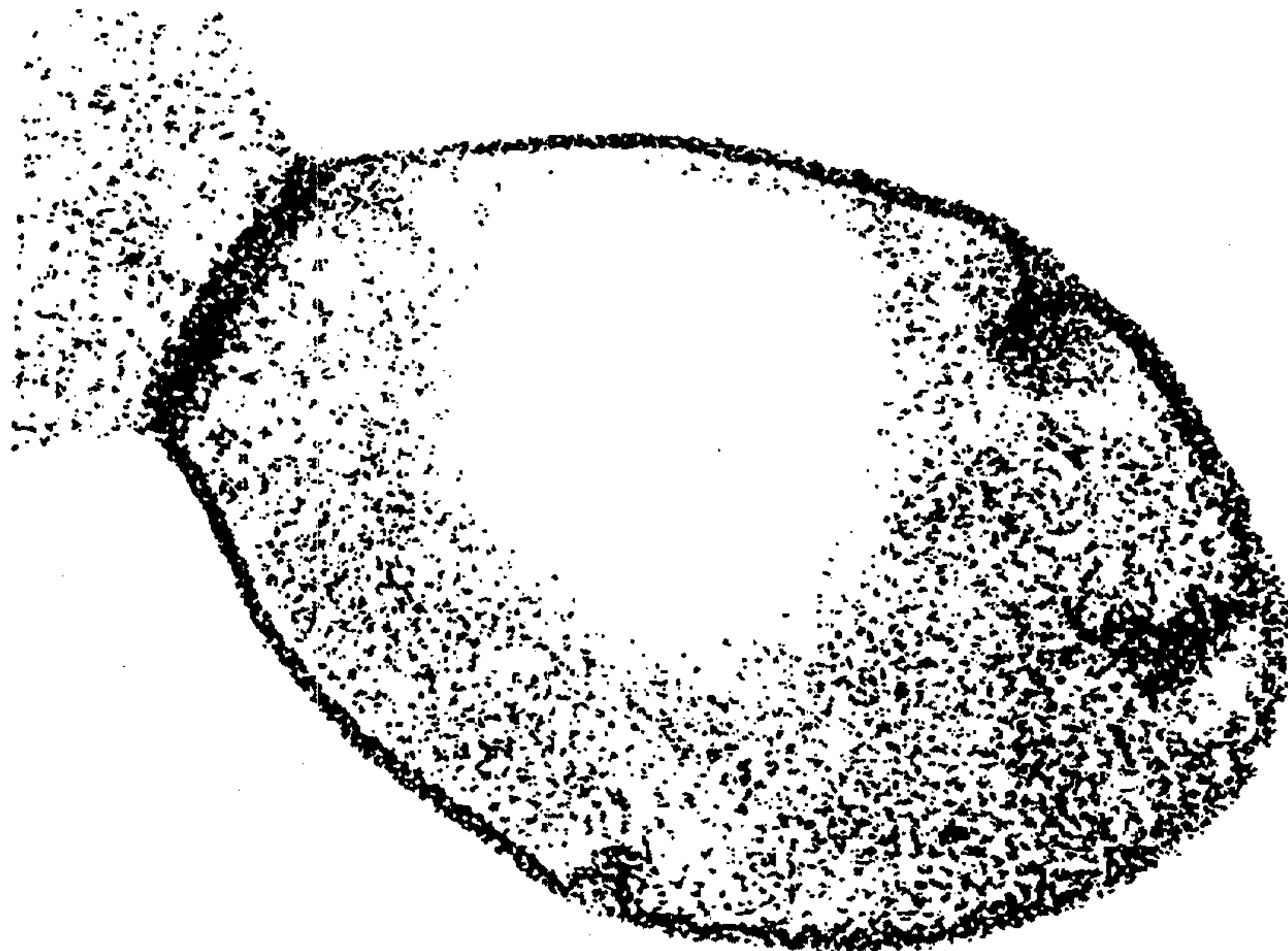


Fig. 8

55/59

WO 00/05378

PCT/NL99/00460

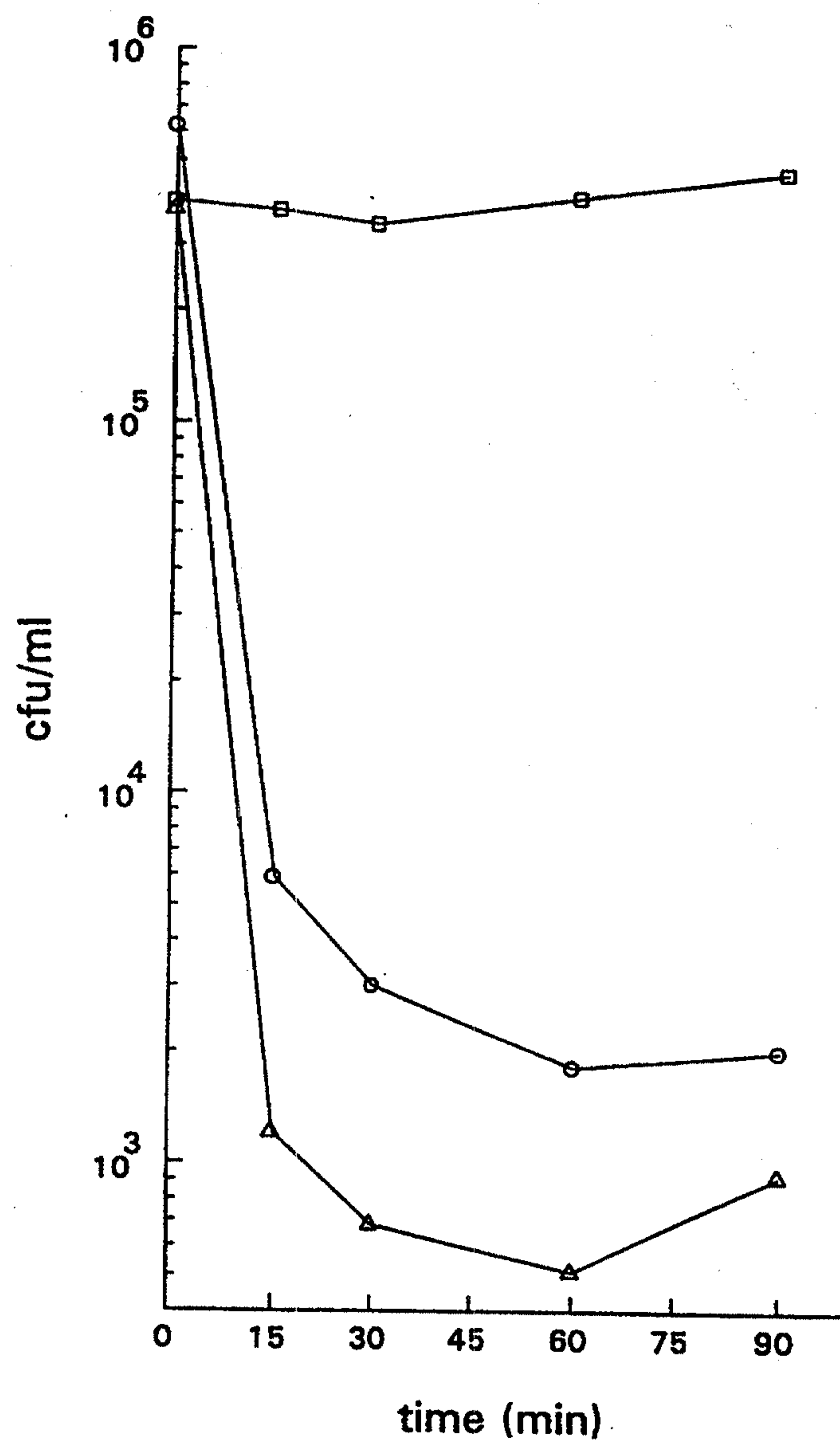


Fig. 9A

56/59

WO 00/05378

PCT/NL99/00460

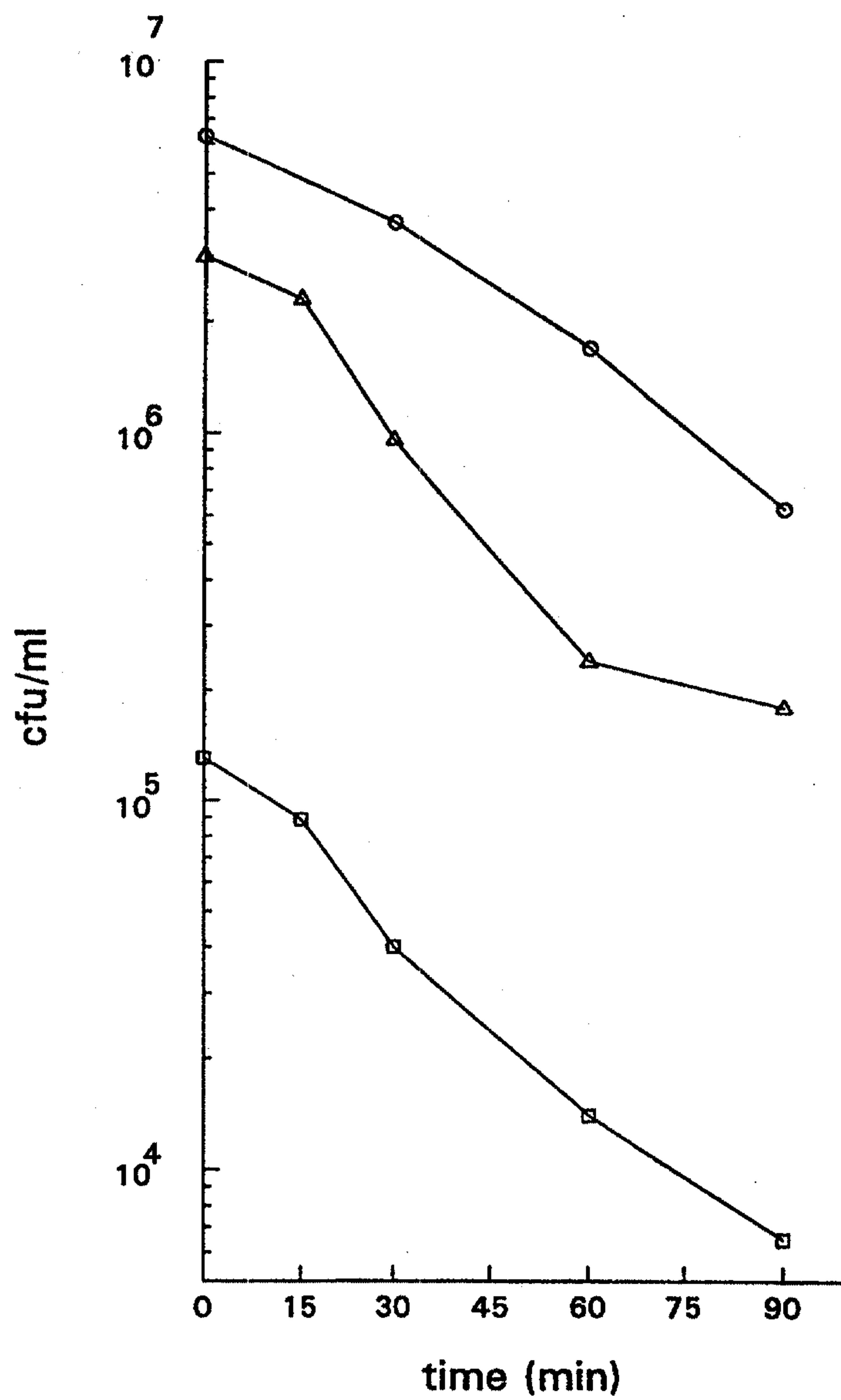


Fig. 9B

(1)	10508	AAGGGCACCT	CTATAAACTC	CCAAAAATGC	GAAATTGGAG	TTACGAAAGC	CTTGTTAAAT	CAA-CATTTTA	AATTTTAGAA	AATTAGTTTT	TAGAGCTCCC	10607
(2)	16985	GGGGCACCT	CTATAAACTC	CCAAAAATGC	GAAATTGGAG	TTACGAAAGC	CTTGTTAAAT	CAA-CATCTTA	AATTTTAGAA	AATTAGTTTT	TAGAGCTCCC	17084
(3)	19803	AAGGGCACCT	CTATAAACTC	CCAAAAATGC	GAAATTGGAG	TTACGAAAGC	CTTGTTAAAT	CAAACATTTTA	AATTTTAGAA	AATTAGTTTT	TAGAGCTCCC	19903

Fig. 10

WO 00/05378

58/59

PCT/NL99/00460

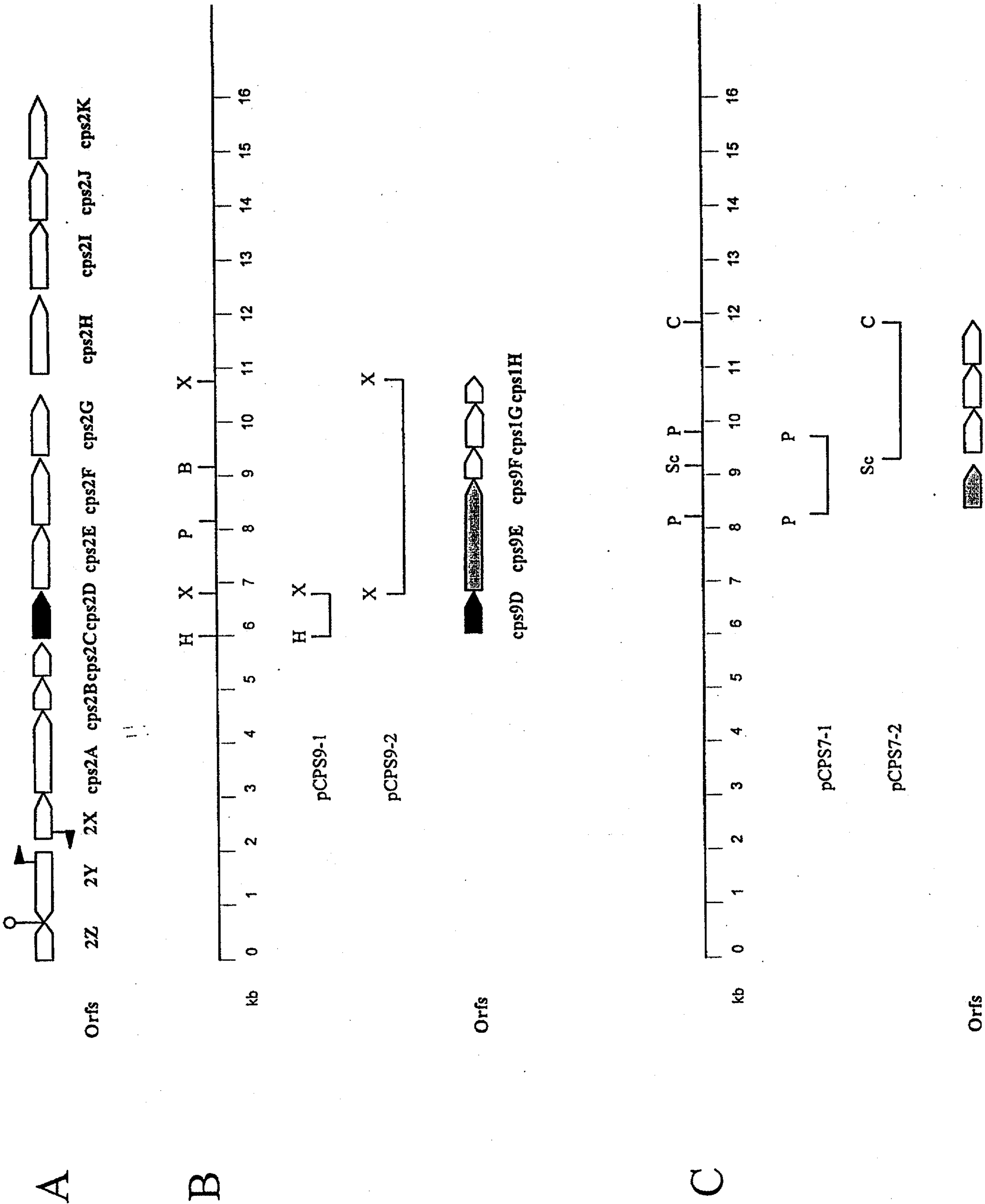


