Part I
Positive sense RNA viruses
2

Coronavirus reverse genetics

Maria Armesto, Kirsten Bentley, Erica Bickerton, Sarah Keep and Paul Britton
Avian Viral Diseases, Institute for Animal Health, Compton Laboratory, Newbury, Berkshire

2.1 The Coronavirinae

The Coronaviridae form part of the order Nidovirales, which comprises two subfamilies, the Coronavirinae and Torovirinae. There are three genera of coronaviruses, alpha-, beta- and gammacoronaviruses (Carstens, 2010), which were so named for their visual resemblance to the corona of the sun in negatively stained preparations (Figure 2.1) (Tyrrell et al., 1968). Representative members of each of the coronavirus genera are shown in Table 2.1. Possibly the most publicised coronavirus of recent years has been the human coronavirus SARS-CoV, which emerged in China in 2002 causing the severe acute respiratory syndrome epidemic (Drosten et al., 2003; Ksiazek et al., 2003; Marra et al., 2003; Rota et al., 2003). Coronaviruses have, however, been isolated from many vertebrates and cause several economically important diseases in livestock species, including pigs, cows and chickens, and domestic species, such as dogs and cats.

Coronaviruses are enveloped viruses with a single-stranded positive-sense RNA genome of 26–32kb, the largest genomes of all RNA viruses currently known, that replicate in the cytoplasm of infected cells. The genome associates with the nucleoprotein (N), forming a helical nucleocapsid within the virus particles. Although common among negative-sense RNA viruses, coronaviruses are the only positive-sense RNA viruses to possess helical nucleocapsids, which are enclosed within lipid envelopes containing the spike (S) glycoprotein, membrane (M) protein and envelope (E) protein (Figure 2.1). For general reviews, see (Siddell, 1995; Lai and Cavanagh, 1997; Enjuanes, 2005; Siddell et al., 2005; Enjuanes et al., 2006; Gorbatenya et al., 2006; Masters, 2006; Thiel, 2007; Britton and Cavanagh, 2008; Perlmutter et al., 2008; Norkin, 2010).
2.2 Infectious bronchitis

The avian coronavirus infectious bronchitis virus (IBV) is the aetiological agent of the disease infectious bronchitis (IB) that affects poultry. IBV replicates primarily in the respiratory tract, causing the highly contagious respiratory disease IB in chickens characterised by nasal discharge, snicking, tracheal ciliostasis and rales (Britton and Cavanagh, 2007). Replication also occurs in other epithelial surfaces including enteric surfaces, oviducts and kidneys (Ambali and Jones, 1990; Cavanagh, 2005; Cavanagh and Gelb, 2008; Jones, 2010). Following an IBV infection, egg production and quality are impaired in layers, and weight gain in broilers is reduced (Cook and Mockett, 1995). Infected birds are predisposed to secondary bacterial infections such as colibacillosis and mortality in young chicks is not uncommon. Faecal excretion of the virus is a consequence of replication in the intestinal tract; however, this does not normally result in clinical disease.

Infectious bronchitis was first described in the US in the 1930s (Schalk and Hawn, 1931; Beach and Schalm, 1936; Beaudette and Hudson, 1937) and is prevalent in poultry farming across the world due to the intensive nature of poultry production, estimated to involve the global production of 55 billion chickens (50 billion broilers and 5 billion layers) on an annual basis. In a report, commissioned by Defra in 2005 (Defra, 2005), IBV was indicated as a major cause of ill health among chickens and was implicated as being responsible for more economic loss in the UK.
2.3 Coronavirus genome organisation

Table 2.1 Coronavirus genera and species.

<table>
<thead>
<tr>
<th>Genus</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>alphacoronavirus</td>
<td>Canine coronavirus (CCoV)</td>
</tr>
<tr>
<td></td>
<td>Feline coronavirus (FCoV)</td>
</tr>
<tr>
<td></td>
<td>Feline infectious peritonitis virus (FIPV)</td>
</tr>
<tr>
<td></td>
<td>Human coronavirus 229E (HCoV-229E)</td>
</tr>
<tr>
<td></td>
<td>Porcine epidemic diarrhoea virus (PEDV)</td>
</tr>
<tr>
<td></td>
<td>Porcine transmissible gastroenteritis virus (TGEV)</td>
</tr>
<tr>
<td>betacoronavirus</td>
<td>Bovine coronavirus (BCoV)</td>
</tr>
<tr>
<td></td>
<td>Human coronavirus HKU1 (HCoV-HKU1)</td>
</tr>
<tr>
<td></td>
<td>Human coronavirus OC43 (HCoV-OC43)</td>
</tr>
<tr>
<td></td>
<td>Human enteric coronavirus (HECoV)</td>
</tr>
<tr>
<td></td>
<td>Murine hepatitis virus (MHV)</td>
</tr>
<tr>
<td></td>
<td>Porcine haemagglutinating encephalomyelitis virus (HEV)</td>
</tr>
<tr>
<td></td>
<td>Rat coronavirus (RtCoV)</td>
</tr>
<tr>
<td></td>
<td>Severe acute respiratory syndrome coronavirus (SARS-CoV)</td>
</tr>
<tr>
<td>gammacoronavirus</td>
<td>IBV-like avian</td>
</tr>
<tr>
<td></td>
<td>Infectious bronchitis virus (IBV)</td>
</tr>
<tr>
<td></td>
<td>Turkey coronavirus (TCoV)</td>
</tr>
<tr>
<td></td>
<td>Pheasant coronavirus (PhCoV)</td>
</tr>
<tr>
<td></td>
<td>Non-IBV-like avian</td>
</tr>
<tr>
<td></td>
<td>Munia coronavirus (MunCoV)</td>
</tr>
<tr>
<td></td>
<td>Bulbul coronavirus (BuCoV)</td>
</tr>
<tr>
<td></td>
<td>Thrush coronavirus (ThCoV)</td>
</tr>
<tr>
<td></td>
<td>Mammalian</td>
</tr>
<tr>
<td></td>
<td>Beluga whale coronavirus SW1 (BeCoV)</td>
</tr>
<tr>
<td></td>
<td>Asian leopard cat coronavirus</td>
</tr>
<tr>
<td></td>
<td>Others*</td>
</tr>
<tr>
<td></td>
<td>Goose coronavirus</td>
</tr>
<tr>
<td></td>
<td>Pigeon coronavirus</td>
</tr>
<tr>
<td></td>
<td>Duck coronavirus</td>
</tr>
</tbody>
</table>

Note: *The derivation of these species of gammacoronavirus has yet to be determined, according to the International Committee on Taxonomy of Viruses, Index of Viruses (Coronaviridae, 2008).

poultry industry than any other disease (Bennett, 2003; Bennett and Jpelaar, 2005); IBV was estimated to cost the UK economy nearly £19 million per year, mainly due to loss of egg production, with serious implications for animal welfare. The cost of control through vaccination is approximately £5 million per year in the UK.

2.3 Coronavirus genome organisation

The genomic RNA has a 5' m7GpppN-cap and a 3' poly(A) tail with untranslated regions (UTRs) at the 5' and the 3' ends that have been shown to be involved in replication and translation (Senanayake and Brian, 1999), reviewed in (Brian and Baric, 2005; Van den Born and Snijder, 2008). The same general genome organisation is shared within the genus: 5' UTR – replicase gene – structural protein genes – UTR 3' (Figure 2.2). The 3'-end of the genome (~8kb) encodes the structural protein genes in the order S – E – M – N, with some betacoronaviruses also producing an haemagglutinin esterase (HE) protein, the gene of which is situated
upstream of the S gene. The replicase gene (gene 1) encompasses the 5’ most two-thirds of the entire genome and consists of two large open reading frames, ORF1a and 1b (Boursnell et al., 1987), reviewed in (Britton and Cavanagh, 2008; Ziebuhr, 2008). ORFs 1a and 1b overlap and the 1b sequence is translated as a result of a -1 frameshift mechanism (Brierley et al., 1987), the signal for which, consisting of a pseudoknot structure and a slippery sequence, lies in the overlapping region between ORF1a and 1b. In addition to the replicase gene and structural protein genes, coronavirus genomes also have several polycistronic genes encoding non-structural or accessory proteins, often referred to as group-specific genes. For example, IBV encodes four accessory proteins, 3a, 3b, 5a and 5b encoded by two polycistronic genes, 3 and 5 (Figure 2.2), the functions of which are as yet unknown.

2.4 The coronavirus replication cycle

The coronavirus replication cycle occurs in the cell cytoplasm as outlined in Figure 2.3, in which the S glycoprotein mediates attachment to host cell receptors and fusion of the virion membrane to the host cell membrane (Koch et al., 1990; Luo and Weiss, 1998). Binding of the S glycoprotein, via the receptor binding domain on the S1 subunit, to the host cell receptor induces conformational changes in
2.4 THE CORONAVIRUS REPLICATION CYCLE

Figure 2.3 The replication cycle of a coronavirus following infection of a susceptible cell. The virus particle attaches to the host cell receptor and fuses with the cell membrane via the S glycoprotein. Genomic RNA is released and acts as a mRNA for the translation of the replicase proteins. Virus-encoded proteinases proteolytically cleave the replicase polyproteins generating 15–16 products, which assemble into replication-transcription complexes (RTCs) associated with virus-induced double membrane vesicles (DMVs). Sub-genomic mRNAs are produced from the genomic RNA for the expression of the structural and accessory proteins at the rough ER. The membrane associated structural proteins assemble into virus particles at the ERGIC. The N protein associates with the genomic RNA to form the nucleocapsid and is incorporated into virions at the ERGIC. Virus particles bud off from the Golgi apparatus and exit the cell by exocytosis.

the S glycoprotein (Zelus et al., 2003; Tripet et al., 2004; Guo et al., 2009; Shulla and Gallagher, 2009), leading to virus–cell fusion and release of the nucleocapsid into the cytoplasm. Some coronaviruses such as SARS-CoV (Inoue et al., 2007) and MHV (Eifart et al., 2007) have been shown to utilise the clathrin-dependent endocytosis pathway for entry before being transported to early endosomes. Conformational changes in the S glycoprotein leading to virus–cell fusion may be pH-dependent, as in the cases of IBV (Chu et al., 2006) and SARS-CoV (Yang et al., 2004), or may be activated by proteases, as in the cases of HCoV-229E (Kawase et al., 2009) and SARS-CoV (Simmons et al., 2005; Matsuyama et al., 2010).
Following virus entry and uncoating, gene 1 of genomic RNA is directly translated into the two large replicase polyproteins, pp1a and pp1ab. Both polyproteins are proteolytically cleaved by two or three virus-encoded proteinases (Ziebuhr et al., 2000; Ziebuhr, 2008) and form replication-transcription complexes (RTC) on virus-induced double membrane vesicles (DMVs); autoproteolytic processing of pp1a and pp1ab polyproteins produces the 15 (IBV) or 16 (other coronaviruses) replicase non-structural proteins (nsp). The RTCs are then responsible for the replication and transcription of genomic and subgenomic mRNAs.

A model for coronavirus transcription and subsequent translation (Figure 2.4) has been described by Sawicki and Sawicki (Sawicki and Sawicki, 1995, 1998, 2005; Sawicki et al., 2007), reviewed in (Pasternak et al., 2006; Van den Born and Snijder, 2008). Minus strand RNA templates are synthesised from the genomic RNA; genome-length RNA copies are then produced by continuous transcription whereas subgenome-length RNAs are produced by a discontinuous transcription mechanism. A transcription regulation sequence (TRS), CTTAACAA for IBV, found upstream of each gene on the genomic RNA and at the leader junction site at the 5’ end of the genome is responsible for the generation of coronavirus subgenomic mRNAs. The generation of coronavirus subgenomic mRNAs starts with the synthesis of a negative-strand copy that initiates from the 3’ end of the genomic RNA and continues until a TRS is reached on the genomic RNA. The RTC either pauses and then continues on to the next TRS or translocates to the TRS comprising the leader junction sequence at the 5’ end of the genome and results in the discontinuous addition of a negative-sense copy of the leader RNA sequence at the 3’ end. The overall process results in a series of negative-sense copies of the sub-genomic RNAs with an anti-leader sequence at the 3’ end. The negative sense genome-length and sub-genomic RNAs are used as templates for synthesis of genomic RNA and a nested set of sub-genomic mRNAs, in which each mRNA has the same 3’ terminus and short 5’ leader sequence, identical to the 5’ end of the genome (Lai et al., 1983). Most sub-genomic mRNAs are structurally polycistronic but functionally monocistronic in which only the ORF at the 5’ end is translated by a cap-dependent mechanism. However, some subgenomic mRNAs are functionally bi- or tricistronic and are subsequently translated via a leaky-scanning mechanism or by internal ribosome entry (Liu and Inglis, 1991, 1992; Le et al., 1994). Newly synthesised viral RNA is found associated with convoluted membranes and DMVs (Gosert et al., 2002), reviewed in (Baker and Denison, 2008), that are thought to originate from the endoplasmic reticulum (ER) (Knoops et al., 2008), although autophagy may also be involved (Prentice et al., 2004). These membranes may serve to protect viral RNA from degradation or provide an optimal environment for viral RNA synthesis (van Hemert et al., 2008).

The nucleocapsid associates with the M protein (Sturman et al., 1980) and structural proteins assemble at the ER-Golgi intermediate compartment (ERGIC) (Klumperman et al., 1994). Complete virus particles bud off from the Golgi apparatus and exit the cell by exocytosis (Tooze et al., 1987).
2.5 DEVELOPMENT OF REVERSE GENETICS SYSTEM

Figure 2.4  Coronavirus replication and transcription. Coronavirus-derived replicase proteins within the RTCs recognise cis-acting elements at the 5' and 3' ends of the genomic RNA and copy the genome into either a genome-length negative-strand template or generate sub-genomic negative-strand templates by a discontinuous process. Negative strands are shown in light grey and are used as templates for genomic and sub-genomic mRNA synthesis, generating a 3'-coterminal nested set of sub-genomic mRNAs. Anti-leader sequences are also shown in light grey. The RTCs age, releasing the minus strand templates for degradation.

Source: Adapted from (Sawicki et al., 2007).

2.5 Development of reverse genetics system for coronaviruses including IBV

Coronaviruses have a single-stranded, non-segmented positive sense RNA genome, requiring the generation of a cDNA that can function as a template for the generation of infectious RNA. The initial stage for a coronavirus-based reverse genetics system involves conversion of the RNA genome into an authentic cDNA that can be
CH02  CORONAVIRUS REVERSE GENETICS

manipulated using standard DNA technologies or utilising homologous recombination. The final stage of the process requires the generation of an infectious RNA from the modified cDNA utilising a DNA-dependent RNA polymerase. Viruses, including coronaviruses, with a positive-sense single-stranded RNA genome have the advantage that the infectious RNA derived from a cDNA copy, like the genomic RNA, can be recognised by a host cell’s transcriptional machinery as an mRNA, resulting in the synthesis of the protein(s) required for replication of the RNA genome, in the case of coronaviruses, this involves 15–16 distinct proteins. Historically, the development of the first reverse genetics system for a single-stranded RNA virus recovered from a cDNA was for the bacteriophage Qb (4.5kb) (Taniguchi et al., 1978). This early success was followed by the recovery of viruses from cDNAs generated from RNA viruses with increasing size as outlined by (Racaniello and Baltimore, 1981; Rice et al., 1987; Liljestrom et al., 1991).

The first reverse genetics system for coronaviruses was developed during the 1990s using targeted RNA recombination rather than recovery of a virus from a full-length cDNA of the virus genome, reviewed in (Masters, 1999; Masters and Rottier, 2005). This technology allowed the modification of a coronavirus genome utilising a recombination event between either a non-replicating or replicating RNA, generated from a modifiable cDNA, introduced into the same cell as the replicating coronavirus genome and a selective marker to differentiate recombinant viruses. Targeted RNA recombination was devised as a method of modifying a coronavirus genome as it was unclear at that time whether the construction of a full-length cDNA and subsequent generation of an infectious RNA were possible for an RNA virus with such a large genome size. The method was originally based on a temperature sensitive (ts) lesion within the N protein gene of MHV (Koetzner et al., 1992) and later utilised selection via retargeting a recombinant coronavirus, by the use of heterologous S glycoprotein, to different cell types (Kuo et al., 2000). For example, modifications were made to MHV by producing a virus, fMHV, which expressed the ectodomain of the S glycoprotein from FIPV allowing for selection on feline cells, modifications were made to the MHV genome with concomitant replacement of the FIPV S glycoprotein with the MHV S glycoprotein allowing selection of the recombinant MHV (rMHV) on murine cells (Kuo et al., 2000). The technology is still a useful tool for specifically modifying some coronavirus genomes; however, the main disadvantage is that it is difficult to modify the replicase region of the genome.

Following the use of targeted recombination to modify a coronavirus genome, cDNAs capable of generating infectious RNAs were being produced for increasingly larger RNA genomes ranging from 15 kb (arteriviruses) (van Dinten et al., 1997) to 20 kb (citrus tristeza virus of the genus Closterovirus) (Satyanarayana et al., 2001). However, despite these successes, it was found that the size of the RNA virus genome was not the main constraint on generating a successful reverse genetics system. The instability of some virus-derived cDNAs in bacteria was identified as a problem that required ingenious strategies for the assembly of full-length cDNAs for generating infectious RNAs. For example, in vitro ligation, without
2.5 DEVELOPMENT OF REVERSE GENETICS SYSTEM

assembly of the full-length cDNA in bacteria, was used to develop an infectious clone system for yellow fever virus (Rice et al., 1989); construction of full-length cDNAs in yeast (Polo et al., 1997), or the introduction of short introns to allow propagation of cDNAs in Escherichia coli (E. coli) (Yamshchikov et al., 2001) have proven successful strategies for the generation of full-length, stable genomic cDNAs of dengue and Japanese encephalitis viruses.

Several groups, including our own, discovered that certain regions of the coronavirus replicase gene proved to be highly unstable in E. coli, therefore preventing the assembly of a full-length coronavirus-derived cDNA. The breakthrough in the development of the first coronavirus reverse genetics system based on a full-length cDNA for generating infectious RNA was reported in 2000 for the porcine coronavirus TGEV (Almazán et al., 2000), reviewed in (Enjuanes et al., 2005; Deming and Baric, 2008). The TGEV full-length cDNA was assembled in a bacterial artificial chromosome (BAC), immediately downstream of a cytomegalovirus (CMV) RNA polymerase II promoter for subsequent generation of infectious RNA. The BAC system was chosen due to the presence of only a single copy per bacterial cell and because it allowed the introduction of the unstable region as a final step. Construction of the TGEV cDNA was initiated from a cDNA representing a defective RNA (D-RNA) that could be rescued by helper TGEV, thus indicating that all the RNA sequences required for replication were present on the D-RNA (Izeta et al., 1999). The authors then sequentially introduced the TGEV sequence absent from the D-RNA to create a full-length cDNA. During this process a sequence, corresponding to part of the replicase sequence, was found to be unstable in E. coli to such an extent that an intact cDNA could not be maintained in the bacteria. The Enjuanes group successfully produced the TGEV cDNA by initially generating a cDNA lacking the sequence that gave rise to instability in E. coli, under the control of the CMV promoter in a BAC. The unstable sequence was introduced into the cDNA as a final cloning step, resulting in a relatively stable full-length cDNA that could be amplified in E. coli. Transfection of the TGEV-BAC construct into susceptible cells resulted in the synthesis of infectious RNA in the nucleus by cellular RNA polymerase II and subsequent amplification in the cytoplasm by virus-encoded enzymes for the recovery of infectious recombinant virus, reviewed in (Enjuanes et al., 2005).

A second reverse genetics system was reported for TGEV involving the in vitro assembly of a full-length cDNA using a series of contiguous cDNAs containing engineered unique restriction sites, dispensing with the requirement for E. coli (Yount et al., 2000). Infectious RNA was produced in vitro using bacteriophage T7-RNA polymerase, utilising a T7-RNA polymerase promoter immediately upstream of the 5’ end of the TGEV cDNA, and electroporated into susceptible cells for the rescue of infectious virus. The authors found that this system required TGEV N protein for the recovery of infectious virus.

A third coronavirus reverse genetics system utilising vaccinia virus (VV) as the vector for the full-length cDNA was reported for the recovery of HCoV 229E (Thiel et al., 2001) and the avian coronavirus IBV (Casais et al., 2001). In both systems,
Figure 2.5  Schematic diagram of the IBV Beau-R full-length cDNA inserted into the vaccinia virus TK gene. The IBV cDNA, representing a full-length copy of the IBV Beau-R genomic RNA (Casais et al., 2001), is shown inserted within a NotI restriction site within the TK gene of VV vNotI/tk (Merchlinsky and Moss, 1992). The IBV cDNA is shown in a 3′–5′ direction as the VV DNA is in the 5′–3′ orientation. The IBV genes are indicated as are the positions of the T7 promoter, the H5R and T7 termination sequences in relation to the IBV cDNA.

Sequential cDNA fragments corresponding to the two genomes were generated and systematically ligated together in vitro before direct cloning into the genome of VV vNotI/tk, via a NotI site introduced into the thymidine kinase (TK) gene of vNotI/tk (Merchlinsky and Moss, 1992). This resulted in a full-length cDNA under the control of a T7 RNA polymerase promoter with a hepatitis δ ribozyme (H5R) placed downstream of the coronavirus poly(A) tail followed by a T7 termination sequence (Figure 2.5), reviewed in (Thiel and Siddell, 2005). Infectious RNA can be generated in vitro from VV templates using T7 RNA polymerase and transfected into permissive cells for the recovery of infectious virus (Thiel et al., 2001). Alternatively, infectious RNA can be generated in situ in which VV DNA is transfected into cells infected with a recombinant fowlpox virus, rFPV-T7 expressing T7 RNA polymerase (Britton et al., 1996), for the recovery of infectious virus (Casais et al., 2001). We found that the IBV N protein is an absolute requirement for the recovery of IBV using primary chick kidney (CK) cells. The requirement of an N protein for the recovery of other coronaviruses is not an absolute requirement, however, recovery is significantly enhanced by the presence of the appropriate N protein (Yount et al., 2000; Yount et al., 2002; Yount et al., 2003; Almazan et al., 2004; Schelle et al., 2005; Schelle et al., 2006; Coley et al., 2005). A possible explanation for this observed enhancement comes from some recent studies in which an interaction between the MHV nsp3 replicase protein and the N protein was found to be critical for replication (Hurst et al., 2010).

Reverse genetics systems for several coronaviruses, belonging to all three genera, have been developed and successfully used to recover infectious viruses (Table 2.2). We will, therefore, describe in more detail how we have used our IBV reverse genetics system, as an example, to modify a coronavirus genome. The use of VV as a vector for a full-length coronavirus cDNA offers a highly stable system for producing and maintaining an authentic cDNA, dispensing with the need for repeated
Table 2.2  Reverse genetics systems for the recovery of infectious coronaviruses.

<table>
<thead>
<tr>
<th>Virus</th>
<th>Genus</th>
<th>System</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGEV</td>
<td>alphacoronavirus</td>
<td>BAC</td>
<td>Almazán et al., 2000</td>
</tr>
<tr>
<td>TGEV</td>
<td>alphacoronavirus</td>
<td>In vitro ligation</td>
<td>Yount et al., 2000</td>
</tr>
<tr>
<td>HCoV 229E</td>
<td>alphacoronavirus</td>
<td>Vaccinia virus</td>
<td>Thiel et al., 2001</td>
</tr>
<tr>
<td>mFIPV</td>
<td>alphacoronavirus</td>
<td>Targeted recombination</td>
<td>Haijema et al., 2003</td>
</tr>
<tr>
<td>FCoV</td>
<td>alphacoronavirus</td>
<td>Vaccinia virus</td>
<td>Tekes et al., 2008</td>
</tr>
<tr>
<td>HCoV NL63</td>
<td>alphacoronavirus</td>
<td>In vitro ligation</td>
<td>Donaldson et al., 2008</td>
</tr>
<tr>
<td>MHV</td>
<td>betacoronavirus</td>
<td>Targeted recombination</td>
<td>Koetzner et al., 1992</td>
</tr>
<tr>
<td>fMHV</td>
<td>betacoronavirus</td>
<td>Targeted recombination</td>
<td>Kuo et al., 2000</td>
</tr>
<tr>
<td>MHV</td>
<td>betacoronavirus</td>
<td>In vitro ligation</td>
<td>Yount et al., 2002</td>
</tr>
<tr>
<td>MHV</td>
<td>betacoronavirus</td>
<td>Vaccinia virus</td>
<td>Coley et al., 2005</td>
</tr>
<tr>
<td>HCoV OC43</td>
<td>betacoronavirus</td>
<td>BAC</td>
<td>St-Jean et al., 2006</td>
</tr>
<tr>
<td>SARS-CoV</td>
<td>betacoronavirus</td>
<td>BAC</td>
<td>Almazan et al., 2006</td>
</tr>
<tr>
<td>SARS-CoV</td>
<td>betacoronavirus</td>
<td>In vitro ligation</td>
<td>Yount et al., 2003</td>
</tr>
<tr>
<td>Bat-S-CoV*</td>
<td>betacoronavirus</td>
<td>In vitro ligation</td>
<td>Becker et al., 2008</td>
</tr>
<tr>
<td>IBV</td>
<td>gammacoronavirus</td>
<td>Vaccinia virus</td>
<td>Casais et al., 2001</td>
</tr>
<tr>
<td>IBV</td>
<td>gammacoronavirus</td>
<td>In vitro ligation</td>
<td>Yount et al., 2005,</td>
</tr>
</tbody>
</table>

Note: *Bat-S-CoV = bat SARS-like coronavirus in which the sequence was synthesised from a consensus Bat-S-CoV genome where the Bat-S-CoV S glycoprotein receptor binding domain (RBD) was replaced with the SARS-CoV RBD for rescue of infectious recombinant virus in Vero cells (Becker et al., 2008).

cloning of cDNA fragments. A major advantage of the VV-based system is that the coronavirus cDNA can be modified or replaced using homologous recombination; the transient dominant selection (TDS) system that we use (Britton et al., 2005) is described below. The resultant rIBVs, apart from the introduced modification, are isogenic as they are derived from the same cDNA sequence.

2.6 Reverse genetics system for IBV

A complete cDNA copy of the IBV Beaudette genome was assembled and inserted into the TK gene of VV (Casais et al., 2001). The IBV cDNA is under the control of a T7 promoter and has a H8R sequence placed downstream of the coronavirus poly(A) tail followed by a T7 termination sequence (Figure 2.5). IBV infectious RNA is generated from the T7 promoter immediately adjacent to the 5’ end of the IBV cDNA using T7 RNA polymerase and terminates at the T7 termination sequence downstream of the H8R sequence, which autocleaves itself and the T7-termination sequence at the end of the poly(A) sequence, resulting in an authentic IBV genomic RNA copy.
2.6.1 Transient dominant selection for modification of the IBV genome

The VV-based TDS recombination method (Falkner and Moss, 1990) we use to modify our IBV cDNA sequence (Britton et al., 2005; Armesto et al., 2008) is outlined in Figure 2.6. The TDS method for modifying sequences in VV and related poxviruses involves the use of a plasmid containing a selective marker gene, *E. coli* guanine phosphoribosyltransferase (GPT); (Mulligan and Berg, 1981; Falkner and Moss, 1988), as well as the IBV sequence to be modified, the donor sequence. A recombinant VV (rVV), containing either an IBV full-length cDNA or an IBV cDNA, in which part of the IBV genome has been deleted, is used as a receiver sequence for modification. As a result of a single-step homologous recombination event, between the donor IBV cDNA sequence in the GPT plasmid and the IBV receiver sequence in the VV genome, the complete plasmid sequence is inserted into the genome of the receiver VV. Resultant rVVs express the selective GPT protein and can be identified by plaque purification in the presence of selection medium containing mycophenolic acid (MPA), xanthine and hypoxanthine. MPA is an inhibitor of purine biosynthesis, therefore only those viruses carrying the GPT gene (Figure 2.7), which provides an alternative pathway for purine biosynthesis, are able to replicate in the presence of MPA and the alternative purine precursors xanthine and hypoxanthine (Falkner and Moss, 1988). Recombinant VVs that are phenotypically GPT⁺ are then plaque purified in the absence of selection medium. The removal of the selection agent results in a second recombination event between repeat sequences in the VV genome causing the loss of the GPT gene (Figure 2.6). This second recombination step results in two possible outcomes; one event will result in the original (unmodified) IBV sequence and the other in the generation of an IBV cDNA containing the desired modification (Figure 2.6). Once the desired IBV cDNA sequence within the VV genome has been identified and confirmed, stocks of this virus are grown in BHK-21 cells for isolation of VV DNA.

2.6.2 Rescue of recombinant IBVs

The rescue of infectious rIBVs is carried out in primary CK cells using the VV DNA containing the modified IBV sequence and a helper rFPV-T7 (Britton et al., 1996) to generate the infectious IBV RNA. The rFPV-T7 infected CK cells are transfected with the VV DNA and a plasmid expressing the IBV N protein (Hiscox et al., 2001), under the control of both the CMV promoter and the T7 RNA promoter. As indicated in Section 2.6.1, we found that the presence of the IBV N protein with the IBV infectious RNA is essential for recovery of infectious virus (Casais et al., 2001). Cell supernatants are filtered to remove any rFPV-T7 (Evans et al., 2000) and potential rIBVs are passaged three times in CK cells to produce stocks of virus for sequence analysis to confirm the presence of the modified IBV sequence.
2.6 REVERSE GENETICS SYSTEM FOR IBV

Figure 2.6  Schematic diagram of the transient dominant selection process for modifying Beau-R cDNA within a vaccinia virus genome. The figure shows the general TDS process for modifying IBV cDNA using a GPT selection plasmid with the modified IBV donor sequence; in this case the addition of IBV structural and accessory genes. The receiver IBV cDNA is within the VV genome and in this example shows an IBV cDNA lacking sequence from the start of the S gene to the end of the N gene. An intermediate rVV is generated, in which the complete donor plasmid DNA sequence is integrated into the truncated IBV cDNA, by a single-step homologous recombination event via a replicase sequence common to both sequences. The rVV has a GPT+ phenotype allowing selection in the presence of MPA. Removal of MPA can result in two types of spontaneous intramolecular recombination events due to the presence of tandem repeat sequences of the IBV cDNA, resulting in either generation of rVVs with a truncated IBV cDNA (no modification) similar to the receiver sequence or a complete full-length IBV cDNA, the desired end product. Both recombination events result in the loss of the GPT gene. The IBV genes representing the structural and accessory genes are shown; a potential recombination event is indicated between the IBV replicase gene sequence common to both constructs.
Figure 2.7 Schematic diagram showing the effect of MPA on purine metabolism. MPA inhibits the enzyme inosine monophosphate dehydrogenase, preventing the formation of xanthine monophosphate required for the generation of guanosine monophosphate, a precursor for nucleic acid biosynthesis, resulting in the intracellular depletion of purine nucleotides and inhibition of cell growth. The inhibition of the *de novo* synthesis of the purines by MPA can be overcome by alternative enzymes such as GPT, which is able to convert the substrates xanthine and hypoxanthine into guanosine monophosphate. Therefore, rVV expressing GPT can grow in the presence of MPA.

Primary CK cells are refractory for growth of most IBV isolates; therefore, rIBVs expressing S glycoproteins from such isolates cannot be recovered using CK cells. In order to recover such rIBVs, the supernatants from the transfected CK cells are used to infect 10-day-old embryonated hen’s eggs. Allantoic fluid is collected and any potential virus passed a further three times in 10-day-old embryos. RNA is extracted from the allantoic fluid of infected eggs and RT-PCR followed by sequencing is used to confirm the identity of the rIBV.

### 2.7 Reverse genetics systems for the modification of coronavirus genomes

This section describes how reverse genetics systems have been used to study the molecular biology of coronaviruses with respect to interactions and functions of the
2.7 REVERSE GENETICS SYSTEMS FOR THE MODIFICATION OF CORONAVIRUS GENOMES

replicase, structural and accessory proteins and to determine whether they play an essential role in the coronavirus replication cycle.

2.7.1 Modifications to coronavirus structural genes

The coronavirus M glycoprotein

The most abundant coronavirus structural protein is the 25 kDa M glycoprotein, which spans the viral envelope three times (Godeke et al., 2000; Hogue and Machamer, 2008). The first membrane-spanning domain targets the M protein to the cis Golgi (Machamer et al., 1990) and has been shown to be sufficient for membrane binding, retention in the Golgi and formation of multimers (Tseng et al., 2010). The M protein has a short amino-terminal ectodomain and a large carboxy-terminal cytoplasmic domain that interacts with the E protein (Corse and Machamer, 2003) and the nucleocapsid and is involved in assembly of virus particles (Narayanan et al., 2003; Hogue and Machamer, 2008). