

# Description

## CORONAVIRUS

**FIELD OF THE INVENTION** The present invention relates to an attenuated coronavirus comprising a variant replicase gene, which causes the virus to have reduced pathogenicity. The present invention also relates to the use of such a coronavirus in a vaccine to prevent and/or treat a disease.

## BACKGROUND TO THE INVENTION

Avian infectious bronchitis virus (IBV), the aetiological agent of infectious bronchitis (IB), is a highly infectious and contagious pathogen of domestic fowl that replicates primarily in the respiratory tract but also in epithelial cells of the gut, kidney and oviduct. IBV is a member of the Order Nidovirales, Family Coronaviridae, Subfamily Coronavirinae and Genus Gammacoronavirus; genetically very similar coronaviruses cause disease in turkeys, guinea fowl and pheasants.

Clinical signs of IB include sneezing, tracheal rales, nasal discharge and wheezing. Meat-type birds have reduced weight gain, whilst egg-laying birds lay fewer eggs and produce poor quality eggs. The respiratory infection predisposes chickens to secondary bacterial infections which can be fatal in chicks. The virus can also cause permanent damage to the oviduct, especially in chicks, leading to reduced egg production and quality; and kidney, sometimes leading to kidney disease which can be fatal. IBV has been reported to be responsible for more economic loss to the poultry industry than any other infectious disease. Although live attenuated vaccines and

inactivated vaccines are universally used in the control of IBV, the protection gained by use of vaccination can be lost either due to vaccine breakdown or the introduction of a new IBV serotype that is not related to the vaccine used, posing a risk to the poultry industry.

Further, there is a need in the industry to develop vaccines which are suitable for use in ovo, in order to improve the efficiency and cost-effectiveness of vaccination programmes. A major challenge associated with in ovo vaccination is that the virus must be capable of replicating in the presence of maternally-derived antibodies against the virus, without being pathogenic to the embryo. Current IBV vaccines are derived following multiple passage in embryonated eggs, this results in viruses with reduced pathogenicity for chickens, so that they can be used as live attenuated vaccines. However such viruses almost always show an increased virulence to embryos and therefore cannot be used for in ovo vaccination as they cause reduced hatchability. A 70% reduction in hatchability is seen in some cases.

Attenuation following multiple passage in embryonated eggs also suffers from other disadvantages. It is an empirical method, as attenuation of the viruses is random and will differ every time the virus is passaged, so passage of the same virus through a different series of eggs for attenuation purposes will lead to a different set of mutations leading to attenuation. There are also efficacy problems associated with the process: some mutations will affect the replication of the virus and some of the mutations may make the virus too attenuated. Mutations can also occur in the S gene which may also affect immunogenicity so that the desired immune response is affected and the potential vaccine may not protect against the required serotype. In addition there are problems associated with reversion to virulence and

stability of vaccines. It is important that new and safer vaccines are developed for the control of IBV. Thus there is a need for IBV vaccines which are not associated with these issues, in particular vaccines which may be used for in ovo vaccination.

## SUMMARY OF ASPECTS OF THE INVENTION

The present inventors have used a reverse genetics approach in order to rationally attenuate IBV. This approach is much more controllable than random attenuation following multiple passages in embryonated eggs because the position of each mutation is known and its effect on the virus, i.e. the reason for attenuation, can be derived.

Using their reverse genetics approach, the present inventors have identified various mutations which cause the virus to have reduced levels of pathogenicity. The levels of pathogenicity may be reduced such that when the virus is administered to an embryonated egg, it is capable of replicating without being pathogenic to the embryo. Such viruses may be suitable for in ovo vaccination, which is a significant advantage and has improvement over attenuated IBV vaccines produced following multiple passage in embryonated eggs.

Thus in a first aspect, the present invention provides a live, attenuated coronavirus comprising a variant replicase gene encoding polyproteins comprising a mutation in one or more of non-structural protein(s) (nsp)-10, nsp-14, nsp-15 or nsp-16. The variant replicase gene may encode a protein comprising one or more amino acid mutations selected from the list of:

Pro to Leu at position 85 of SEQ ID NO: 6,

Val to Leu at position 393 of SEQ ID NO: 7;

Leu to lie at position 183 of SEQ ID NO: 8;

Val to lie at position 209 of SEQ ID NO: 9.

The replicase gene may encode a protein comprising the amino acid mutation Pro to Leu at position 85 of SEQ ID NO: 6.

The replicase gene may encode a protein comprising the amino acid mutations Val to Leu at position 393 of SEQ ID NO: 7; Leu to Ile at position 183 of SEQ ID NO: 8; and Val to lie at position 209 of SEQ ID NO: 9.

The replicase gene may encodes a protein comprising the amino acid mutations Pro to Leu at position 85 of SEQ ID NO: 6; Val to Leu at position 393 of SEQ ID NO:7; Leu to lie at position 183 of SEQ ID NO:8; and Val to lie at position 209 of SEQ ID NO: 9. The replicase gene may comprise one or more nucleotide substitutions selected from the list of:

C to T at nucleotide position 12137;

G to C at nucleotide position 18114;

T to A at nucleotide position 19047; and

G to A at nucleotide position 20139;

compared to the sequence shown as SEQ ID NO: 1.

The coronavirus may be an infectious bronchitis virus (IBV). The coronavirus may be IBV M41.

The coronavirus may comprise an S protein at least part of which is

from an IBV serotype other than M41. For example, the S1 subunit or the entire S protein may be from an IBV serotype other than M41. The coronavirus according to the first aspect of the invention has reduced pathogenicity compared to a coronavirus expressing a corresponding wild-type replicase, such that when the virus is administered to an embryonated egg, it is capable of replicating without being pathogenic to the embryo.

In a second aspect, the present invention provides a variant replicase gene as defined in connection with the first aspect of the invention. In a third aspect, the present invention provides a protein encoded by a variant coronavirus replicase gene according to the second aspect of the invention.

In a fourth aspect, the present invention provides a plasmid comprising a replicase gene according to the second aspect of the invention.

In a fifth aspect, the present invention provides a method for making the coronavirus according to the first aspect of the invention which comprises the following steps:

- (i) transfecting a plasmid according to the fourth aspect of the invention into a host cell;
- (ii) infecting the host cell with a recombining virus comprising the genome of a coronavirus strain with a replicase gene;
- (iii) allowing homologous recombination to occur between the replicase gene sequences in the plasmid and the corresponding sequences in the recombining virus genome to produce a modified replicase gene; and

(iv) selecting for recombining virus comprising the modified replicase gene.

The recombining virus may be a vaccinia virus.

The method may also include the step:

(v) recovering recombinant coronavirus comprising the modified replicase gene from the DNA from the recombining virus from step (iv).

In a sixth aspect, the present invention provides a cell capable of producing a coronavirus according to the first aspect of the invention.

In a seventh aspect, the present invention provides a vaccine comprising a coronavirus according to the first aspect of the invention and a pharmaceutically acceptable carrier. In an eighth aspect, the present invention provides a method for treating and/or preventing a disease in a subject which comprises the step of administering a vaccine according to the seventh aspect of the invention to the subject.

Further aspects of the invention provide:

- the vaccine according to the seventh aspect of the invention for use in treating and/or preventing a disease in a subject.
- use of a coronavirus according to the first aspect of the invention in the manufacture of a vaccine for treating and/or preventing a disease in a subject.

The disease may be infectious bronchitis (IB).

The method of administration of the vaccine may be selected from the group consisting of; eye drop administration, intranasal administration, drinking water administration, post- hatch injection and in ovo injection.

Vaccination may be by in ovo vaccination. The present invention also provides a method for producing a vaccine according to the seventh aspect of the invention, which comprises the step of infecting a cell according to the sixth aspect of the invention with a coronavirus according to the first aspect of the invention. DESCRIPTION OF THE FIGURES

Figure 1 - Growth kinetics of M41-R-6 and M41 -R-12 compared to M41-CK (M41 EP4) on CK cells Figure 2 - Clinical signs, snicking and wheezing, associated with M41 -R-6 and M41 -R-12 compared to M41 -CK (M41 EP4) and Beau-R (Bars show mock, Beau-R, M41-R 6, M41-R 12, M41-CK EP4 from left to right of each timepoint).

Figure 3 - Ciliary activity of the viruses in tracheal rings isolated from tracheas taken from infected chicks. 100% ciliary activity indicates no effect by the virus; apathogenic, 0% activity indicates complete loss of ciliary activity, complete ciliostasis, indicating the virus is pathogenic (Bars show mock, Beau-R, M41-R 6, M41 -R 12, M41 -CK EP4 from left to right of each timepoint). Figure 4 - Clinical signs, snicking, associated with M41 R-nsp10rep and M41 R- nsp14,15,16rep compared to M41 -R-12 and M41 -CK (M41 EP5) (Bars show mock, M41 - R12; M41 R-nsp10rep; M41 R-nsp14,15,16rep and M41 -CK EP5 from left to right of each timepoint).

Figure 5 - Ciliary activity of M41 R-nsp10rep and M41 R-nsp14,15,16rep compared to M41-R-12 and M41-CK in tracheal rings

isolated from tracheas taken from infected chicks (Bars show mock; M41-R12; M41 R-nsp10rep; M41 R-nsp14,15,16rep and M41 -CK EP5 from left to right of each timepoint).

Figure 6 - Clinical signs, snicking, associated with M41 R-nsp10, 15rep, M41 R-nsp10, 14, 15rep, M41 R-nsp10, 14, 16rep, M41 R-nsp10, 15, 16rep and M41 -K compared to M41 -CK (Bars show mock, M41 R-nsp10,15rep1 ; M41 R-nsp10,14,16rep4; M41 R-nsp10,15,16rep8; M41 R-nsp10,14,15rep10; M41 -K6 and M41 -CK EP4 from left to right of each timepoint).

Figure 7 - Clinical signs, wheezing, associated with M41 R-nsp10, 15rep, M41 R-nsp10, 14, 15rep, M41 R-nsp10, 14, 16rep, M41 R-nsp10, 15, 16rep and M41-K compared to M41-CK (Bars show mock, M41 R-nsp10,15rep1 ; M41 R-nsp10,14,16rep4; M41 R-nsp10,15,16rep8; M41 R-nsp10,14,15rep10; M41 -K6 and M41 -CK EP4 from left to right of each timepoint).

Figure 8 - Ciliary activity of M41 R-nsp10, 15rep, M41 R-nsp10, 14, 15rep, M41 R-nsp10,

14, 16rep, M41 R-nsp10, 15, 16rep and M41 -K compared to M41-CK in tracheal rings isolated from tracheas taken from infected chicks (Bars show mock, M41 R-nsp10,15rep1 ;

M41 R-nsp10,14,16rep4; M41 R-nsp10,15,16rep8; M41 R-nsp10,14,15rep10; M41-K6 and M41-CK EP4 from left to right of each timepoint).

Figure 9 - Growth kinetics of rIBVs compared to M41 -CK on CK cells. Fig 9A shows the results for M41 -R and M41 -K. Fig 9B shows the results for M41 -nsp10 rep; M41 R-nsp14,

15, 16 rep; M41 R-nsp10, 15 rep; M41 R-nsp10, 15, 16 rep; M41 R-nsp10, 14, 15 rep; and M41 R-nsp10, 14, 16.

Figure 10 - Position of amino acid mutations in mutated nsp10, nsp14, nsp15 and nsp16 sequences. Figure 11 - A) Snicking; B) Respiratory symptoms (wheezing and rales combined) and C) Ciliary activity of rIBV M41 R-nsp 10,14 rep and rIBV M41 R-nsp 10,16 rep compared to M41 -CK (Bars show mock, M41 R-nsp10,14rep; M41 R-nsp10,16rep and M41 -K from left to right of each timepoint).

## DETAILED DESCRIPTION

The present invention provides a coronavirus comprising a variant replicase gene which, when expressed in the coronavirus, causes the virus to have reduced pathogenicity compared to a corresponding coronavirus which comprises the wild-type replicase gene.

## CORONAVIRUS

Gammacoronavirus is a genus of animal virus belonging to the family Coronaviridae. Coronaviruses are enveloped viruses with a positive-sense single-stranded RNA genome and a helical symmetry.

The genomic size of coronaviruses ranges from approximately 27 to 32 kilobases, which is the longest size for any known RNA virus.

Coronaviruses primarily infect the upper respiratory or gastrointestinal tract of mammals and birds. Five to six different currently known strains of coronaviruses infect humans. The most publicized human coronavirus, SARS-CoV which causes severe acute respiratory syndrome (SARS), has a unique pathogenesis because it causes both upper and lower respiratory tract infections and can also cause

gastroenteritis. Middle East respiratory syndrome coronavirus (MERS-CoV) also causes a lower respiratory tract infection in humans. Coronaviruses are believed to cause a significant percentage of all common colds in human adults. Coronaviruses also cause a range of diseases in livestock animals and domesticated pets, some of which can be serious and are a threat to the farming industry. Economically significant coronaviruses of livestock animals include infectious bronchitis virus (IBV) which mainly causes respiratory disease in chickens and seriously affects the poultry industry worldwide; porcine coronavirus (transmissible gastroenteritis, TGE) and bovine coronavirus, which both result in diarrhoea in young animals. Feline coronavirus has two forms, feline enteric coronavirus is a pathogen of minor clinical significance, but spontaneous mutation of this virus can result in feline infectious peritonitis (FIP), a disease associated with high mortality. There are also two types of canine coronavirus (CCoV), one that causes mild gastrointestinal disease and one that has been found to cause respiratory disease. Mouse hepatitis virus (MHV) is a coronavirus that causes an epidemic murine illness with high mortality, especially among colonies of laboratory mice.

Coronaviruses are divided into four groups, as shown below:

#### Alpha

- Canine coronavirus (CCoV)
- Feline coronavirus (FeCoV)
- Human coronavirus 229E (HCoV-229E)
- Porcine epidemic diarrhoea virus (PEDV)

- Transmissible gastroenteritis virus (TGEV)
- Human Coronavirus NL63 (NL or New Haven)

## Beta

- Bovine coronavirus (BCoV)
- Canine respiratory coronavirus (CRCoV) - Common in SE Asia and Micronesia
- Human coronavirus OC43 (HCoV-OC43)
- Mouse hepatitis virus (MHV)
- Porcine haemagglutinating encephalomyelitis virus (HEV)
- Rat coronavirus (RCV). Rat Coronavirus is quite prevalent in Eastern Australia where, as of March/April 2008, it has been found among native and feral rodent colonies.
- (No common name as of yet) (HCoV-HKU1 )

## Severe acute respiratory syndrome coronavirus (SARS-CoV)

- Middle East respiratory syndrome coronavirus (MERS-CoV)

## Gamma

- Infectious bronchitis virus (IBV)
- Turkey coronavirus (Bluecomb disease virus)
- Pheasant coronavirus

- Guinea fowl coronavirus

Delta Bulbul coronavirus (BuCoV)

Thrush coronavirus (ThCoV)

Munia coronavirus (MuCoV)

Porcine coronavirus (PorCov) HKU15

The variant replicase gene of the coronavirus of the present invention may be derived from an alphacoronavirus such as TGEV; a betacoronavirus such as MHV; or a gammacoronavirus such as IBV.

As used herein the term "derived from" means that the replicase gene comprises substantially the same nucleotide sequence as the wild-type replicase gene of the relevant coronavirus. For example, the variant replicase gene of the present invention may have up to 80%, 85%, 90%, 95%, 98% or 99% identity with the wild type replicase sequence. The variant coronavirus replicase gene encodes a protein comprising a mutation in one or more of non-structural protein (nsp)-10, nsp-14, nsp-15 or nsp-16 when compared to the wild-type sequence of the non-structural protein.

IBV

Avian infectious bronchitis (IB) is an acute and highly contagious respiratory disease of chickens which causes significant economic losses. The disease is characterized by respiratory signs including gasping, coughing, sneezing, tracheal rales, and nasal discharge. In young chickens, severe respiratory distress may occur. In layers, respiratory distress, nephritis, decrease in egg production, and loss of

internal egg quality and egg shell quality are common.

In broilers, coughing and rattling are common clinical signs, rapidly spreading in all the birds of the premises. Morbidity is 100% in non-vaccinated flocks. Mortality varies depending on age, virus strain, and secondary infections but may be up to 60% in non-vaccinated flocks.

The first IBV serotype to be identified was Massachusetts, but in the United States several serotypes, including Arkansas and Delaware, are currently circulating, in addition to the originally identified Massachusetts type. The IBV strain Beaudette was derived following at least 150 passages in chick embryos. IBV Beaudette is no longer pathogenic for hatched chickens but rapidly kills embryos.

H120 is a commercial live attenuated IBV Massachusetts serotype vaccine strain, attenuated by approximately 120 passages in embryonated chicken eggs. H52 is another Massachusetts vaccine, and represents an earlier and slightly more pathogenic passage virus (passage 52) during the development of H120. Vaccines based on H120 are commonly used. IB QX is a virulent field isolate of IBV. It is sometimes known as "Chinese QX" as it was originally isolated following outbreaks of disease in the Qingdao region in China in the mid 1990s. Since that time the virus has crept towards Europe. From 2004, severe egg production issues have been identified with a very similar virus in parts of Western Europe, predominantly in the Netherlands, but also reported from Germany, France, Belgium, Denmark and in the UK.

The virus isolated from the Dutch cases was identified by the Dutch Research Institute at Deventer as a new strain that they called D388. The Chinese connection came from further tests which showed that

the virus was 99% similar to the Chinese QX viruses. A live attenuated QX-like IBV vaccine strain has now been developed.

IBV is an enveloped virus that replicates in the cell cytoplasm and contains a non-segmented, single-stranded, positive sense RNA genome. IBV has a 27.6 kb RNA genome and like all coronaviruses contains the four structural proteins; spike glycoprotein (S), small membrane protein (E), integral membrane protein (M) and nucleocapsid protein (N) which interacts with the genomic RNA.

The genome is organised in the following manner: 5'UTR - polymerase (replicase) gene - structural protein genes (S-E-M-N) - UTR 3'; where the UTR are untranslated regions (each ~ 500 nucleotides in IBV).

The lipid envelope contains three membrane proteins: S, M and E. The IBV S protein is a type I glycoprotein which oligomerizes in the endoplasmic reticulum and is assembled into homotrimer inserted in the virion membrane via the transmembrane domain and is associated through non-covalent interactions with the M protein. Following incorporation into coronavirus particles, the S protein is responsible for binding to the target cell receptor and fusion of the viral and cellular membranes. The S glycoprotein consists of four domains: a signal sequence that is cleaved during synthesis; the ectodomain, which is present on the outside of the virion particle; the transmembrane region responsible for anchoring the S protein into the lipid bilayer of the virion particle; and the cytoplasmic tail. All coronaviruses also encode a set of accessory protein genes of unknown function that are not required for replication in vitro, but may play a role in pathogenesis. IBV encodes two accessory genes, genes 3 and 5, which both express two accessory proteins 3a, 3b and 5a, 5b, respectively. The variant replicase gene of the coronavirus of the

present invention may be derived from an IBV. For example the IBV may be IBV Beaudette, H120, H52, IB QX, D388 or M41.

The IBV may be IBV M41. M41 is a prototypic Massachusetts serotype that was isolated in the USA in 1941. It is an isolate used in many labs throughout the world as a pathogenic lab stain and can be obtained from ATCC (VR-21™). Attenuated variants are also used by several vaccine producers as IBV vaccines against Massachusetts serotypes causing problems in the field. The present inventors chose to use this strain as they had worked for many years on this virus, and because the sequence of the complete virus genome is available. The M41 isolate, M41-CK, used by the present inventors was adapted to grow in primary chick kidney (CK) cells and was therefore deemed amenable for recovery as an infectious virus from a cDNA of the complete genome. It is representative of a pathogenic IBV and therefore can be analysed for mutations that cause either loss or reduction in pathogenicity.

The genome sequence of IBV M41 -CK is provided as SEQ ID NO: 1.

SEQ ID NO: 1 IBV M41 -CK Sequence

ACTTAAGATAGATATTA

ATATATATCTATCACACTAGCCTTGCGCTAGATTTCCAACCTTA

ACAAAACGGACTTAAATACCTACAGCTGGTCCTCATAGGTGTTCCATTG  
CAGTGCACTTT

AGTGCCCTGGATGGCACCTGGCCACCTGTCAGGTTTTTGTATTATAAAT  
CTTATTGTTGC

TGGTATCACTGCTTGTTTTGCCGTGTCTCACTTTATACATCCGTTGCTTG  
GGCTACCTAG

TATCCAGCGTCCTACGGGCGCCGTGGCTGGTTTCGAGTGCGAAGAACC

TCTGGTTCATCTA  
GCGGTAGGCGGGTGTGTGGAAGTAGCACTTCAGACGTACCGGTTCTGT  
TGTGTGAAATAC  
GGGGTCACCTCCCCCACATACCTCTAAGGGCTTTTGAGCCTAGCGTT  
GGGCTACGTTCT  
CGCATAAGGTTCGGCTATACGACGTTTGTAGGGGGTAGTGCCAAACAAC  
CCCTGAGGTGAC  
AGGTTCTGGTGGTGTTTAGTGAGCAGACATACAATAGACAGTGACAACA  
TGGCTTCAAGC  
CTAAAACAGGGAGTATCTCCCAAATAAGGGATGTCATTCTTGTATCCA  
AAGACATTCCT  
GAACAACCTTTGTGACGCTTTGTTTTTCTATACGTCACACAACCCTAAGG  
ATTACGCTGAT  
GCTTTTGCAGTTAGGCAGAAGTTTGATCGTAATCTGCAGACTGGGAAAC  
AGTTCAAATTT  
GAAACTGTGTGTGGTCTCTTCCTCTTGAAGGGAGTTGACAAAATAACAC  
CTGGCGTCCCA  
GCAAAAGTCTTAAAAGCCACTTCTAAGTTGGCAGATTTAGAAGACATCT  
TTGGTGTCTCT  
CCCTTTGCAAGAAAATATCGTGAACCTTTTGAAGACAGCATGCCAGTGGT  
CTCTTACTGTA  
GAAACACTGGATGCTCGTGCAAACTCTTGATGAAATTTTTGACCCTA  
CTGAAATACTT  
TGGCTTCAGGTGGCAGCAAAAATCCAAGTTTCGGCTATGGCGATGCGC  
AGGCTTGTTGGA  
GAAGTAACTGCAAAAGTCATGGATGCTTTGGGCTCAAATATGAGTGCTC  
TTTTCCAGATT  
TTTAAACAACAATAGTCAGAATTTTTCAAAAAGCGCTGGCTATTTTTGA  
GAATGTGAGT  
GAATTACCACAGCGTATTGCAGCACTTAAGATGGCTTTTGCTAAGTGTG

CCAAGTCCATT  
ACTGTTGTGGTTATGGAGAGGACTCTAGTTGTTAGAGAGTTTCGCAGGAA  
CTTGTCTTGCA  
AGCATTAAATGGTGCTGTTGCAAATTCTTTGAAGAACTCCCAAATGGTT  
TCATGGGTGCT  
AAAATTTTCACTACACTTGCCTTCTTTAGGGAGGCTGCAGTGAAAATTG  
TGGATAACATA  
CCAAATGCACCGAGAGGCACTAAAGGGTTTGAAGTCGTTGGTAATGCC  
AAAGGTACACAA  
GTTGTTGTGCGTGGCATGCGAAATGACTTAACACTGCTTGACCAAAA  
GCTGAAATTCCT  
GTGGAGTCAGAAGGTTGGTCTGCAATTTTGGGTGGACATCTTTGCTATG  
TCTTTAAGAGT  
GGTGATCGCTTTTACGCGGCACCTCTTTCAGGAAATTTTGCATTGCATG  
ATGTGCATTGT  
TGTGAGCGTGTTGTCTGTCTTTCTGATGGTGTAACACCGGAGATAAATG  
ATGGACTTATT  
CTTGCAGCAATCTACTCTTCTTTTAGTGTCGCAGAACTTGTGGCAGCCA  
TTAAAAGGGGT  
GAACCATTTAAGTTTCTGGGTCATAAATTTGTGTATGCAAAGGATGCAG  
CAGTTTCTTTT  
ACATTAGCGAAGGCTGCTACTATTGCAGATGTTTTGAAGCTGTTTCAAT  
CAGCGCGTGTG  
AAAGTAGAAGATGTTTGGTCTTCACTTACTGAAAAGTCTTTTGAATTCTG  
GAGGCTTGCA  
TATGGAAAAGTGCGTAATCTCGAAGAATTTGTTAAGACTTGTTTTTGTAA  
GGCTCAAATG  
GCGATTGTGATTTTAGCGACAGTGCTTGGAGAGGGCATTGTCATCTT  
GTTTCGCAAGTC  
ATCTATAAAGTAGGTGGTCTTTTTACTAAAGTTGTTGACTTTTGTGAAAA

ATATTGGAAA  
GGTTTTTGTGCACAGTTGAAAAGAGCTAAGCTCATTGTCACTGAAACCC  
TCTGTGTTTTG  
AAAGGAGTTGCACAGCATTGTTTTCAACTATTGCTGGATGCAATACAGT  
TTATGTATAAA  
AGTTTTAAGAAGTGTGCACTTGGTAGAATCCATGGAGACTTGCTCTTCT  
GGAAAGGAGGT  
GTGCACAAAATTATTCAAGAGGGCGATGAAATTTGGTTTGACGCCATTG  
ATAGTATTGAT  
GTTGAAGATCTGGGTGTTGTTCAAGAAAATTGATTGATTTTGATGTTT  
GTGATAATGTG  
ACACTTCCAGAGAACCAACCCGGTCATATGGTTCAAATCGAGGATGAC  
GGAAAGA ACTAC  
ATGTTCTTCCGCTTCAAAAAGGATGAGAACATTTATTATACACCAATGTC  
ACAGCTTGGT  
GCTATTAATGTGGTTTGCAAAGCAGGCGGTAAACTGTCACCTTTGGAG  
AAACTACTGTG  
CAAGAAATACCACCACCTGATGTTGTGTTTATTAAGGTTAGCATTGAGT  
GTTGTGGTGAA  
CCATGGAATACAATCTTCAAAAAGGCTTATAAGGAGCCCATTGAAGTAG  
AGACAGACCTC  
ACAGTTGAACAATTGCTCTCTGTGGTCTATGAGAAAATGTGTGATGATCT  
CAAGCTGTTT  
CCGGAGGCTCCAGAACCACCACCTTTGAGAATGTCACACTTGTTGATA  
AGAATGGTAAA  
GATTTGGATTGCATAAAATCATGCCATCTGATCTATCGTGATTATGAGAG  
CGATGATGAC  
ATCGAGGAAGAAGATGCAGAAGAATGTGACACGGATTCAGGTGATGCT  
GAGGAGTGTGAC  
ACTAATTCAGAATGTGAAGAAGAAGATGAGGATACTAAAGTGTTGGCTC

TTATAACAAGAC  
CCGGCAAGTAACAAATATCCTCTGCCTCTTGATGATGATTATAGCGTCTA  
CAATGGATGT  
ATTGTTCATAAGGACGCTCTCGATGTTGTGAATTTACCATCTGGTGAAG  
AAACCTTTGTT  
GTCAATAACTGCTTTGAAGGGGCTGTTAAAGCTCTTCCGCAGAAAGTTA  
TTGATGTTCTA  
GGTGA CTGGGGTGAGGCTGTTGATGCGCAAGAACAATTGTGTCAACAA  
GAATCAACTCGG  
GTCATATCTGAGAAATCAGTTGAGGGTTTTACTGGTAGTTGTGATGCAAT  
GGCTGAACAA  
GCTATTGTTGAAGAGCAGGAAATAGTACCTGTTGTTGAACAAAGTCAGG  
ATGTAGTTGTT  
TTTACACCTGCAGACCTAGAAGTTGTTAAAGAAACAGCAGAAGAGGTTG  
ATGAGTTTATT  
CTCATTTCTGCTGTCCCTAAAGAAGAAGTTGTGTCTCAGGAGAAAGAG  
GAGCCACAGGTT  
GAGCAAGAGCCTACCCTAGTTGTTAAAGCACAAACGTGAGAAGAAGGCT  
AAAAAGTTCAAA  
GTTAAACCAGCTACATGTGAAAAACCCAAATTTTTGGAGTACAAAACAT  
GTGTGGGTGAT  
TTGGCTGTTGTAATTGCCAAAGCATTGGATGAGTTTAAAGAGTTCTGCA  
TTGTAAACGCT  
GCAAATGAGCACATGTTCGCATGGTGGTGGCGTTGCAAAGGCAATTGCA  
GACTTTTGTGGA  
CCGGACTTTGTTGAATATTGCGCGGACTATGTTAAGAAACATGGTCCAC  
AGCAAAA ACTT  
GTCACACCTTCATTTGTTAAAGGCATTCAATGTGTGAATAATGTTGTAGG  
ACCTCGCCAT  
GGAGACAGCAACTTGCGTGAGAAGCTTGTTGCTGCTTACAAGAGTGTT

CTTG TAGGTGGA  
GTGGT TAACTATGTTGTGCCAGTTCTCTCATCAGGGATTTTTGGTGTAG  
ATTTTAAAATA  
TCAATAGATGCTATGCGCGAAGCTTTTAAAGGTTGTGCCATACGCGTTC  
TTTTATTTTCT  
CTGAGTCAAGAACACATCGATTATTTTCGATGCAACTTGTAAGCAGAAGA  
CAATTTATCTT  
ACGGAGGATGGTGT TAAATACCGCTCTGTTGTTTTAAACCTGGTGATT  
CTTTGGGTCAA  
TTTGGACAGGTTTTTGCAAGAAATAAGGTAGTCTTTTCGGCTGATGATG  
TTGAGGATAAA  
GAAATCCTCTTTATACCCACA ACTGACAAGACTATTCTTGAATATTATGG  
TTTAGATGCG  
CAAAGTATGTAACATATTTGCAAACGCTTGCGCAGAAATGGGATGTTC  
AATATAGAGAC  
AATTTTGTTATATTAGAGTGGCGTGACGGAAATTGCTGGATTAGTTCAG  
CAATAGTTCTC  
CTTCAAGCTGCTAAAATTAGATTTAAAGGTTTTCTTGCAGAAGCATGGG  
CTAAACTGTTG  
GGTGGAGATCCTACAGACTTTGTTGCCTGGTGTTATGCAAGTTGCAATG  
CTAAAGTAGGT  
GATTTTTTCAGATGCTAATTGGCTTTTGGCCAATTTAGCAGAACATTTTGA  
CGCAGATTAC  
ACAAATGCACTTCTTAAGAAGTGTGTGTCGTGCAATTGTGGTGTTAAGA  
GTTATGAACTT  
AGGGGTCTTGAAGCCTGTATTCAGCCAGTTTCGAGCACCTAATCTTCTAC  
ATTTTAAAACG  
CAATATTCAAATTGCCCAACCTGTGGTGCAAGTAGTACGGATGAAGTAA  
TAGAAGCTTCA  
TTACCGTACTTATTGCTTTTTTGCTACTGATGGTCCTGCTACAGTTGATTG

TGATGAAAAT  
GCTGTAGGGACTGTTGTTTTTCATTGGCTCTACTAATAGTGGCCATTGTT  
ATACACAAGCC  
GATGGTAAGGCTTTTGACAATCTTGCTAAGGATAGAAAATTTGGAAGGA  
AGTCGCCTTAC  
ATTACAGCAATGTATACACGTTTTTCTCTTAGGAGTGAAAATCCCCTACT  
TGTTGTTGAA  
CATAGTAAGGGTAAAGCTAAAGTAGTAAAAGAAGATGTTTCTAACCTTG  
CTACTAGTTCT  
AAAGCCAGTTTTTGACGATCTTACTGACTTTGAACAGTGGTATGATAGCA  
ACATCTATGAG  
AGTCTTAAAGTGCAGGAGACACCTGATAATCTTGATGAATATGTGTCATT  
TACGACAAAG  
GAAGATTCTAAGTTGCCACTGACACTTAAAGTTAGAGGTATCAAATCAG  
TTGTTGACTTT  
AGGTCTAAGGATGGTTTTACTTATAAGTTAACACCTGATACTGATGAAAA  
TTCAAAAACA  
CCAGTCTACTACCCAGTCTTGGATTCTATTAGTCTTAGGGCAATATGGGT  
TGAAGGCAGT  
GCTAATTTTGTTGTTGGGCATCCAAATTATTATAGTAAGTCTCTCCGAAT  
TCCCACGTTT  
TGGGAAAATGCCGAGAGCTTTGTTAAAATGGGTATATAAAATTGATGGTG  
TAACTATGGGC  
CTTTGGCGTGCAGAACACCTTAATAAACCTAATTTGGAGAGAATTTTAA  
ACATTGCTAAG  
AAAGCTATTGTTGGATCTAGTGTTGTTACTACGCAGTGTGGTAAAATACT  
AGTTAAAGCA  
GCTACATACGTTGCCGATAAAGTAGGTGATGGTG TAGTTCGCAATATTAC  
AGATAGAATT  
AAGGGTCTTTGTGGATTCACACGTGGCCATTTTGAAAAGAAAATGTCC

CTACAATTTCTA  
AAGACACTTGTGTTCTTTTTCTTTTATTTCTTAAAGGCTAGTGCTAAGAG  
TTTAGTTTCT  
AGCTATAAGATTGTGTTATGTAAGGTGGTGTGGTGGCTACCTTACTTATAGT  
GTGGTTTATA  
TACACAAGTAATCCAGTAGTGTTTACTGGAATACGTGTGCTAGACTTCCT  
ATTTGAAGGT  
TCTTTATGTGGTCCTTATAATGACTACGGTAAAGATTCTTTTGATGTGTT  
ACGCTATTGT  
GCAGGTGATTTTACTTGTCGTGTGTGTTTACATGATAGAGATTCACTTCA  
TCTGTACAAA  
CATGCTTATAGCGTAGAACAAATTTATAAGGATGCAGCTTCTGGCATTAA  
CTTTAATTGG  
AATTGGCTTTATTTGGTCTTTCTAATATTATTTGTTAAGCCAGTGGCAGG  
TTTTGTTATT  
ATTTGTTATTGTGTTAAGTATTTGGTATTGAGTTCAACTGTGTTGCAAAC  
TGGTGTAGGT  
TTTCTAGATTGGTTTGTA AAAACAGTTTTTACCCATTTTAATTTTATGGG  
AGCGGGATTT  
TATTTCTGGCTCTTTTACAAGATATACGTACAAGTGCATCATATATTGTAC  
TGTAAGGAT  
GTAACATGTGAAGTGTGCAAGAGAGTTGCACGCAGCAACAGGCAAGAG  
GTTAGCGTTGTA  
GTTGGTGGACGCAAGCAAATAGTGCATGTTTACACTAATTCTGGCTATA  
ACTTTTGTAAG  
AGACATAATTGGTATTGTAGAAATTGTGATGATTATGGTCACCAAATAC  
ATTTATGTCC  
CCTGAAGTTGCTGGCGAGCTTTCTGAAAAGCTTAAGCGCCATGTTAAA  
CCTACAGCATAT  
GCTTACCACGTTGTGTATGAGGCATGCGTGGTTGATGATTTTGTTAATT

TAAAATATAAG  
GCTGCAATTCCTGGTAAGGATAATGCATCTTCTGCTGTTAAGTGTTTCA  
GTGTTACAGAT  
TTTTTAAAGAAAGCTGTTTTTCTTAAGGAGGCATTGAAATGTGAACAAA  
TATCTAATGAT  
GGTTTTATAGTGTGTAATACACAGAGTGCGCATGCACTAGAGGAAGCAA  
AGAATGCAGCC  
GTCTATTATGCGCAATATCTGTGTAAGCCAATACTTATACTTGACCAGGC  
ACTTTATGAG  
CAATTAATAGTAGAGCCTGTGTCTAAGAGTGTTATAGATAAAGTGTGTAG  
CATTTTGTCT  
AATATAATATCTGTAGATACTGCAGCTTTAAATTATAAGGCAGGCACACTT  
CGTGATGCT  
CTGCTTTCTATTACTAAAGACGAAGAAGCCGTAGATATGGCTATCTTCTG  
CCACAATCAT  
GAAGTGGAATACACTGGTGACGGTTTTACTAATGTGATACCGTCATATG  
GTATGGACACT  
GATAAGTTGACACCTCGTGATAGAGGGTTTTTGATAAATGCAGATGCTT  
CTATTGCTAAT  
TTAAGAGTCAAAAATGCTCCTCCGGTAGTATGGAAGTTTTCTGATCTTAT  
TAAATTGTCT  
GACAGTTGCCTTAAATATTTAATTTTCAGCTACTGTCAAGTCAGGAGGTC  
GTTTCTTTATA  
ACAAAGTCTGGTGCTAAACAAGTTATTTCTTGTCATACCCAGAACTGT  
TGGTAGAGAAA  
AAGGCAGGTGGTGTTATTAATAACACTTTTAAATGGTTTATGAGTTGTTT  
TAAATGGCTT  
TTTGTCTTTTATATACTTTTTACAGCATGTTGTTTGGGTTACTACTATATG  
GAGATGAAT  
AAAAGTTTTGTTCCACCCCATGTATGATGTAAACTCCCACTGCATGTTG

AAGGGTTCAAA  
GTTATAGACAAAGGTGTTATTAGAGAGATTGTGTCAGAAGATAATTGTTT  
CTCTAATAAG  
TTTGTTAATTTTGACGCCTTTTGGGGTAAATCATATGAAAATAATAAAAA  
CTGTCCAATT  
GTTACAGTTGTTATAGATGGTGACGGGACAGTAGCTGTTGGTGTTCCTG  
GTTTTGTATCA  
TGGGTATGGATGGTGTATGTTTGTGCATATGACACAGACTGATCGTAG  
ACCTTGGTAC  
ATTCCTACCTGGTTTAATAGAGAAATTGTTGGTTACACTCAGGATTCAAT  
TATCACTGAG  
GGTAGTTTTTATACATCTATAGCATTATTTTCTGCTAGATGTTTATATTTAA  
CAGCCAGC  
AATACACCTCAATTGTATTGTTTTAATGGCGACAATGATGCACCTGGAG  
CCTTACCATTT  
GGTAGTATTATTCCTCATAGAGTATACTTCCAACCTAATGGTGTTAGGCT  
TATAGTTCCA  
CAACAAATACTGCATACACCCTACATAGTGAAGTTTGTTCAGACAGCTA  
TTGTAGAGGT  
AGTGTATGTGAGTATACTAAACCAGGTTACTGTGTGTCACTAGACTCCCA  
ATGGGTTTTG  
TTAATGATGAATACATTAGTAAACCTGGCGTTTTTCTGTGGTTCTACTGT  
TAGAGAACTT  
ATGTTTAATATGGTTAGTACATTCTTTACTGGTGTCAACCCTAATATTTAT  
ATTCAGCTA  
GCAACTATGTTTTTAATACTAGTTGTTATTGTGTTAATTTTTGCAATGGTT  
ATAAAGTTT  
CAAGGTGTTTTTAAAGCTTATGCGACCATTGTGTTTACAATAATGTTAGT  
TTGGGTTATT  
AATGCATTTGTTTTGTGTGTACATAGTTATAATAGTGTTTTTAGCTGTTATA

TTATTAGTA  
CTCTATTGCTATGCATCATTGGTTACAAGTCGCAATACTGCTATAATAATG  
CATTGTTGG  
CTTGTTTTTACCTTTGGTTTAATAGTACCCACATGGTTGGCTTGTTGCTA  
TCTGGGATTT  
ATTCTTTATATGTACACACCGTTGGTTTTCTGGTGTTACGGTACTACTAA  
AAATACTCGT  
AAGTTGTATGATGGCAACGAGTTTGTGGTAATTATGACCTTGCTGCGA  
AGAGCACTTTT  
GTTATTCGTGGTACTGAATTTGTTAAGCTTACGAATGAGATAGGTGATAA  
ATTTGAAGCC  
TATCTTTCTGCGTATGCTAGACTTAAATACTATTCAGGCACTGGTAGTGA  
GCAAGATTAC  
TTGCAAGCTTGTCGTGCATGGTTAGCTTATGCTTTGGACCAATATAGAA  
ATAGTGGTGTT  
GAGGTTGTTTATACCCACCGCGTACTCTATTGGTGTTAGTAGACTAC  
ACGCTGGTTTT  
AAAAAACTAGTTTCTCCTAGTAGTGCTGTTGAGAAGTGCATTGTTAGTG  
TCTCTTATAGA  
GGCAATAATCTTAATGGACTGTGGCTGGGTGATTCTATTTACTGCCCAC  
GCCATGTGTTA  
GGTAAGTTTAGTGGTGACCAGTGGGGTGACGTACTAAACCTTGCTAATA  
ATCATGAGTTT  
GAAGTTGTAACCTCAAATGGTGTTACTTTGAATGTTGTCAGCAGGCGGC  
TTAAAGGAGCA  
GTTTTAATTTTACAAACTGCAGTTGCCAATGCTGAAACTCCTAAGTATAA  
GTTTGTTAAA  
GCTAATTGTGGTGATAGTTTCACTATAGCTTGTTCTTATGGTGGTACAGT  
TATAGGACTT  
TACCCTGTCACCTATGCGTTCTAATGGTACTATTAGAGCATCTTTCCCTAGC

AGGAGCCTGT  
GGCTCAGTTGGTTTTAATATAGAAAAGGGTGTAGTTAATTTCTTTTATAT  
GCACCATCTT  
GAGTTACCTAATGCATTACACACTGGAAGTACCTAATGGGTGAGTTTT  
ATGGTGGTTAT  
GTAGATGAAGAGGTTGCGCAAAGAGTGCCACCAGATAATCTAGTTACTA  
ACAATATTGTA  
GCATGGCTCTATGCGGCAATTATTAGTGTTAAAGAAAGTAGTTTTTCACA  
ACCTAAATGG  
TTGGAGAGTACTACTGTTTCTATTGAAGATTACAATAGGTGGGCTAGTGA  
TAATGGTTTT  
ACTCCATTTTCCACTAGTACTGCTATTACTAAATTAAGTGCTATAACTGG  
GGTTGATGTT  
TGTAAGTCCCTTCGCACTATTATGGTAAAAAGTGCTCAATGGGGTAGTG  
ATCCCATTTTA  
GGACAATATAATTTTGAAGACGAATTGACACCAGAATCTGTATTTAATCA  
AGTTGGTGGT  
GTTAGGTTACAGTCTTCTTTTGTAAGAAAAGCTACATCTTGGTTTTGGA  
GTAGATGTGTA  
TTAGCTTGCTTCTTGTTTGTGTTGTGTGCTATTGTCTTATTTACGGCAGT  
GCCACTTAAG  
TTTTATGTACATGCAGCTGTTATTTTGTGATGGCTGTGCTCTTTATTTT  
TTTTACTGTT  
AAACATGTTATGGCATAACATGGACACTTTCCTATTGCCTACATTGATTAC  
AGTTATTATT  
GGAGTTTGTGCTGAAGTCCCTTTCATATAACAATACTCTAATTAGTCAAGT  
TGTTATTTTC  
TTAAGCCAATGGTATGATCCTGTAGTCTTTGATACTATGGTACCATGGAT  
GTTATTGCCA  
TTAGTGTTGTACTGCTTTTTAAGTGTGTACAAGGCTGCTATATGAATTC

TTTCAATACT  
TCTTTGTTAATGCTGTATCAGTTTATGAAGTTAGGTTTTGTTATTTACACC  
TCTTCAAAC  
ACTCTTACTGCATATACAGAAGGTAATTGGGAGTTATTCTTTGAGTTGGT  
TCACACTATT  
GTGTTGGCTAATGTTAGTAGTAATTCCTTAATTGGTTTAATTGTTTTTAA  
GTGTGCTAAG  
TGGATTTTATATTATTGCAATGCAACATACTTTAATAATTATGTGTTAATG  
GCAGTCATG  
GTTAATGGCATAGGCTGGCTTTGCACCTGTTACTTTGGATTGTATTGGT  
GGGTTAATAAA  
GTTTTTGGTTTAAACCTTAGGTAAATACAATTTTAAAGTTTCAGTAGATCA  
ATATAGGTAT  
ATGTGTTTGCATAAGGTAAATCCACCTAAAACCTGTGTGGGAGGTCTTTA  
CTACAAATATA  
CTTATACAAGGAATTGGAGGCGATCGTGTGTTGCCTATAGCTACAGTGC  
AATCTAAATTG  
AGTGATGTAAAGTGTACAACCTGTTGTTTTAATGCAGCTTTTGACTAAGCT  
TAATGTTGAA  
GCAAATTCAAAAATGCATGCTTATCTTGTTGAGTTACACAATAAAAATCCT  
CGCATCTGAT  
GATGTTGGAGAGTGCATGGATAATTTATTGGGTATGCTTATAACACTATT  
TTGTATAGAT  
TCTACTATTGATTTGGGTGAGTATTGTGATGATATACTTAAGAGGTCAAC  
TGTATTACAA  
TCGGT TACTCAAGAGTTTTTCGCACATACCCTCGTATGCTGAATATGAAA  
GAGCTAAGAGT  
ATTTATGAAAAGGTTTTAGCCGATTCTAAAAATGGTGGTGTAAACACAGC  
AAGAGCTTGCT  
GCATATCGTAAAGCTGCCAATATTGCAAAGTCAGTTTTTTGATAGAGACTT

GGCTGTTCAA  
AAGAAGTTAGATAGCATGGCAGAACGTGCTATGACAACAATGTATAAAG  
AGGCGCGTGTA  
ACTGATAGAAGAGCAAATTAGTTTCATCATTACATGCACTACTTTTTTC  
AATGCTTAAG  
AAAATAGATTCTGAGAAGCTTAATGTCTTATTTGACCAGGCGAATAGTG  
GTGTTGTACCC  
CTAGCAACTGTTCCAATTGTTTGTAGTAATAAGCTTACCCTTGTTATAACC  
AGACCCAGAG  
ACGTGGGTCAAGTGTGTGGAGGGTGTGCATGTTACATATTCAACAGTTG  
TTTGGAAATATA  
GACTGTGTTACTGATGCCGATGGCACAGAGTTACACCCCACTTCTACAG  
GTAGTGGATTG  
ACTTACTGTATAAGTGGTGATAATATAGCATGGCCTTTAAAGGTAACTT  
GACTAGGAAT  
GGGCATAATAAGGTTGATGTTGCCTTGCAAATAATGAGCTTATGCCTC  
ACGGTGTAAG  
ACAAAGGCTTGCGTAGCAGGTGTAGATCAAGCACATTGTAGCGTTGAGT  
CTAAATGTTAT  
TATACAAGTATTAGTGGCAGTTCAGTTGTAGCTGCTATTACCTCTTCAA  
TCCTAATCTG  
AAAGTAGCCTCTTTTTTTGAATGAGGCAGGTAATCAGATTTATGTAGACTT  
AGACCCACCA  
TGTAATTTGGTATGAAAGTGGGTGATAAGGTTGAAGTTGTTTACCTGT  
ATTTTATAAAA  
AATACGAGGTCTATTGTAAGAGGTATGGTACTTGGTGCTATATCTAATGT  
TGTTGTGTTA  
CAATCTAAAGGTCATGAGACAGAGGAAGTGGATGCTGTAGGCATTCTCT  
CACTTTGTTCT  
TTTGCAGTAGATCCTGCGGATACATATTGTAAATATGTGGCAGCAGGTAA

TCAACCTTTA  
GGTAACTGTGTTAAAATGTTGACAGTACATAATGGTAGTGGTTTTGCAAT  
AACATCAAAG  
CCAAGTCCAACCTCCGGATCAGGATTCTTATGGAGGAGCTTCTGTGTGTC  
TTTATTGTAGA  
GCACATATAGCACACCCTGGCGGAGCAGGAAATTTAGATGGACGCTGT  
CAATTTAAAGGT  
TCTTTTGTGCAAATACCTACTACGGAGAAAGATCCTGTTGGATTCTGTC  
TACGTAACAAG  
GTTTGCAGTGTGTTGTCAGTGTGGATTGGTTATGGATGTCAGTGTGATT  
CACTTAGACAA  
CCTAAACCTTCTGTTTCAGTCAGTTGCTGTTGCATCTGGTTTTGATAAGA  
ATTATTTAAAC  
GGGTACGGGGTAGCAGTGAGGCTCGGCTGATACCCCTAGCTAATGGAT  
GTGACCCCGATG  
TTGTAAAGCGAGCCTTTGATGTTTGTAAATAAGGAATCAGCCGGTATGTT  
TCAAATTTGA  
AGCGTAACTGTGCACGATTCCAAGAAGTACGTGATACTGAAGATGGAAA  
TCTTGAGTATT  
GTGATTCTTATTTTGTGGTTAAACAAACCACTCCTAGTAATTATGAACAT  
GAGAAAGCTT  
GTTATGAAGACTTAAAGTCAGAAGTAACAGCTGATCATGATTTCTTTGTG  
TTCAATAAGA  
ACATTTATAATATTAGTAGGCAGAGGCTTACTAAGTATACTATGATGGATT  
TTTGCTATG  
CTTTGCGGCACTTTGACCCAAAGGATTGCGAAGTTCTTAAAGAAATACT  
TGTCACCTTATG  
GTTGTATAGAAGATTATCACCCCTAAGTGGTTTGAAGAGAATAAGGATTGG  
TACGACCCAA  
TAGAAAACCCCTAAATATTATGCCATGTTGGCTAAAATGGGACCTATTGTA

CGACGTGCTT  
TATTGAATGCTATTGAGTTCGGAAACCTCATGGTTGAAAAAGGTTATGTT  
GGTGTATTATA  
CACTTGATAACCAAGATCTTAATGGCAAATTTTATGATTTTGGTGATTTT  
CAGAAGACAG  
CGCCTGGTGCTGGTGTTCCTGTTTTTGATACGTATTATTCTTACATGATG  
CCCATCATAG  
CCATGACTGATGCGTTGGCACCTGAGAGGTATTTTGAATATGATGTGCA  
TAAGGGTTATA  
AATCTTATGATCTCCTCAAGTATGATTATACTGAGGAGAAACAAGATTTG  
TTTCAGAAGT  
ACTTTAAGTATTGGGATCAAGAGTATCACCCCTAACTGTCGCGACTGTAG  
TGATGACAGGT  
GTTTGATACATTGTGCAAACCTTCAACATCTTGTTTTCTACACTTGTACCG  
CAGACTTCTT  
TCGGTAATTTGTGTAGAAAGGTTTTTGTGATGGTGTACCATTTATAGCT  
ACTTGTGGCT  
ATCATTCTAAGGAACTTGGTGTATTATGAATCAAGATAACACCATGTCA  
TTTTCAAAAA  
TGGGTTTGAGTCAACTCATGCAGTTTGTGGAGATCCTGCCTTGTTAGT  
GGGGACATCCA  
ATAAATTAGTGGATCTTAGAACGTCTTGTTTTAGTGTGTTGTGCTTTAGCG  
TCTGGTATTA  
CTCATCAAACGGTAAAACCAGGTCACCTTAAACAAGGATTTCTACGATTT  
TGCAGAGAAGG  
CTGGTATGTTTAAGGAAGGTTCTTCTATACCACTTAAACATTTCTTCTAC  
CCACAGACTG  
GTAATGCTGCTATAAACGATTATGATTATTATCGTTATAACAGGCCTACCA  
TGTTTGATA  
TACGTCAACTTTTATTTTGTGTTAGAAAGTGACTTCTAAATATTTTGAATGTT

ATGAAGGCG  
GCTGTATACCAGCAAGCCAAGTTGTAGTTAACAATTTAGATAAGAGTGC  
AGGTTATCCGT  
TCAATAAGTTTGGAAAGGCCCGTCTCTATTATGAAATGAGTCTAGAGGA  
GCAGGACCAAC  
TCTTTGAGAGTACAAAGAAGAACGTCCTGCCTACTATAACTCAGATGAA  
TTTAAAATATG  
CCATATCCGCGAAAAATAGAGCGCGTACAGTGGCAGGTGTGTCTATCCT  
TTCTACTATGA  
CTAATAGGCAGTTTCATCAGAAGATTCTTAAGTCTATAGTCAACACTAGA  
AACGCTCCTG  
TAGTTATTGGAACAACCAAGTTTTATGGCGGTTGGGATAACATGTTGAG  
AAACCTTATTC  
AGGGTGTTGAAGACCCGATTCTTATGGGTTGGGATTATCCAAAGTGTGA  
TAGAGCAATGC  
CTAATTTGTTGCGTATAGCAGCATCTTTAGTACTCGCTCGTAAACACACT  
AATTGTTGTA  
CTTGGTCTGAACGCGTTTATAGGTTGTATAATGAATGCGCTCAGGTTTT  
ATCTGAAACTG  
TCTTAGCTACAGGTGGTATATATGTGAAACCTGGTGGTACTAGCAGTGG  
AGATGCTACTA  
CTGCTTATGCAAACAGTGTTTTCAACATAATACAAGCCACATCTGCTAAT  
GTTGCGCGTC  
TTTTGAGTGTTATAACGCGTGATATTGTATATGATGACATTAAGAGCTTG  
CAGTATGAAT  
TGTACCAGCAGGTTTATAGGCGAGTCAATTTTGACCCAGCATTTGTTGA  
Ai<sup>3/4</sup>AGTTTTATT  
CTTATTTGTGTAAGAATTTCTCATTGATGATCTTGTCTGACGACGGTGTT  
GTTTGTATA  
ACAACACATTAGCCAAACAAGGTCTTGTAGCAGATATTTCTGGTTTTAGA

GAAGTTCTCT  
ACTATCAGAACAATGTTTTTATGGCTGATTCTAAATGTTGGGTTGAACCA  
GATTTAGAAA  
AAGGCCACATGAATTTTGTTCACAGCACACAATGTTAGTGGAGGTTGA  
TGGTGAGCCTA  
GATACTTGCCATATCCAGACCCATCACGTATTTTGTGTGCATGTGTTTTT  
GTAGATGATT  
TGGATAAGACAGAATCTGTGGCTGTTATGGAGCGTTATATCGCTCTTGC  
CATAGATGCGT  
ACCCACTAGTACATCATGAAAATGAGGAGTACAAGAAGGTATTCTTTGT  
GCTTCTTTCAT  
ACATCAGAAAACCTCTATCAAGAGCTTCTCAGAATATGCTTATGGACTAC  
TCTTTTGTA  
TGGATATAGATAAGGGTAGTAAATTTTGGGAACAGGAGTTCTATGAAAAT  
ATGTATAGAG  
CCCCTACAACATTACAGTCTTGTGGCGTTTGTGTAGTGTGTAATAGTCA  
AACTATATTGC  
GCTGTGGTAATTGTATTCGCAAACCATTTTTGTGTTGTAAGTGTTGCTAT  
GACCATGTCA  
TGCACACAGACCACAAAATGTTTTGTCTATAAATCCTTACATTTGCTCA  
CAGCCAGGTT  
GTGGTGAAGCAGATGTTACTAAATTGTACCTCGGAGGTATGTCATACTT  
CTGCGGTAATC  
ATAAACCAAAGTTATCAATACCGTTAGTATCTAATGGTACAGTGTTTGA  
ATTTACAGGG  
CTAATTGTGCAGGTAGCGAAAATGTTGATGATTTAATCAACTAGCTACT  
ACTAATTGGT  
CTACTGTGGAACCTTATATTTTGGCAAATCGTTGTGTAGATTCGTTGAGA  
CGCTTTGCTG  
CAGAGACAGTAAAAGCTACAGAAGAATTACATAAGCAACAATTTGCTAG

TGCAGAAGTGA  
GAGAAGTACTCTCAGATCGTGAATTGATTCTGTCTTGGGAGCCAGGTAA  
AACCAGGCCTC  
CATTGAATAGAAATTATGTTTTCACTGGCTTTCACTTTACTAGAACTAGT  
AAAGTTCAGC  
TCGGTGATTTTACATTTGAAAAAGGTGAAGGTAAGGACGTTGTCTATTA  
TCGAGCGACGT  
CTACTGCTAAATTGTCTGTTGGAGACATTTTTGTTTTAACCTCACACAAT  
GTTGTTTCTC  
TTATAGCGCCAACGTTGTGTCCTCAGCAAACCTTTTCTAGGTTTGTGAA  
TTTAAGACCTA  
ATGTGATGGTACCTGCGTGTTTTGTAATAACATTCCATTGTACCATTTA  
GTAGGCAAGC  
AGAAGCGTACTACAGTACAAGGCCCTCCTGGCAGTGGTAAATCCCATTT  
TGCTATAGGAT  
TGGCGGCTTACTTTAGTAACGCCCGTGTCGTTTTTACTGCATGCTCTCA  
TGCAGCTGTTG  
ATGCTTTATGTGAAAAAGCTTTTAAAGTTTCTTAAAGTAGATGATTGCACT  
CGTATAGTAC  
CTCAAAGGACTACTATCGATTGCTTCTCTAAGTTTAAAGCTAATGACACA  
GGCAAAAAGT  
ACATTTTTAGTACTATTAATGCCTTGCCAGAAGTTAGTTGTGACATTCTT  
TTGGTTGACG  
AGGTTAGTATGTTGACCAATTACGAATTGTCTTTTATTAATGGTAAGATAA  
ACTATCAAT  
ATGTTGTGTATGTAGGTGATCCTGCTCAATTACCGGCGCCTCGTACGTT  
GCTTAACGGTT  
CACTCTCTCAAAGGATTATAATGTTGTCACAAACCTTATGGTTTGTGTT  
AAACCTGACA  
TTTTCTTGCAAAGTGTTACCGTTGTCCTAAAGAAATTGTAGATACTGTT

TCTACTCTTG  
TATATGATGGAAAGTTTATTGCAAATAACCCGGAATCACGTCAGTGTTTC  
AAGGTTATAG  
TTAATAATGGTAATTCTGATGTAGGACATGAAAGTGGCTCAGCCTACAAC  
ATAACTCAAT  
TAGAATTTGTGAAAGATTTTGTCTGTCGCAATAAGGAATGGCGGGAAGC  
AACATTCATTT  
CACCTTATAATGCTATGAACCAGAGAGCCTACCGTATGCTTGGACTTAAT  
GTTTCAGACAG  
TAGACTCGTCTCAAGGTTCCGGAGTATGATTATGTTATCTTTTGTGTTACT  
GCAGATTCGC  
AGCATGCACTGAATATTAACAGATTCAATGTAGCGCTTACAAGAGCCAA  
GCGTGGTATAC  
TAGTTGTCATGCGTCAGCGTGATGAACTATATTCAGCTCTTAAGTTTATA  
GAGCTTGATA  
GTGTAGCAAGTCTGCAAGGTACAGGCTTGTTTAAAATTTGCAACAAAGA  
GTTTAGTGGTG  
TTCACCCAGCTTATGCAGTCACAACCTAAGGCTCTTGCTGCAACTTATAA  
AGTTAATGATG  
AACTTGCTGCACTTGTTAACGTGGAAGCTGGTTCAGAAATAACATATAA  
ACATCTTATTT  
CTTTGTTAGGGTTTAAAGATGAGTGTTAATGTTGAAGGCTGCCACAACAT  
GTTTATAACAC  
GTGATGAGGCTATCCGCAACGTAAGAGGTTGGGTAGGTTTTGATGTAGA  
AGCAACACATG  
CTTGCGGTACTAACATTGGTACTAACCTGCCTTTCCAAGTAGGTTTCTC  
TACTGGTGCAG  
ACTTTGTAGTTACGCCTGAGGGACTTGTAGATACTTCAATAGGCAATAAT  
TTTGAGCCTG  
TGAATTCTAAAGCACCTCCAGGTGAACAATTTAATCACTTGAGAGCGTT

ATTCAAAAGTG  
CTAAACCTTGGCATGTTGTAAGGCCAAGGATTGTGCAAATGTTAGCGGA  
TAACCTGTGCA  
ACGTTTCAGATTGTGTAGTGTTTGTACGTGGTGTGCATGGCCTAGA  
AACCACCTTTGC  
GCTATTTTGTAAATAGGCAAGGACCAAGTTTGTCTTGCGGTTCTAG  
AGCAACAACCTT  
TTAATTCTCATACTCAGGCTTATGCTTGTTGGAAGCATTGCTTGGGTTTT  
GATTTTGT  
ATAATCCACTCTTAGTGGATATTCAACAGTGGGGTTATTCTGGTAACCTA  
CAATTTAACC  
ATGATTTGCATTGTAATGTGCATGGACACGCACATGTAGCTTCTGCGGA  
TGCTATTATGA  
CGCGTTGTCTTGCAATTAATAATGCATTTTGTCAAGATGTCAACTGGGA  
TTTAACTTACC  
CTCATATAGCAAATGAGGATGAAGTCAATTCTAGCTGTAGATATTTACAA  
CGCATGTATC  
TTAATGCATGTGTTGATGCTCTTAAAGTTAACGTTGTCTATGATATAGGC  
AACCTAAAG  
GTATAZ^AATGTGTTAGACGTGGAGACTTAAATTTTAGATTCTATGATAAG  
AATCCAATAG  
TACCCAATGTCAAGCAGTTTGTAGTATGACTATAATCAGCACAAAGATAAG  
TTTGCTGATG  
GTCTTTGTATGTTTTGGAATTGTAATGTGGATTGTTATCCCGACAATTCC  
TTAGTTTGTA  
GGTACGACACACGAAATTTGAGTGTGTTAACCTACCTGGTTGTAATGG  
TGGTAGCTTGT  
ATGTTAACAAGCATGCATTCCACACACCTAAATTTGATCGCACTAGCTTT  
CGTAATTTGA  
AAGCTATGCCATTCTTTTTCTATGACTCATCGCCTTGCGAGACCATTCA

ATTGGATGGAG  
TTGCGCAAGACCTTGTGTCATTAGCTACGAAAGATTGTATCACAAAATG  
CAACATAGGCG  
GTGCTGTTTGTAAAAAGCACGCACAAATGTATGCAGATTTTGTGACTTC  
TTATAATGCAG  
CTGTTACTGCTGGTTTTACTTTTTGGGTTACTAATAATTTTAACCCATATA  
ATTTGTGGA  
AAAGTTTTTCAGCTCTCCAGTCTATCGACAATATTGCTTATAATATGTATA  
AGGGTGGTC  
ATTATGATGCTATTGCAGGAGAAATGCCCACTATCGTAACTGGAGATAAA  
GTTTTTGTTA  
TAGATCAAGGCGTAGAAAAAGCAGTTTTTTTTAATCAAACAATTCTGCC  
TACATCTGTAG  
CGTTTGAGCTGTATGCGAAGAGAAATATTCGCACACTGCCAAACAACC  
GTATTTTGAAAG  
GTTTGGGTGTAGATGTGACTAATGGATTTGTAATTTGGGATTACACGAA  
CCAAACACCAC  
TATACCGTAATACTGTTAAGGTATGTGCATATACAGACATAGAACCAAATG  
GCCTAATAG  
TGCTGTATGATGATAGATATGGTGATTACCAGTCTTTTCTAGCTGCTGAT  
AATGCTGTTT  
TAGTTTCTACACAGTGTTACAAGCGGTATTCGTATGTAGAAATACCGTCA  
AACCTGCTTG  
TTCAGAACGGTATTCCGTTAAAAGATGGAGCGAACCTGTATGTTTATAA  
GCGTGTTAATG  
GTGCGTTTGTACGCTACCTAACACATTAACACACAGGGTTCGCAGTTA  
TGAAACTTTTG  
AACCTCGTAGTGATGTTGAGCGTGATTTTCTCGACATGTCTGAGGAGAG  
TTTTGTAGAAA  
AGTATGGTAAAGAATTAGGTCTACAGCACATACTGTATGGTGAAGTTGAT

AAGCCCCAAT  
TAGGTGGTTTACACACTGTTATAGGTATGTGCAGACTTTTACGTGCGAAT  
AAGTTGAACG  
CAAAGTCTGTTACTAATTCTGATTCTGATGTCATGCAAAATTATTTTGTAT  
TGGCAGACA  
ATGGTTCCTACAAGCAAGTGTGTACTGTTGTGGATTTGCTGCTTGATGA  
TTTCTTAGAAC  
TTCTTAGGAACATACTGAAAGAGTATGGTACTAATAAGTCTAAAGTTGTA  
ACAGTGTCAA  
TTGATTACCATAGCATAAATTTTATGACTTGGTTTGAAGATGGCATTATTA  
AAACATGTT  
ATCCACAGCTTCAATCAGCATGGACGTGTGGTTATAATATGCCTGAACT  
TTATAAAGTTC  
AGAATTGTGTTATGGAACCTTGCAACATTCCTAATTATGGTGTGGAATA  
GCGTTGCCAA  
GTGGTATTATGATGAATGTGGCAAAGTATACACAACCTCTGTCAATACCTT  
TCGAAAACAA  
CAATGTGTGTACCGCATAATATGCGAGTAATGCATTTTGGAGCTGGAAG  
TGACAAAGGAG  
TGGCTCCAGGTAGTACTGTTCTTAAACAATGGCTCCCAGAAGGGACAC  
TCCTTGTCGATA  
ATGATATTGTAGACTATGTGTCTGATGCACATGTTTCTGTGCTTTCAGAT  
TGCAATAAAT  
ATAAGACAGAGCACAAGTTTGTGATCTTGATATCTGATATGTATACAGAC  
AATGATTCAA  
AAAGAAAGCATGAAGGCGTGATAGCCAATAATGGCAATGATGACGTTTT  
CATATATCTCT  
CAAGTTTTCTTCGTAATAATTTGGCTCTAGGTGGTAGTTTTGCTGTAAAA  
GTGACAGAGA  
CAAGTTGGCACGAAGTTTTATATGACATTGCACAGGATTGTGCATGGTG

GACAATGTTTT  
GTACAGCAGTGAATGCCTCTTCTTCAGAAGCATTCTTGGTTGGTGTTAA  
TTATTTGGGTG  
CAAGTGAAAAGGTTAAGGTTAGTGGAAAAACGCTGCACGCAAATTATAT  
ATTTTGGAGGA  
ATTGTAATTATTTACAAACCTCTGCTTATAGTATATTTGACGTTGCTAAGT  
TTGATTTGA  
GATTGAAAGCAACACCAGTTGTTAATTTGAAAACCTGAACAAAAGACAGA  
CTTAGTCTTTA  
ATTTAATTAAGTGTGGTAAGTTACTGGTAAGAGATGTTGGTAACACCTCT  
TTTACTAGTG  
ACTCTTTTGTGTGTACTATGTAGTGCTGCTTTGTATGACAGTAGTTCTTA  
CGTTTACTAC  
TACCAAAGTGCCTTTAGACCACCTAATGGTTGGCATTACACGGGGGT  
GCTTATGCGGTA  
GTTAATATTTCTAGCGAATCTAATAATGCAGGCTCTTCACCTGGGTGTAT  
TGTTGGTACT  
ATTCATGGTGGTCGTGTTGTTAATGCTTCTTCTATAGCTATGACGGCACC  
GTCATCAGGT  
ATGGCTTGGTCTAGCAGTCAGTTTTGTACTGCACACTGTAACCTTTTCAG  
ATACTACAGTG  
TTTGTTACACATTGTTATAAATATGATGGGTGTCCTATAACTGGCATGCT  
TCAAAAGAAT  
TTTTTACGTGTTTCTGCTATGAAAAATGGCCAGCTTTTCTATAATTTAAC  
AGTTAGTGTA  
GCTAAGTACCCTACTTTTAAATCATTTTCAGTGTGTTAATAATTTAACATCC  
GTATATTTA  
AATGGTGATCTTGTTTACACCTCTAATGAGACCACAGATGTTACATCTGC  
AGGTGTTTAT  
TTTAAAGCTGGTGGACCTATAACTTATAAAGTTATGAGAGAAGTTAAAGC

CCTGGCTTAT  
TTTGTTAATGGTACTGCACAAGATGTTATTTTGTGTGATGGATCACCTAG  
AGGCTTGTTA  
GCATGCCAGTATAATACTGGCAATTTTTCAGATGGCTTTTATCCTTTTAT  
TAATAGTAGT  
TTAGTTAAGCAGAAGTTTATTGTCTATCGTGAAAATAGTGTTAATACTACT  
TTTACGTTA  
CACAATTTCACTTTTCATAATGAGACTGGCGCCAACCCTAATCCTAGTG  
GTGTTTCAGAAT  
ATTCAAACCTACCAAACACAAACAGCTCAGAGTGGTTATTATAATTTTAA  
TTTTTCCTTT  
CTGAGTAGTTTTGTTTATAAGGAGTCTAATTTTATGTATGGATCCTTATCAC  
CCAAGTTGT  
AATTTTAGACTAGAACTATTAATAATGGCTTGTGGTTTAATTCACCTTCA  
GTTTCAATT  
GCTTACGGTCCTCTTCAAGGTGGTTGCAAGCAATCTGTCTTTAGTGGTA  
GAGCAACTTGT  
TGTTATGCTTATTCATATGGAGGTCCTTCGCTGTGTAAAGGTGTTTATTC  
AGGTGAGTTA  
GATCTTAATTTTGAATGTGGACTGTTAGTTTATGTTACTAAGAGCGGTGG  
CTCTCGTATA  
CAAACAGCCACTGAACCGCCAGTTATAACTCGACACAATTATAATAATAT  
TACTTTAAAT  
ACTTGTGTTGATTATAATATATATGGCAGAACTGGCCAAGGTTTTATTACT  
AATGTAACC  
GACTCAGCTGTTAGTTATAATTATCTAGCAGACGCAGGTTTGGCTATTTT  
AGATACATCT  
GGTTCCATAGACATCTTTGTTGTACAAGGTGAATATGGTCTTACTTATTA  
TAAGGTTAAC  
CCTTGCGAAGATGTCAACCAGCAGTTTGTAGTTTCTGGTGGTAAATTAG

TAGGTATTCTT  
ACTTCACGTAATGAGACTGGTTCTCAGCTTCTTGAGAACCAGTTTTACA  
TTAAAATCACT  
AATGGAACACGTCGTTTTAGACGTTCTATTACTGAAAATGTTGCAAATT  
GCCCTTATGTT  
AGTTATGGTAAGTTTTGTATAAAACCTGATGGTTCAATTGCCACAATAGT  
ACCAAACAA  
TTGGAACAGTTTTGTGGCACCTTTACTTAATGTTACTGAAAATGTGCTCAT  
ACCTAACAGT  
TTTAATTTAACTGTTACAGATGAGTACATACAAACGCGTATGGATAAGGT  
CCAAATTAAT  
TGTCTGCAGTATGTTTGTGGCAATTCTCTGGATTGTAGAGATTTGTTTCA  
ACAATATGGG  
CCTGTTTGTGACAACATATTGTCTGTAGTAAATAGTATTGGTCAAAAAGA  
AGATATGGAA  
CTTTTGAATTTCTATTCTTCTACTAAACCGGCTGGTTTTAATACACCATT  
TCTTAGTAAT  
GTTAGCACTGGTGAGTTTAATATTTCTCTTCTGTTAACAACCTCCTAGTAG  
TCCTAGAAGG  
CGTTCTTTTATTGAAGACCTTCTATTTACAAGCGTTGAATCTGTTGGATT  
ACCAACAGAT  
GACGCATACAAAATTGCACTGCAGGACCTTTAGGTTTTCTTAAGGACC  
TTGCGTGTGCT  
CGTGAATATAATGGTTTGCTTGTGTTGCCTCCCATTATAACAGCAGAAAT  
GCAAATTTTG  
TATACTAGTTCTCTAGTAGCTTCTATGGCTTTTGGTGGTATTACTGCAGC  
TGGTGCTATA  
CCTTTTGCCACACAACCTGCAGGCTAGAATTAATCACTTGGGTATTACCC  
AGTCACTTTTG  
TTGAAGAATCAAGAAAAAATTGCTGCTTCCTTTAATAAGGCCATTGGTC

GTATGCAGGAA  
GGTTTTAGAAGTACATCTCTAGCATTACAACAATTCAAGATGTTGTTAA  
TAAGCAGAGT  
GCTATTCTTACTGAGACTATGGCATCACTTAATAAAAATTTTGGTGCTAT  
TTCTTCTATG  
ATTCAAGAAATCTACCAGCAACTTGACGCCATACAAGCAAATGCTCAAG  
TGGATCGTCTT  
ATAACTGGTAGATTGTCATCACTTTCTGTTTTAGCATCTGCTAAGCAGGC  
GGAGCATATT  
AGAGTGTCACAACAGCGTGAGTTAGCTACTCAGAAAATTAATGAGTGTG  
TTAAGTCACAG  
TCTATTAGGTACTCCTTTTGTGGTAATGGACGACATGTTCTAACCATACC  
GCAAATGCA  
CCTAATGGTATAGTGTTTATACACTTTTCTTATACTCCAGATAGTTTTGTT  
AATGTTACT  
GCAATAGTGGGTTTTTGTGTAAAGCCAGCTAATGCTAGTCAGTATGCAA  
TAGTACCCGCT  
AATGGTAGGGGTATTTTTATACAAGTTAATGGTAGTTACTACATCACAGC  
ACGAGATATG  
TATATGCCAAGAGCTATTACTGCAGGAGATATAGTTACGCTTACTTCTTG  
TCAAGCAAAT  
TATGTAAGTGTAATAAGACCGTCATTACTACATTCGTAGACAATGATGAT  
TTTGATTTT  
AATGACGAATTGTCAAATGGTGGAAATGACACTAAGCATGAGCTACCAG  
ACTTTGACAAA  
TTCAATTACACAGTACCTATACTTGACATTGATAGTGAAATTGATCGTATT  
CAAGGCGTT  
ATACAGGGTCTTAATGACTCTTTAATAGACCTTGAAAACTTTCAATACT  
CAAACCTTAT  
ATTAAGTGGCCTTGGTATGTGTGGTTAGCCATAGCTTTTGCCACTATTAT

CTTCATCTTA  
ATACTAGGATGGGTTTTCTTCATGACTGGATGTTGTGGTTGTTGTTGTG  
GATGCTTTGGC  
ATTATGCCTCTAATGAGTAAGTGTGGTAAGAAATCTTCTTATTACACGAC  
TTTTGATAAC  
GATGTGGTAACTTAACAATACAGACCTAAAAAGTCTGTTTAATGATTCAA  
AGTCCCACGT  
CCTTCCTAATAGTATTAATTTTTCTTTGGTGTAAACTTGTACTAAGTTGTT  
TTAGAGAGT  
TTATTATAGCGCTCCAACAATAACAAGTTTTACTCCAAATTATCAATA  
GTAECTTAC  
AGCCTAGACTGACCCTTTGTCACAGTCTAGACTAATGTTAACTTAGAA  
GCAATTATTGA  
AACTGGTGAGCAAGTGATTCAAAAAATCAGTTTCAATTTACAGCATATTT  
CAAGTGTATT  
AAACACAGAAGTATTTGACCCCTTTGACTATTGTTATTACAGAGGAGGTA  
ATTTTTGGGA  
AATAGAGTCAGCTGAAGATTGTTTCAGGTGATGATGAATTTATTGAATAAG  
TCGCTAGAGG  
AAAATGGAAGTTTTCTAACAGCGCTTTATATATTTGTAGGATTTTTAGCA  
CTTTATCTTC  
TAGGTAGAGCACTTCAAGCATTTGTACAGGCTGCTGATGCTTGTTGTTT  
ATTTTGGTATA  
CATGGGTAGTAATTCCAGGAGCTAAGGGTACAGCCTTTGTATATAAGTAT  
ACATATGGTA  
GAAACTTAACAATCCGGAATTAGAAGCAGTTATTGTCAACGAGTTTCC  
TAAGAACGGTT  
GGAATAATAAAAATCCAGCAAATTTTCAAGATGTCCAACGAGACAAATT  
GACTCTTGAC  
TTTGAACAGTCAGTTGAGCTTTTTTAAAGAGTATAATTTATTTATAACTGC

ATTCTTGTTG  
TTCTTAACCATAATACTTCAGTATGGCTATGCAACAAGAAGTAAGTTTAT  
TTATATACTG  
AAAATGATAGTGTTATGGTGCTTTTGGCCCCTTAACATTGCAGTAGGTG  
TAATTTTCATGT  
ATATACCCACCAAACACAGGAGGTCTTGTCGCAGCGATAATACTTACAG  
TGTTTGCGTGT  
CTGTCTTTTGTAGGTTATTGGATCCAGAGTATTAGACTCTTTAAGCGGTG  
TAGGTCATGG

TGGTCATTTAACCCAGAATCTAATGCCGTAGGTTCAATACTCCTAACTAA  
TGGTCAACAA  
TGTAATTTTGCTATAGAGAGTGTGCCAATGGTGCTTTCTCCAATTATAAA  
GAATGGTGTT  
CTTTATTGTGAGGGTCAGTGGCTTGCTAAGTGTGAACCAGACCACTTG  
CCTAAAGATATA  
TTTGTTTGTACACCGGATAGACGTAATATCTACCGTATGGTGCAGAAATA  
TACTGGTGAC  
CAAAGCGGAAATAAGAAACGGTTTGCTACGTTTGTCTATGCAAAGCAGT  
CAGTAGATACT  
GGCGAGCTAGAAAGTGTAGCAACAGGAGGGAGTAGTCTTTACACCTAA  
ATGTGTGTGTGT  
AGAGAGTATTTAAAATTATTCTTTAATAGTGCCTCTATTTTAAGAGCGCAT  
AATAGTATT  
ATTTTTGAGGATATTAATATAAATCCTCTCTGTTTTATACTCTCTTTTCAA  
GAGCTATTA  
TTTAAAAACAGTTTTTCCACTCTTTTGTGCCAAAACTATTGTTGTTAA  
TGGTGTAACC  
TTTCAAGTAGATAATGGAAAAGTCTACTACGAAGGAAAACCAATTTTCA  
GAAAGGTTGT

TGTAGGTTGTGGTTGAGTTATAAAAAAGATTAAACTACCTACTACTTA  
TTTTTATAAG  
AGGCGTTTTATCTTACAAGCGCTTAATAAATACGGACGATGAAATGGCT  
GACTAGTTTTG  
TAAGGGCAGTTATTTTCATGTTATAAACCCCTATTATTA ACTCAATTAAGAG  
TATTAGATA  
GGTTAATCTTAGATCATGGACCAAACACATCTTAACGTGTGTTAGGTG  
CGTGATTTTTGT  
TTCAATTAGATTTAGTTTATAGGTTGGCGTATACGCCTACTCAATCGCTG  
GTATGAATAA  
TAGTAAAGATAATCCTTTTTGCGGAGCAATAGCAAGAAAAGCGCGAATT  
TATCTGAGAGA  
AGGATTAGATTGTGTTTACTTTCTTAACAAAGCAGGACAAGCAGAGTCT  
TGTCCCGCGTG  
TACCTCTCTAGTATTCCAGGGGAAA ACTTGTGAGGAACACAAATATAATA  
ATAATCTTTT  
GTCATGGCAAGCGGTAAGGCAACTGGAAAGACAGATGCCCCAGCTCCA  
GTCATCAA ACTA  
GGAGGACCAAAGCCACCTAAAGTTGGTTCTTCTGGAAATGTATCTTGGT  
TTCAAGCAATA  
AAAGCCAAGAAGTTAAATTCACCTCCGCCTAAGTTTGAAGGTAGCGGT  
GTTCTGATAAT  
GAAAATCTAAAACCAAGTCAGCAGCATGGATATTGGAGACGCCAAGCTA  
GGTTTAAGCCA  
GGTAAAGGTGGAAGAAAACCAAGTCCCAGATGCTTGGTATTTTTACTATA  
CTGGAACAGGA  
CCAGCCGCTAACCTGAATTGGGGTGATAGCCAAGATGGTATAGTGTGG  
GTTGCTGGTAAG  
GGTGCTGATACTAAATTTAGATCTAATCAGGGTACTCGTGACTCTGACAA  
GTTTGACCAA

TATCCGCTACGGTTTTTCAGACGGAGGACCTGATGGTAATTTCCGTTGG  
GATTTTCATTCCT  
CTGAATCGTGGCAGGAGTGGGAGATCAACAGCAGCTTCATCAGCAGCA  
TCTAGTAGAGCA  
CCATCACGTGAAGTTTCGCGTGGTCGCAGGAGTGGTTCTGAAGATGAT  
CTTATTGCTCGT  
GCAGCAAGGATAATTCAGGATCAGCAGAAGAAGGGTTCTCGCATTACA  
AAGGCTAAGGCT  
GATGAAATGGCTCACCGCCGGTATTGCAAGCGCACTATTCCACCTAATT  
ATAAGGTTGAT  
CAAGTGTTTGGTCCCCGTACTAAAGGTAAGGAGGGAAATTTTGGTGAT  
GACAAGATGAAT  
GAGGAAGGTATTAAGGATGGGCGCGTTACAGCAATGCTCAACCTAGTT  
CCTAGCAGCCAT  
GCTTGTCTTTTTCGGAAGTAGAGTGACGCCCAGACTTCAACCAGATGGG  
CTGCACTTGAAA  
TTTGAATTTACTACTGTGGTCCCACGTGATGATCCGCAGTTTGATAATTA  
TGTA AAAAATT  
TGTGATCAGTGTGTTGATGGTGTAGGAACACGTCCAAAAGATGATGAAC  
CAAGACCAAAG  
TCACGCTCAAGTTCAAGACCTGCAACAAGAGGAAATTCTCCAGCGCCA  
AGACAGCAGCGC  
CCTAAGAAGGAGAAAAAGCCAAAGAAGCAGGATGATGAAGTGGATAAA  
GCATTGACCTCA  
GATGAGGAGAGGAACAATGCACAGCTGGAATTTGATGATGAACCCAAG  
GTAATTA ACTGG  
GGGATTCAGCCCTAGGAGAGAATGAACTTTGAGTAAAATTCAATAGTA  
AGAGTTAAGGA  
AGATAGGCATGTAGCTTGATTACCTACATGTCTATCGCCAGGGAAATGTC  
TAATTTGTCT

ACTTAGTAGCCTGGAAACGAACGGTAGACCCTTAGATTTTAATTTAGTT  
TAATTTTTAGT  
TTAGTTTAAGTTAGTTT TAGAGTAGGTATAAAGATGCCAGTGCCGGGGGCC  
ACGCGGAGTAC  
GACCGAGGGTACAGCACTAGGACGCCCAT TAGGGGAAGAGCTAAATTT  
TAGTTTAAGTTA  
AGTTTAATTGGCTATGTATAGTTAAAATTTATAGGCTAGTATAGAGTTAGA  
GCAAAAAAA AAAAAAAAAAAAAAAAAAAAAA

## REPLICASE

In addition to the structural and accessory genes, two-thirds of a coronavirus genome comprises the replicase gene (at the 5' end of the genome), which is expressed as two polyproteins, ppla and ppl ab, in which pplab is an extension product of ppla as a result of a -1 ribosomal shift mechanism. The two polyproteins are cleaved by two types of virus-encoded proteinases usually resulting in 16 non-structural proteins (Nsp1 -16); IBV lacks Nsp1 thereby encoding Nsp2-16.

Thus Gene 1 in IBV encodes 15 (16 in other coronaviruses) non-structural proteins (nsp2- 16), which are associated with RNA replication and transcription.

The term 'replicase protein' is used herein to refer to the ppl a and pplab polyproteins or individual nsp subunits. The term 'replicase gene' is used herein to refer to a nucleic acid sequence which encodes for replicase proteins.

A summary of the functions of coronavirus nsp proteins is provided in Table 1.

The variant replicase gene encoded by the coronavirus of the present invention comprises a mutation in one or more of the sections of sequence encoding nsp-10, nsp-14, nsp-15 or nsp-16.

Nsp10 has RNA-binding activity and appears to be involved in homo and/or heterotypic interactions within other nsps from the pp1a/pp1ab region. It adopts an  $\alpha/\beta$  fold comprised of five  $\alpha$ -helices, one 3<sub>10</sub>-helix and three  $\beta$ -strands. Two zinc-binding sites have been identified that are formed by conserved cysteine residues and one histidine residue (Cys-74/Cys-77/His-83/Cys-90; Cys-117/Cys-120/Cys-128/Cys-130). The protein has been confirmed to bind single-stranded and double-stranded RNA and DNA without obvious specificity. Nsp-10 can be cross-linked with nsp-9, suggesting the existing of a complex network of protein-protein interactions involving nsp-7, -8, -9 and -10. In addition, nsp-10 is known to interact with nsp-14 and nsp-16.

Nsp-14 comprises a 3'-to-5' exoribonuclease (ExoN) active domain in the amino-terminal region. SARS-CoV ExoN has been demonstrated to have metal ion-dependent 3'-to-5' exoribonuclease activity that acts on both single-stranded and double-stranded RNA, but not on DNA. Nsp-14 has been shown to have proof-reading activity. This nsp has also been shown to have N7-methyltransferase (MT) activity in the carboxyl-terminal region. Nsp-15 associated NendoU (nidoviral endoribonuclease, specific for U) RNase activity has been reported for a number of coronaviruses, including SARS-CoV, MHV and IBV. The activities were consistently reported to be significantly enhanced by Mn<sup>2+</sup> ions and there was little activity in the presence of Mg<sup>2+</sup> and Ca<sup>2+</sup>. NendoU cleaves at the 3' side of uridylate residues in both single-stranded and double-stranded RNA. The biologically relevant substrate(s) of coronavirus NendoUs remains to be identified.

Nsp-16 has been predicted to mediate ribose-2'-O-methyltransferase (2'-O-MTase) activity and reverse-genetics experiments have shown that the 2'-O-MTase domain is essential for viral RNA synthesis in HCoV-229E and SARS-CoV. The enzyme may be involved in the production of the cap 1 structures of coronavirus RNAs and it may also cooperate with NendoU and ExoN in other RNA processing pathways. 2'-O-MTase might also methylate specific RNAs to protect them from NendoU-mediated cleavage. The genomic and protein sequences for nsp-10, -14, -15 and -16 are provided as SEQ ID NO: 2-5 and 6-9, respectively.

SEQ ID NO: 2 (nsD-10 nucleotide sequence - nucleotides [11884-12318](#) of SEQ ID NO:1 )

```
TCTA3/4AGGTCATGAGACAGAGGAAGTGGATGCTGTAGGCATTCTCTCA
CTTTGTTCTTTTGCAGTA
GATCCTGCGGATACATATTGTAAATATGTGGCAGCAGGTAATCAACCTTT
AGGTAACTGTGTTAAA
ATGTTGACAGTACATAATGGTAGTGGTTTTGCAATAACATCAAAGCCAAG
TCCAACTCCGGATCAG
GATTCTTATGGAGGAGCTTCTGTGTGTCTTTATTGTAGAGCACATATAGC
ACACCTTGGCGGAGCA
GGAAATTTAGATGGACGCTGTCAATTTAAAGGTTCTTTTGTGCAAATAC
CTACTACGGAGAAAGAT
CCTGTTGGATTCTGTCTACGTAACAAGGTTTGCACTGTTTGTCAGTGTT
GGATTGGTTATGGATGT
CAGTGTGATTCACTTAGACAACCTAAACCTTCTGTTCAG
```

SEQ ID NO: 3 (nsp-14 nucleotide sequence - nucleotides [16938-18500](#) of SEQ ID NO:1)

```
GGTACAGGCTTGTTTAAATTTGCAACAAAGAGTTTAGTGGTGTTCACC
```

CAGCTTATGCAGTCACA  
ACTAAGGCTCTTGCTGCAACTTATAAAGTTAATGATGAACTTGCTGCACT  
TGTTAACGTGGAAGCT  
GGTTCAGAAATAACATATAAACATCTTATTTCTTTGTTAGGGTTTAAGAT  
GAGTGTTAATGTTGAA  
GGCTGCCACAACATGTTTATAACACGTGATGAGGCTATCCGCAACGTAA  
GAGGTTGGGTAGGTTTT  
GATGTAGAAGCAACACATGCTTGCGGTACTAACATTGGTACTAACCTGC  
CTTTCCAAGTAGGTTTC  
TCTACTGGTGCAGACTTTGTAGTTACGCCTGAGGGACTTGTAGATACTT  
CAATAGGCAATAATTTT  
GAGCCTGTGAATTCTAAAGCACCTCCAGGTGAACAATTTAATCACTTGA  
GAGCGTTATTCAAAGT  
GCTAAACCTTGGCATGTTGTAAGGCCAAGGATTGTGCAAATGTTAGCG  
GATAACCTGTGCAACGTT  
TCAGATTGTGTAGTGTTTGTACCGTGGTGTCATGGCCTAGAACTAACCA  
CTTTGCGCTATTTTGTT  
AAAATAGGCAAGGACCAAGTTTGTCTTGCGGTTCTAGAGCAACAACCTT  
TTAATTCTCATACTCAG  
GCTTATGCTTGTTGGAAGCATTGCTTGGGTTTTGATTTTGTTTATAATCC  
ACTCTTAGTGGATATT  
CAACAGTGGGGTTATTCTGGTAACCTACAATTTAACCATGATTTGCATTG  
TAATGTGCATGGACAC  
GCACATGTAGCTTCTGCGGATGCTATTATGACGCGTTGTCTTGCAATTA  
ATAATGCATTTTGTCAA  
GATGTCAACTGGGATTTAACTTACCCTCATATAGCAAATGAGGATGAAGT  
CAATTCTAGCTGTAGA  
TATTTACAACGCATGTATCTTAATGCATGTGTTGATGCTCTTAAAGTTAA  
CGTTGTCTATGATATA  
GGCAACCCTAAAGGTATTAATGTGTTAGACGTGGAGACTTAAATTTTA

GATTCTATGATAAGAAT  
CCAATAGTACCCAATGTCAAGCAGTTTGAGTATGACTATAATCAGCACAA  
AGATAAGTTTGCTGAT  
GGTCTTTGTATGTTTTGGAATTGTAATGTGGATTGTTATCCCGACAATTC  
CTTACTTTGTAGGTAC  
GACACACGAAATTTGAGTGTGTTAACCTACCTGGTTGTAATGGTGGTA  
GCTTGTATGTTAACAAG  
CATGCATTCCACACACCTAAATTTGATCGCACTAGCTTTCGTAATTTGAA  
AGCTATGCCATTCTTT  
TTCTATGACTCATCGCCTTGCGAGACCATTCAATTGGATGGAGTTGCGC  
AAGACCTTGTGTCATTA  
GCTACGAAAGATTGTATCACAAAATGCAACATAGGCGGTGCTGTTTGTA  
AAAAGCACGCACAAATG  
TATGCAGATTTTGTGACTTCTTATAATGCAGCTGTTACTGCTGGTTTTAC  
TTTTTGGGTTACTAAT  
AATTTTAACCCATATAATTTGTGGAAAAGTTTTTCAGCTCTCCAG

SEQ ID NO: 4 (nsp-15 nucleotide sequence - nucleotides [18501-19514](#)  
of SEQ ID NO:1)

TCTATCGACAATATTGCTTATAATATGTATAAGGGTGGTCATTATGATGCT  
ATTGCAGGAGAAATG  
CCCCTATCGTAACTGGAGATAAAGTTTTTGTATAGATCAAGGCGTAGA  
AAAAGCAGTTTTTTTTT  
AATCAAACAATTCTGCCTACATCTGTAGCGTTTGAGCTGTATGCGAAGA  
GAAATATTCGCACACTG  
CCAAACAACCGTATTTTGAAAGGTTTGGGTGTAGATGTGACTAATGGAT  
TTGTAATTTGGGATTAC  
ACGAACCAAACACCACTATACCGTAATACTGTTAAGGTATGTGCATATAC  
AGACATAGAACCAAAT

GGCCTAATAGTGCTGTATGATGATAGATATGGTGATTACCAGTCTTTTCT  
AGCTGCTGATAATGCT  
GTTTTAGTTTCTACACAGTGTTACAAGCGGTATTCGTATGTAGAAATACC  
GTCAAACCTGCTTGTT  
CAGAACGGTATTCCGTAAAAGATGGAGCGAACCTGTATGTTTATAAGC  
GTGTTAATGGTGCGTTT  
GTTACGCTACCTAACACAATAAACACACAGGGTCGAAGTTATGAACTT  
TTGAACCTCGTAGTGAT  
GTTGAGCGTGATTTTCTCGACATGTCTGAGGAGAGTTTTGTAGAAAAGT  
ATGGTAAAGAATTAGGT  
CTACAGCACATACTGTATGGTGAAGTTGATAAGCCCCAATTAGGTGGTT  
TCCACACTGTTATAGGT  
ATGTGCAGACTTTTACGTGCGAATAAGTTGAACGCAAAGTCTGTTACTA  
ATTCTGATTCTGATGTC  
ATGCAAAATTATTTTGTATTGGCAGACAATGGTTCCTACAAGCAAGTGT  
GTAAGTGTGGATTTG  
CTGCTTGATGATTTCTTAGAACTTCTTAGGAACATACTGAAAGAGTATGG  
TACTAATAAGTCTAAA  
GTTGTAACAGTGTCAATTGATTACCATAGCATAAATTTTATGACTTGGTT  
TGAAGATGGCATTATT AAAACATGTTATCCACAGCTTCAA

SEQ ID NO: 5 (nsp-16 nucleotide sequence - nucleotides [19515-20423](#) of SEQ ID NO:1)

TCAGCATGGACGTGTGGTTATAATATGCCTGAACTTTATAAAGTTCAGAA  
TTGTGTTATGGAACCT  
TGCAACATTCCTAATTATGGTGTGGGAATAGCGTTGCCAAGTGGTATTAT  
GATGAATGTGGCAAAG  
TATACACAACCTCTGTCAATACCTTTCGAAAACAACAATGTGTGTACCGCA  
TAATATGCGAGTAATG

CATTTTGGAGCTGGAAGTGACAAAGGAGTGGTGCCAGGTAGTACTGTT  
CTTAAACAATGGCTCCCA  
GAAGGGGACACTCCTTGTTCGATAATGATATTGTAGACTATGTGTCTGATGC  
ACATGTTTCTGTGCTT  
TCAGATTGCAATAAATAAGACAGAGCACAAGTTTGATCTTGTGATATC  
TGATATGTATACAGAC  
AATGATTCAAAAAGAAAGCATGAAGGCGTGATAGCCAATAATGGCAATG  
ATGACGTTTTTCATATAT  
CTCTCAAGTTTTCTTCGTAATAATTTGGCTCTAGGTGGTAGTTTTGCTGT  
AAAAGTGACAGAGACA  
AGTTGGCACGAAGTTTTATATGACATTGCACAGGATTGTGCATGGTGGA  
CAATGTTTTGTACAGCA  
GTGAATGCCTCTTCTTCAGAAGCATTCTTGATTGGTGTTAATTATTTGGG  
TGCAAGTGAAAAGGTT  
AAGGTTAGTGGAAAAACGCTGCACGCAAATTATATATTTTTGGAGGAATT  
GTAATTATTTACAAACC  
TCTGCTTATAGTATATTTGACGTTGCTAAGTTTGATTTGAGATTGAAAGC  
AACGCCAGTTGTTAAT  
TTGAAAACCTGAACAAAAGACAGACTTAGTCTTTAATTTAATTAAGTGTGG  
TAAGTTACTGGTAAGA  
GATGTTGGTAACACCTCTTTTACTAGTGACTCTTTTGTGTGTACTATGTA  
G

SEQ ID NO: 6 (nsp-10 amino acid sequence)

SKGHETEEVDAVGILSLCSFAVDPADTYCKYVAAGNQPLGNCVKMLTVHN  
GSGFAITSKPSPTPDQ  
DSYGGASVCLYCRAHIAHPGGAGNLDGRCQFKGSFVQIPTTEKDPVGFCL  
RNKVCTVCQCWIGYGC QCDSL RQPKPSVQ

SEQ ID NO: 7 (nsp-14 amino acid sequence) GTGL FKICNKE

FSGVHPAYAVTTKALAATYKVNDELAALVNVEAGSE IT YKHL I  
SLLGFKMSVNVE  
GCHNMFITRDEAIRNVRGWVGFDFVEATHACGTNIGTNLFPQVGFSTGADF  
WTPEGLVDTSIGNNF  
EPVNSKAPPGEQFNHLRALFKSAKPWHWRPRIVQMLADNLCNVSDCWF  
VTWCHGLELTTLRVYFV  
KIGKDQVCSCGSRATTFNSHTQAYACWKHCLGDFVYNPLLVDIQQWGY  
SGNLQFNHDLHCNVHGH  
AHVASADAIMTRCLAINNAFCQDVNWDLTYPHIANEDEVNSSCRYLQRM  
YLNACVDALKVNVVYDI  
GNPKGIKVRRGDLNFRFYDKNPIVPNVKQFEYDYNQHKDKFADGLCMF  
WNCNVDCYPDNSLVCRY  
DTRNLSVFNLPGCNGGSLYVNKHAFHTPKFDRTSFRNLKAMPFFFYDSS  
PCETIQLDGVAQDLVSL AT KDC I T KCN I GGAVC KKHAQMY AD FVT S  
YNAAVT AG FT FWVTNNFNPYNLWKS FSALQ

SEQ ID NO: 8 (nsp-15 amino acid sequence)

SIDNIAYNMYKGGHYDAIAGEMPTIVTGDECVFVIDQGVEKAVFFNQITILPT  
SVAFELYAKRNIRTL  
PNNRILKGLGVDVTNGFVIWDYTNQTPLYRNTVKVCAYTDIEPNGLIVLYD  
DRYGDYQSFLAADNA  
VLVSTQCYKRYSYVEIPSNLLVQNGIPLKDGANLYVYKRVNGAFVTLPNTL  
NTQGRSYETFEPKSD  
VERDFLDMSEESFVEKYGKELGLQHILYGEVDKPQLGGLHTVIGMCROLLRA  
NKLNAKSVTNSDSV  
MQNYFVLADNGSYKQVCTWDLNLLDFLELLRNILKEYGTNKSQVTVSID  
YHSINFMTWFEDGII KTCYPQLQ SEQ ID NO: 9 (nsp-16 amino acid  
sequence)

SAWTCGYNMPELYKVQNCVMEPCNIPNYGVGIALPSGIMMNVAKYTQLC

QYLSKTTMCPHNMRVM  
HFGAGSDKGVAPGSTVLKQWLPEGTLLVDNDIVDYVSDAHVSVLSDCNK  
YKTEHKFDLVISDMYTD  
NDSKRKHEGVIANNGNDDVFIYLSSFLRNNLALGGFAVICVTETSWHEVL  
YDIAQDCAWWTMFCTA  
VNASSEAFLVGVNYLGASEKVKVSGKTLHANYIFWRNCNYLQTSAYSIF  
DVAKFDLRLKATPWN LKTEQKTDLVFNLI KCGKLLVRDVGNTS FT SDS  
FVCTM

**REDUCED PATHOGENICITY** The live, attenuated coronavirus of the present invention comprises a variant replicase gene which causes the virus to have reduced pathogenicity compared to a coronavirus expressing the corresponding wild-type gene.

The term "attenuated" as used herein, refers to a virus that exhibits said reduced pathogenicity and may be classified as non-virulent. A live, attenuated virus is a weakened replicating virus still capable of stimulating an immune response and producing immunity but not causing the actual illness.

The term "pathogenicity" is used herein according to its normal meaning to refer to the potential of the virus to cause disease in a subject. Typically the pathogenicity of a coronavirus is determined by assaying disease associated symptoms, for example sneezing, snicking and reduction in tracheal ciliary activity. The term "reduced pathogenicity" is used to describe that the level of pathogenicity of a coronavirus is decreased, lessened or diminished compared to a corresponding, wild-type coronavirus. In one embodiment, the coronavirus of the present invention has a reduced pathogenicity compared to the parental M41-CK virus from which it was derived or a control coronavirus. The control coronavirus may be a coronavirus

with a known pathogenicity, for example a coronavirus expressing the wild-type replicase protein. The pathogenicity of a coronavirus may be assessed utilising methods well-known in the art. Typically, pathogenicity is assessed by assaying clinical symptoms in a subject challenged with the virus, for example a chicken.

As an illustration, the chicken may be challenged at 8-24 days old by nasal or ocular inoculation. Clinical symptoms, associated with IBV infection, may be assessed 3-10 days post-infection. Clinical symptoms commonly assessed to determine the pathogenicity of a coronavirus, for example an IBV, include gasping, coughing, sneezing, snicking, depression, ruffled feathers and loss of tracheal ciliary activity. The variant replicase of the present invention, when expressed in a coronavirus, may cause a reduced level of clinical symptoms compared to a coronavirus expressing a wild-type replicase.

For example a coronavirus expressing the variant replicase may cause a number of snicks per bird per minute which is less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20% or less than 10% of the number of snicks caused by a virus expressing the wild type replicase.

A coronavirus expressing a variant replicase according to the present invention may cause wheezing in less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20% or less than 10% of the number of birds in a flock infected with the a virus expressing the wild type replicase.

A coronavirus expressing a variant replicase according to the present invention may result in tracheal ciliary activity which is at least 60%, at

least 70%, at least 80%, at least 90% or at least 95% of the level of tracheal ciliary activity in uninfected birds. A coronavirus expressing a variant replicase according to the present invention may cause clinical symptoms, as defined in Table 2, at a lower level than a coronavirus expressing the wild type replicase.

The variant replicase of the present invention, when expressed in a coronavirus, may cause the virus to replicate at non-pathogenic levels in ovo.

While developing vaccines to be administered in ovo to chicken embryos, attention must be paid to two points: the effect of maternal antibodies on the vaccines and the effect of the vaccines on the embryo. Maternal antibodies are known to interfere with active immunization. For example, vaccines with mild strains do not induce protective antibody levels when administered to broiler chickens with maternal antibodies as these strains are neutralized by the maternal antibody pool. Thus a viral particle must be sufficiently efficient at replicating and propagating to ensure that it is not neutralized by the maternally-derived antibodies against the virus. Maternally-derived antibodies are a finite pool of effective antibodies, which decrease as the chicken ages, and neutralization of the virus in this manner does not equate to the establishment of long-term immunity for the embryo/chick. In order to develop long-term immunity against the virus, the embryo and hatched chicken must develop an appropriate protective immune response which is distinct to the effect of the maternally-derived antibodies. To be useful for in ovo vaccination, the virus must also not replicate and propagate at a level which causes it to be pathogenic to the embryo.

Reduced pathogenicity in terms of the embryo may mean that the

coronavirus causes less reduction in hatchability compared to a corresponding, wild-type control coronavirus. Thus the term "without being pathogenic to the embryo" in the context of the present invention may mean "without causing reduced hatchability" when compared to a control coronavirus.

A suitable variant replicase may be identified using methods which are known in the art. For example comparative challenge experiments following in ovo vaccination of embryos with or without maternally-derived antibodies may be performed (i.e. wherein the layer has or has not been vaccinated against IBV).

If the variant replicase enables the virus to propagate at a level which is too high, the embryo will not hatch or will not be viable following hatching (i.e. the virus is pathogenic to the embryo). A virus which is pathogenic to the embryo may kill the embryo.

If the variant replicase causes a reduction in viral replication and propagation which is too great, the virus will be neutralised by the maternally-derived antibodies. Subsequent challenge of the chick with IBV will therefore result in the development of clinical symptoms (for example wheezing, snicking, loss of ciliary activity) and the onset of disease in the challenged chick; as it will have failed to develop effective immunity against the virus. VARIANT As used herein, the term 'variant' is synonymous with 'mutant' and refers to a nucleic acid or amino acid sequence which differs in comparison to the corresponding wild-type sequence. A variant/mutant sequence may arise naturally, or may be created artificially (for example by site-directed mutagenesis). The mutant may have at least 70, 80, 90, 95, 98 or 99% sequence identity with the corresponding portion of the wild type sequence. The mutant may have less than 20, 10, 5, 4, 3, 2 or

1 mutation(s) over the corresponding portion of the wild-type sequence.

The term "wild type" is used to mean a gene or protein having a nucleotide or amino acid sequence which is identical with the native gene or protein respectively (i.e. the viral gene or protein). Identity comparisons can be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs can calculate % identity between two or more sequences. A suitable computer program for carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U.S.A.; Devereux et al., 1984, *Nucleic Acids Research* 12:387). Examples of other software that can perform sequence comparisons include, but are not limited to, the BLAST package (see Ausubel et al., 1999 *ibid* - Chapter 18), FASTA (Atschul et al., 1990, *J. Mol. Biol.*, 403-410) and the GENWORKS suite of comparison tools, ClustalX (see Larkin et al. (2007) *Clustal W and Clustal X version 2.0. Bioinformatics*, 23:[2947-2948](#)). Both BLAST and FASTA are available for offline and online searching (see Ausubel et al., 1999 *ibid*, pages 7-58 to 7-60). However, for some applications, it is preferred to use the GCG Bestfit program. A new tool, called BLAST 2 Sequences is also available for comparing protein and nucleotide sequence (see *FEMS Microbiol Lett* 1999 174(2): 247-50; *FEMS Microbiol Lett* 1999 177(1): 187-8 and [tatiana@ncbi.nlm.nih.gov](mailto:tatiana@ncbi.nlm.nih.gov)).

The sequence may have one or more deletions, insertions or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent molecule. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature

of the residues as long as the activity is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine, valine, glycine, alanine, asparagine, glutamine, serine, threonine, phenylalanine, and tyrosine.

Conservative substitutions may be made, for example according to the Table below, Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other:

The coronavirus of the present invention may comprise a variant replicase gene which encodes a protein which comprises a mutation compared to any one of SEQ ID NO: 6, 7, 8 or 9 which, when expressed in a coronavirus, causes the virus to have reduced pathogenicity compared to a coronavirus expressing the corresponding wild-type replicase. The variant replicase gene may encode a protein which comprises at least one or more amino acid mutations in any combination of nsp-10, nsp-14, nsp-15 and nsp-16.

The variant replicase gene of the coronavirus of the present invention may encode a protein comprising a mutation as defined in the M41 mod sequences presented in Figure 10.

The variant replicase gene of the coronavirus of the present invention may encode a protein which comprises one or more amino acid mutations selected from the list of:

Pro to Leu at position 85 of SEQ ID NO: 6,

Val to Leu at position 393 of SEQ ID NO: 7;

Leu to lie at position 183 of SEQ ID NO: 8;

Val to lie at position 209 of SEQ ID NO: 9. The variant replicase gene of the coronavirus of the present invention may encode a protein which does not comprise a mutation in nsp-2, nsp-3, nsp-6 or nsp-13.

The variant replicase gene of the coronavirus of the present invention may encode a protein which does not comprise a mutation in nspl 0 which corresponds to the threonine to isoleucine mutation caused by a mutation at nucleotide position 12,008 in the gene reported by Ammayappan et al. (Arch Virol (2009) 154:495-499).

Ammayappan et al (as above) reports the identification of sequence changes responsible for the attenuation of IBV strain Arkansas DPI. The study identified 17 amino acid changes in a variety of IBV proteins following multiple passages, approx. 100, of the virus in embryonated eggs. It was not investigated whether the attenuated virus (Ark DP1 101) is capable of replicating in the presence of maternally-derived antibodies against the virus in ovo, without being pathogenic to the embryo. Given that this virus was produced by multiple passage in SPF embryonated eggs, similar methodology for classical IBV vaccines, it is likely that this virus is pathogenic for embryos. The virus may also be sensitive to maternally-derived antibodies if the hens were vaccinated with a similar serotype. The variant replicase gene of the coronavirus of the present invention may encode a protein which comprises any combination of one or more amino acid mutations provided in the list above.

The variant replicase gene may encode a protein which comprises the

amino acid mutation Pro to Leu at position 85 of SEQ ID NO: 6.

The variant replicase gene may encode a protein which comprises the amino acid mutation Val to Leu at position 393 of SEQ ID NO: 7. The variant replicase gene may encode a protein which comprises the amino acid mutation Leu to lie at position 183 of SEQ ID NO: 8.

The variant replicase gene may encode a protein which comprises the amino acid mutation Val to lie at position 209 of SEQ ID NO: 9. The variant replicase gene may encode a protein which comprises the amino acid mutations Pro to Leu at position 85 of SEQ ID NO: 6, and Val to Leu at position 393 of SEQ ID NO: 7. The variant replicase gene may encode a protein which comprises the amino acid mutations Pro to Leu at position 85 of SEQ ID NO: 6 Leu to lie at position 183 of SEQ ID NO: 8.

The variant replicase gene may encode a protein which comprises the amino acid mutations Pro to Leu at position 85 of SEQ ID NO: 6 and Val to lie at position 209 of SEQ ID NO: 9.

The variant replicase gene may encode a protein which comprises the amino acid mutations Val to Leu at position 393 of SEQ ID NO: 7 and Leu to lie at position 183 of SEQ ID NO: 8.

The variant replicase gene may encode a protein which comprises the amino acid mutations Val to Leu at position 393 of SEQ ID NO: 7 and Val to lie at position 209 of SEQ ID NO: 9.

The variant replicase gene may encode a protein which comprises the amino acid mutations Leu to lie at position 183 of SEQ ID NO: 8 and Val to Ile at position 209 of SEQ ID NO: 9. The variant replicase gene may encode a protein which comprises the amino acid mutations Pro

to Leu at position 85 of SEQ ID NO: 6, Val to Leu at position 393 of SEQ ID NO: 7 and Leu to lie at position 183 of SEQ ID NO: 8.

The variant replicase gene may encode a protein which comprises the amino acid mutations Pro to Leu at position 85 of SEQ ID NO: 6 Leu to lie at position 183 of SEQ ID NO: 8 and Val to lie at position 209 of SEQ ID NO: 9.

The variant replicase gene may encode a protein which comprises the amino acid mutations Pro to Leu at position 85 of SEQ ID NO: 6, Val to Leu at position 393 of SEQ ID NO: 7 and Val to lie at position 209 of SEQ ID NO: 9. The variant replicase gene may encode a protein which comprises the amino acid mutations Val to Leu at position 393 of SEQ ID NO: 7, Leu to lie at position 183 of SEQ ID NO: 8 and Val to lie at position 209 of SEQ ID NO: 9. The variant replicase gene may encode a protein which comprises the amino acid mutations Pro to Leu at position 85 of SEQ ID NO: 6, Val to Leu at position 393 of SEQ ID NO: 7, Leu to lie at position 183 of SEQ ID NO: 8 and Val to lie at position 209 of SEQ ID NO: 9. The variant replicase gene may also be defined at the nucleotide level.

For example the nucleotide sequence of the variant replicase gene of the coronavirus of the present invention may comprise one or more nucleotide substitutions within the regions selected from the list of: [11884-12318](#), [16938-18500](#), [18501-19514](#) and 19515- 20423 of SEQ ID NO:1.

For example the nucleotide sequence of the variant replicase gene of the coronavirus of the present invention may comprise one or more nucleotide substitutions selected from the list of:

C to T at nucleotide position 12137;

G to C at nucleotide position 181 14;

T to A at nucleotide position 19047; and

G to A at nucleotide position 20139;

compared to the sequence shown as SEQ ID NO: 1.

As used herein, the term "substitution" is synonymous with the term mutation and means that the nucleotide at the identified position differs to that of the wild-type nucleotide sequence. The nucleotide sequence may comprise any combination of the nucleotide substitutions selected from the list of:

C to T at nucleotide position 12137;

G to C at nucleotide position 181 14;

T to A at nucleotide position 19047; and

G to A at nucleotide position 20139;

compared to the sequence shown as SEQ ID NO: 1. The nucleotide sequence may comprise the substitution C12137T. The nucleotide sequence may comprise substitution G181 14C. The nucleotide sequence may comprise the substitution T19047A. The nucleotide sequence may comprise the substitution G20139A. The nucleotide sequence may comprise the substitutions C12137T and G181 14C.

The nucleotide sequence may comprise the substitutions C12137T and T19047A. The nucleotide sequence may comprise the

substitutions C12137T and G20139A. The nucleotide sequence may comprise the substitutions G18114C and T19047A. The nucleotide sequence may comprise the substitutions G18114C and G20139A. The nucleotide sequence may comprise the substitutions T19047A and G20139A.

The nucleotide sequence may comprise the substitutions C12137T, G181 14C and T19047A.

The nucleotide sequence may comprise the substitutions C12137T, T19047A and G20139A.

The nucleotide sequence may comprise the substitutions C12137T, G181 14C and G20139A. The nucleotide sequence may comprise the substitutions G18114C, T19047A and G20139A.

The nucleotide sequence may comprise the substitutions C12137T, G181 14C, T19047A and G20139A.

The nucleotide sequence may not comprise a substitution which corresponds to the C12008T substitution reported by Ammayappan et al. (as above). The nucleotide sequence may be natural, synthetic or recombinant. It may be double or single stranded, it may be DNA or RNA or combinations thereof. It may, for example, be cDNA, PCR product, genomic sequence or mRNA.

The nucleotide sequence may be codon optimised for production in the host/host cell of choice.

It may be isolated, or as part of a plasmid, virus or host cell.

PLASMID

A plasmid is an extra-chromosomal DNA molecule separate from the chromosomal DNA which is capable of replicating independently of the chromosomal DNA. They are usually circular and double-stranded.

Plasmids, or vectors (as they are sometimes known), may be used to express a protein in a host cell. For example a bacterial host cell may be transfected with a plasmid capable of encoding a particular protein, in order to express that protein. The term also includes yeast artificial chromosomes and bacterial artificial chromosomes which are capable of accommodating longer portions of DNA.

The plasmid of the present invention comprises a nucleotide sequence capable of encoding a defined region of the replicase protein. It may also comprise one or more additional coronavirus nucleotide sequence(s), or nucleotide sequence(s) capable of encoding one or more other coronavirus proteins such as the S gene and/or gene 3.

The plasmid may also comprise a resistance marker, such as the guanine xanthine phosphoribosyltransferase gene (gpt) from *Escherichia coli*, which confers resistance to mycophenolic acid (MPA) in the presence of xanthine and hypoxanthine and is controlled by the vaccinia virus P7.5 early/late promoter.

**RECOMBINANT VACCINIA VIRUS** The present invention also relates to a recombinant vaccinia virus (rVV) comprising a variant replicase gene as defined herein. The recombinant vaccinia virus (rVV) may be made using a vaccinia-virus based reverse genetics system.

In this respect, the present invention also provides a method for making a viral particle by:

(i) transfecting a plasmid as described in the previous section into a

host cell;

(ii) infecting the host cell with a recombining virus comprising the genome of a coronavirus strain with a replicase gene;

(iii) allowing homologous recombination to occur between the replicase gene sequences in the plasmid and the corresponding sequences in the recombining virus genome to produce a modified replicase gene;

(iv) selecting for recombining virus comprising the modified replicase gene.

The term 'modified replicase gene' refers to a replicase gene which comprises a variant replicase gene as described in connection with the first aspect of the present invention. Specifically, the term refers to a gene which is derived from a wild-type replicase gene but comprises a nucleotide sequence which causes it to encode a variant replicase protein as defined herein.

The recombination may involve all or part of the replicase gene. For example the recombination may involve a nucleotide sequence encoding for any combination of nsp- 10, nsp-14, nsp-15 and/or nsp-16. The recombination may involve a nucleotide sequence which encodes for an amino acid mutation or comprises a nucleotide substitution as defined above. The genome of the coronavirus strain may lack the part of the replicase protein corresponding to the part provided by the plasmid, so that a modified protein is formed through insertion of the nucleotide sequence provided by the plasmid.

The recombining virus is one suitable to allow homologous recombination between its genome and the plasmid. The vaccinia

virus is particularly suitable as homologous recombination is routinely used to insert and delete sequences for the vaccinia virus genome.

The above method optionally includes the step:

(v) recovery of recombinant coronavirus comprising the modified replicase gene from the DNA from the recombining virus from step (iv). Methods for recovering recombinant coronavirus, such as recombinant IBV, are known in the art (See Britton et al (2005) see page 24; and PCT/GB2010/001293).

For example, the DNA from the recombining virus from step (iv) may be inserted into a plasmid and used to transfect cells which express cytoplasmic T7 RNA polymerase. The cells may, for example be pre-infected with a fowlpox virus expressing T7 RNA polymerase.

Recombinant coronavirus may then be isolated, for example, from the growth medium. When the plasmid is inserted into the vaccinia virus genome, an unstable intermediate is formed. Recombinants comprising the plasmid may be selected for e.g. using a resistance marker on the plasmid.

Positive recombinants may then be verified to contain the modified replicase gene by, for example, PCR and sequencing.

Large stocks of the recombining virus including the modified replicase gene (e.g. recombinant vaccinia virus, (rVV) may be grown up and the DNA extracted in order to carry out step (v)).

Suitable reverse genetics systems are known in the art (Casais et al (2001) J. Virol 75:[12359-12369](#); Casais et al (2003) J. Virol. 77:[9084-9089](#); Britton et al (2005) J. Virological Methods 123:203-211 ; Armesto et al (2008) Methods in Molecular Biology 454:255-273).

## CELL

The coronavirus may be used to infect a cell. Coronavirus particles may be harvested, for example from the supernatant, by methods known in the art, and optionally purified.

The cell may be used to produce the coronavirus particle. Thus the present invention also provides a method for producing a coronavirus which comprises the following steps:

- (i) infection of a cell with a coronavirus according to the invention; (ii) allowing the virus to replicate in the cell; and
- (iii) harvesting the progeny virus.

The present invention also provides a cell capable of producing a coronavirus according to the invention using a reverse genetics system. For example, the cell may comprise a recombining virus genome comprising a nucleotide sequence capable of encoding the replicase gene of the present invention.

The cell may be able to produce recombinant recombining virus (e.g. vaccinia virus) containing the replicase gene.

Alternatively the cell may be capable of producing recombinant coronavirus by a reverse genetics system. The cell may express or be induced to express T7 polymerase in order to rescue the recombinant viral particle.

## VACCINE

The coronavirus may be used to produce a vaccine. The vaccine may be a live attenuated form of the coronavirus of the present invention

and may further comprise a pharmaceutically acceptable carrier. As defined herein, "pharmaceutically acceptable carriers" suitable for use in the invention are well known to those of skill in the art. Such carriers include, without limitation, water, saline, buffered saline, phosphate buffer, alcohol/aqueous solutions, emulsions or suspensions. Other conventionally employed diluents and excipients may be added in accordance with conventional techniques. Such carriers can include ethanol, polyols, and suitable mixtures thereof, vegetable oils, and injectable organic esters. Buffers and pH adjusting agents may also be employed. Buffers include, without limitation, salts prepared from an organic acid or base. Representative buffers include, without limitation, organic acid salts, such as salts of citric acid, e.g., citrates, ascorbic acid, gluconic acid, histidine-Hel, carbonic acid, tartaric acid, succinic acid, acetic acid, or phthalic acid, Tris, trimethanmine hydrochloride, or phosphate buffers. Parenteral carriers can include sodium chloride solution, Ringer's dextrose, dextrose, trehalose, sucrose, and sodium chloride, lactated Ringer's or fixed oils. Intravenous carriers can include fluid and nutrient replenishes, electrolyte replenishers, such as those based on Ringer's dextrose and the like. Preservatives and other additives such as, for example, antimicrobials, antioxidants, chelating agents (e.g., EDTA), inert gases and the like may also be provided in the pharmaceutical carriers. The present invention is not limited by the selection of the carrier. The preparation of these pharmaceutically acceptable compositions, from the above-described components, having appropriate pH isotonicity, stability and other conventional characteristics is within the skill of the art. See, e.g., texts such as Remington: The Science and Practice of Pharmacy, 20th ed, Lippincott Williams & Wilkins, publ., 2000; and The Handbook of Pharmaceutical Excipients, 4.sup.th edit., eds. R. C. Rowe et al, APhA Publications, 2003.

The vaccine of the invention will be administered in a "therapeutically effective amount", which refers to an amount of an active ingredient, e.g., an agent according to the invention, sufficient to effect beneficial or desired results when administered to a subject or patient. An effective amount can be administered in one or more administrations, applications or dosages. A therapeutically effective amount of a composition according to the invention may be readily determined by one of ordinary skill in the art. In the context of this invention, a "therapeutically effective amount" is one that produces an objectively measured change in one or more parameters associated Infectious Bronchitis condition sufficient to effect beneficial or desired results .An effective amount can be administered in one or more administrations. For purposes of this invention, an effective amount of drug, compound, or pharmaceutical composition is an amount sufficient to reduce the incidence of Infectious Bronchitis. As used herein, the term "therapeutic" encompasses the full spectrum of treatments for a disease, condition or disorder. A "therapeutic" agent of the invention may act in a manner that is prophylactic or preventive, including those that incorporate procedures designed to target animals that can be identified as being at risk (pharmacogenetics); or in a manner that is ameliorative or curative in nature; or may act to slow the rate or extent of the progression of at least one symptom of a disease or disorder being treated.

The present invention also relates to a method for producing such a vaccine which comprises the step of infecting cells, for example Vero cells, with a viral particle comprising a replicase protein as defined in connection with the first aspect of the invention.

**VACCINATION METHOD** The coronavirus of the present invention may

be used to treat and/or prevent a disease. To "treat" means to administer the vaccine to a subject having an existing disease in order to lessen, reduce or improve at least one symptom associated with the disease and/or to slow down, reduce or block the progression of the disease. To "prevent" means to administer the vaccine to a subject who has not yet contracted the disease and/or who is not showing any symptoms of the disease to prevent or impair the cause of the disease (e.g. infection) or to reduce or prevent development of at least one symptom associated with the disease. The disease may be any disease caused by a coronavirus, such as a respiratory disease and and/or gastroenteritis in humans and hepatitis, gastroenteritis, encephalitis, or a respiratory disease in other animals.

The disease may be infectious bronchitis (IB); Porcine epidemic diarrhoea; Transmissible gastroenteritis; Mouse hepatitis virus; Porcine haemagglutinating encephalomyelitis; Severe acute respiratory syndrome (SARS); or Bluecomb disease.

The disease may be infectious bronchitis. The vaccine may be administered to hatched chicks or chickens, for example by eye drop or intranasal administration. Although accurate, these methods can be expensive e.g. for large broiler flocks. Alternatives include spray inoculation of administration to drinking water but it can be difficult to ensure uniform vaccine application using such methods. The vaccine may be provided in a form suitable for its administration, such as an eye- dropper for intra-ocular use.

The vaccine may be administered by in ovo inoculation, for example by injection of embryonated eggs. In ovo vaccination has the advantage that it provides an early stage resistance to the disease. It also facilitates the administration of a uniform dose per subject, unlike

spray inoculation and administration via drinking water.

The vaccine may be administered to any suitable compartment of the egg, including allantoic fluid, yolk sac, amnion, air cell or embryo. It may be administered below the shell (aircell) membrane and chorioallantoic membrane. Usually the vaccine is injected into embryonated eggs during late stages of embryonic development, generally during the final quarter of the incubation period, such as 3-4 days prior to hatch. In chickens, the vaccine may be administered between day 15-19 of the 21 -day incubation period, for example at day 17 or 18.

The process can be automated using a robotic injection process, such as those described in WO [2004/078203](#).

The vaccine may be administered together with one or more other vaccines, for example, vaccines for other diseases, such as Newcastle disease virus (NDV). The present invention also provides a vaccine composition comprising a vaccine according to the invention together with one or more other vaccine(s). The present invention also provides a kit comprising a vaccine according to the invention together with one or more other vaccine(s) for separate, sequential or simultaneous administration.

The vaccine or vaccine composition of the invention may be used to treat a human, animal or avian subject. For example, the subject may be a chick, chicken or mouse (such as a laboratory mouse, e.g. transgenic mouse). Typically, a physician or veterinarian will determine the actual dosage which will be most suitable for an individual subject or group of subjects and it will vary with the age, weight and response of the particular subject(s).

The composition may optionally comprise a pharmaceutically acceptable carrier, diluent, excipient or adjuvant. The choice of pharmaceutical carrier, excipient or diluent can be selected with regard to the intended route of administration and standard pharmaceutical practice. The pharmaceutical compositions may comprise as (or in addition to) the carrier, excipient or diluent, any suitable binder(s), lubricant(s), suspending agent(s), coating agent(s), solubilising agent(s), and other carrier agents that may aid or increase the delivery or immunogenicity of the virus.

The invention will now be further described by way of Examples, which are meant to serve to assist one of ordinary skill in the art in carrying out the invention and are not intended in any way to limit the scope of the invention.

#### EXAMPLES EXAMPLE 1 - Generation of an IBV reverse genetics system based on M41 -CK

A M41-CK full-length cDNA was produced by replacement of the Beaudette cDNA in the Vaccinia virus reverse genetics system previously described in PCT/GB2010/001293 (herein incorporated by reference) with synthetic cDNA derived from the M41 consensus sequence.

The IBV cDNA within recombinant Vaccinia virus (rVV) rVV-BeauR-Rep-M41 structure described in Armesto, Cavanagh and Britton (2009). PLoS ONE 4(10): e7384. doi:10.1371/journal.pone.0007384, which consisted of the replicase derived from IBV Beaudette strain and the structural and accessory genes and 3' UTR from IBV M41 -CK, was further modified by replacement of the Beaudette 5' UTR-Nsp2-Nsp3 sequence with the corresponding sequence from IBV M41 -CK.

The resulting IBV cDNA consisted of 5' UTR-Nsp2-Nsp3 from M41 , Nsp4-Nsp16 from Beaudette and the structural and accessory genes and 3' UTR from M41. This cDNA was further modified by the deletion of the Beaudette Nsp4-Nsp16 sequence. The resulting cDNA, lacking Nsp4-16, was modified in four further steps in which the deleted Nsps were sequentially replaced with the corresponding sequences from M41 -CK, the replacement cDNAs represented M41 - CK Nsp4-8, Nsp9-12, Nsp12-14 and finally Nsp15-16. Each replacement cDNA contained approx. 500 nucleotides at the 5' end corresponding to the 3' most M41 sequence previously inserted and approx. 500 nucleotides at the 3' end corresponding to the M41 S gene sequence. This allowed insertion of the M41 cDNA sequence by homologous recombination and sequential addition of contiguous M41 replicase gene sequence. The synthetic cDNAs containing the M41 -derived Nsp sequences were added by homologous recombination utilising the inventor's previous described transient dominant selection (TDS) system (see PCT/GB2010/001293). The M41 -derived cDNAs containing sequence corresponding to the M41 Nsps-10, -14, -15 and -16 contained the modified amino acids at positions 85, 393, 183 and 209, respectively, as indicated in Figure 10.

A full-length cDNA representing the genome of M41 -CK was generated in Vaccinia virus representing the synthetic sequences. Two rIBVs, M41 -R-6 and M41 -R-12, were rescued and shown to grow in a similar manner as M41 -CK (Fig. 1 ).

#### EXAMPLE 2 - Determining the pathogenicity of rescued M41 viruses

The viruses rescued in Example 1 were used to infect 8-day-old specific pathogen free

(SPF) chicks by ocular and nasal inoculation to test them for pathogenicity, as observed by clinical signs on a daily basis 3-7 days post-infection and for ciliary activity days 4 and 6 post-infection. Loss of ciliary activity is a well-established method for determining the pathogenicity of IBV. The two M41 -R viruses were found to be apathogenic when compared to M41-CK though they did show some clinical signs in comparison to uninfected control chicks (Fig. 2) and some but inconsistent loss in ciliary activity (Fig. 3).

Thus, the M41 -R molecular clones of M41 -CK were not pathogenic when compared to the parental virus M41 -CK. The inventors identified several nucleotide differences in the M41 -R compared to the M41 -CK sequences. The majority of these were synonymous mutations, as the nucleotide change did not affect the amino acid sequence of the protein associated with the sequence. However, four non-synonymous mutations were identified in the IBV replicase gene specific to Nsp-10, Nsp-14, Nsp-15 and Nsp-16 components of the replicase gene, these mutations resulted in amino acid changes (Table 3).

Table 3. Non-Synonymous mutations identified in the NSDS of M41-R full-length genome

### EXAMPLE 3 - Repair of M41 -R rIBVs

In order to determine whether the identified mutations were responsible for the loss of pathogenicity associated with M41 -R, the Nsp10 mutation was repaired and the mutations in Nsp-14, -15 & -16 were repaired and shown to grow in a similar manner as M41 -CK (Fig 9). The inventors thus generated the rIBVs, M41 R-nsp10rep and M41 R-nsp14, 15, 16rep, using synthetic cDNAs containing the correct nucleotides utilising the inventor's previously described (TDS) system

(see PCT/GB2010/001293). The rIBVs were assessed for pathogenicity in chicks as described previously. Both rIBVs showed increased pathogenicity when compared to M41 -R but not to the level observed with M41 -CK (Figs 4 and 5). M41 R-nsp14, 15, 16rep gave more clinical signs and more reduction in ciliary activity than M41 R-nsp10rep, overall these results indicated that the changes associated with the four Nsp's appear to affect pathogenicity.

To determine the roles of the Nsp's in pathogenicity the full-length cDNA corresponding to M41 R-nsp10rep was used to repair the mutations in Nsp's 14, 15 & 16 using a synthetic cDNA containing the correct nucleotides utilising the TDS system.

The following rIBVs were produced:-

M41 R-nsp10, 15rep - M41-R with the mutations in Nsp-10 and Nsp-15 repaired

M41 R-nsp10, 14, 15rep - M41-R with mutations in Nsp-10, -14 and -15 repaired

M41 R-nsp10, 14, 16rep - M41-R with mutations in Nsp-10, -14 and -16 repaired

M41 R-nsp10, 15, 16rep - M41 -R with mutations in Nsp-10, -15 and -16 repaired

M41-K - All four mutations, Nsp-10,-14,-15 & -16 repaired in M41 -R

The rIBVs were shown to grow in a similar manner as M41 -CK (Fig 9) and assessed for pathogenicity as described previously. M41 -K (in which all four mutations had been repaired) resulted in clinical signs

and 100% loss of ciliary activity (complete ciliostasis) by 4 days post-infection (Fig. 6, 7 & 8). The other rIBVs demonstrated varying levels of pathogenicity, apart from M41 R-nsp10, 15, 16rep, which was essentially apathogenic. These results confirmed that repair of all four Nsps restored pathogenicity to M41 -R; again supporting the previous evidence that the mutations described in the four Nsps are implicated in attenuating M41-CK.

The inventors also generated rIBV M41 R-nsp 10, 14 rep (nsp 10 and 14 are repaired, nsp 15 and 16 contain mutations) and rIBV M41 R-nsp 10, 16 rep (nsp 10 and 16 are repaired, nsp 14 and 15 contain mutations) and assessed the pathogenicity of these viruses. rIBV M41 R-nsp 10, 14 rep less pathogenic than M41-K but caused around 50% ciliostasis on days 4-6 post-infection. rIBV M41 R-nsp 10, 16 rep was almost apathogenic and caused no ciliostasis (see Figure 11 a-c). Thus the genome associated with M41-R is a potential backbone genome for a rationally attenuated IBV. EXAMPLE 4 - Vaccination/Challenge Study with M41 -R

Candidate vaccine viruses were tested in studies in which fertilized chicken eggs were vaccinated in ovo at 18 days embryonation and in which the hatchability of the inoculated eggs was determined. The clinical health of the chickens was investigated and the chickens were challenged at 21 days of age with a virulent IB M41 challenge virus at 103.65 EID50 per dose.

Clinical signs were investigated after challenge protection by the vaccine and a ciliostasis test was performed at 5 days after challenge to investigate the effect of the challenge viruses on movement of the cilia and protection by the vaccine against ciliostasis (inhibition of cilia movement).

In ovo vaccination in commercial broiler eggs The design of the experiment is given in Table 4 and the clinical results are given in Table 5. Hatchability of the eggs inoculated with IB M41 -R was good and chickens were healthy. IB M41-R protected against clinical signs after challenge in the broilers (placebo: 19/19 affected, IB M41 -R: 3/18 affected and 1 dead). The results of the ciliostasis test are given in Table 6. IB M41-R generated protection against ciliostasis.

Table 4 - Design of a hatchability, safety, efficacy study in commercial eggs

1 Dose volume 0.1 ml, NA, not applicable.

2 103.65 EID50 per dose.

Table 5 - Hatch percentages and clinical data before and after challenge in commercial chickens, for design see Table 1.

1 Disturbed respiratory system

2 Whizzing

3 Change of voice

4 Breathing difficult

5 Swollen intra-orbital sinuses

6 Uneven growth

7 Weak

Table 6 - Results of the ciliostasis test after challenge for design see

Table 1.

In ovo vaccination in specific pathogen-free (SPF) eggs

The design of the study in SPF eggs is given in Table 7 and is similar with the design of the studies with commercial broilers, but the vaccination dose for IB M41 -R was higher, (10<sup>5</sup> EID<sub>50</sub> per dose).

The results (Table 8) show that the hatch percentage for IB M41 -R hatch was low, and 19 of 40 hatched and the chicks were weak. Eight chicks died. The remaining 11 chickens were challenged and 11 of the chicks hatched from the eggs which had been inoculated with saline were challenged.

In the ciliostasis test after challenge it appeared that all chickens vaccinated in ovo with IB M41-R were protected, whereas none of the controls was protected, see Table 9. Table 7. Design of a hatchability, safety, efficacy study in SPF eggs

Dose volume 0.1 ml, NA, not applicable.

2 Challenge dose 10<sup>3.3</sup> EID<sub>50</sub> in 0.2 ml.

Table 8. Hatch percentages and clinical data before and after challenge in SPF chickens, for design see Table 7.

Table 9. Results of the ciliostasis test after challenge, for design see Table 7.

In conclusion, IB M41 -R was safe in commercial eggs, generated protection against clinical signs and to an extent against ciliostasis.

In SPF eggs vaccinated with IB M41 R a relatively low number of

chickens hatched. This may be due to the 10s EID50 per egg of IB M41 -R used. This was 10-fold higher than the dose used in earlier studies in which there was a higher level of hatchability. The lower hatch percentages may also be caused by a particularly high susceptibility of the batch of SPF eggs for viruses, as in other studies the level of embryo mortality was also higher than had previously been observed.

After challenge all surviving chickens after hatch were completely protected against ciliostasis. It is concluded that IB M41 -R has great potential as vaccine to be administered in ovo. All publications mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described methods and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology, virology or related fields are intended to be within the scope of the following claims.