Fig. 11

WO 00/05378

PCT/NL99/00460

59/59

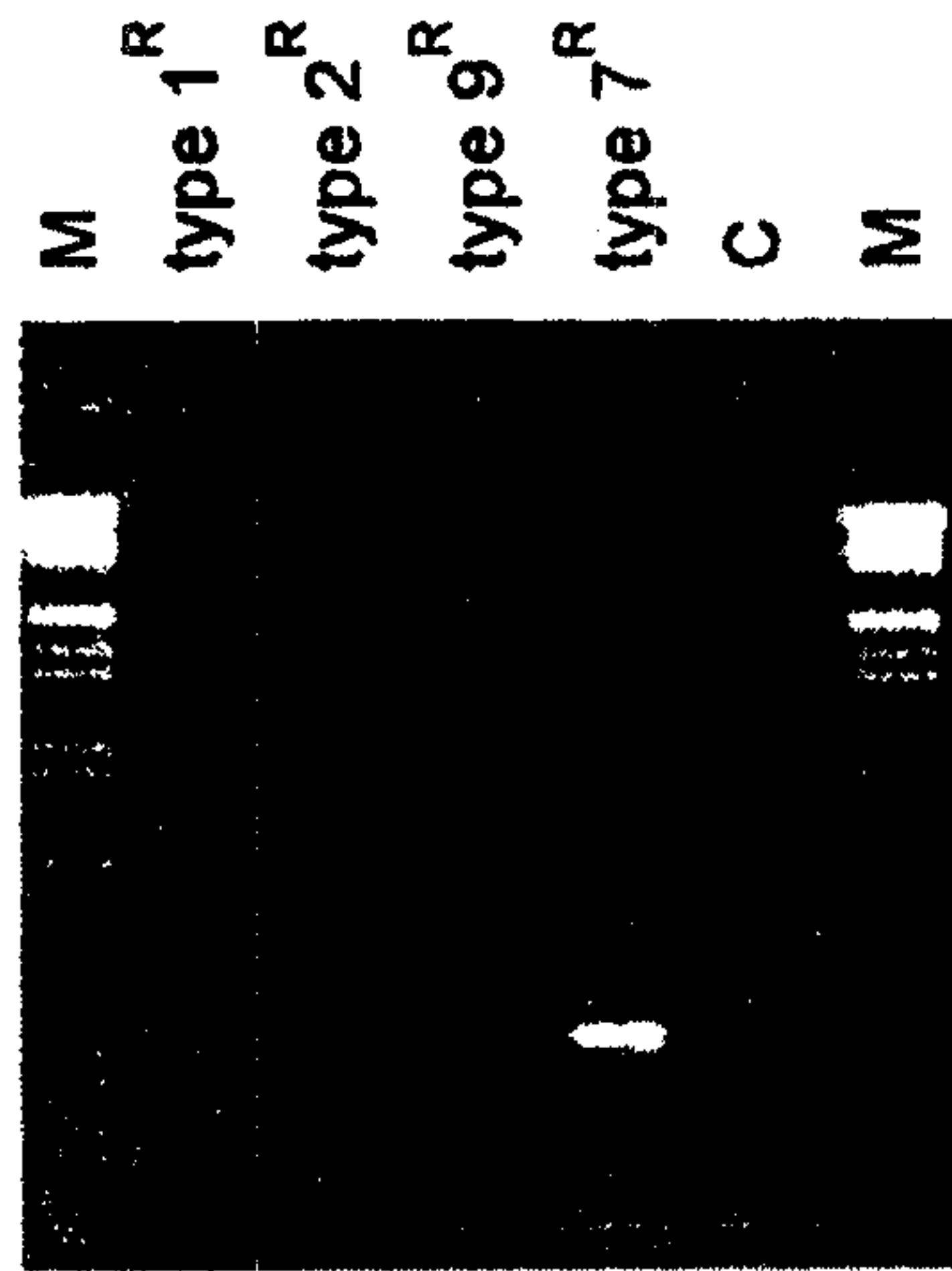
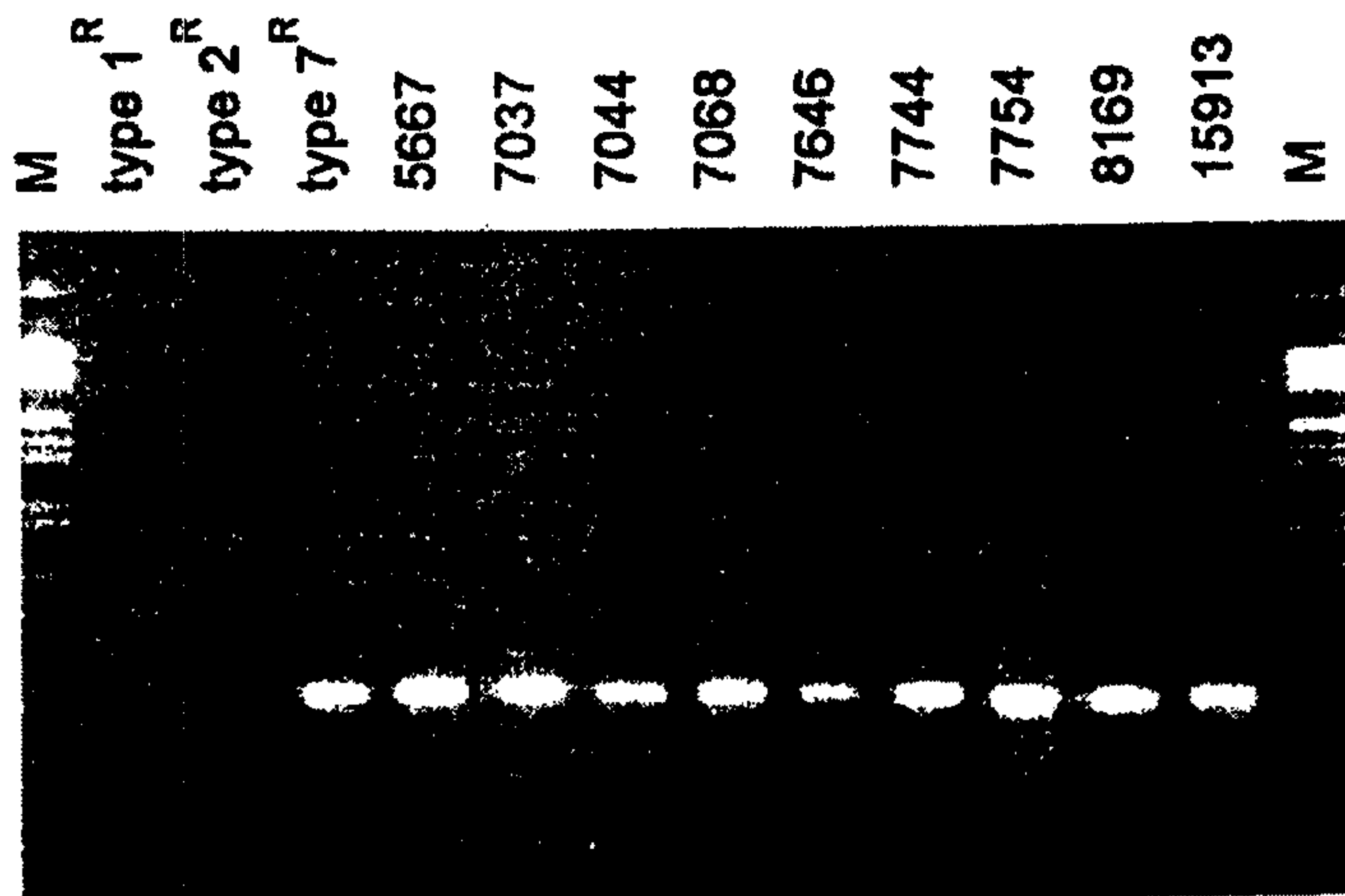
A**B**

Fig. 12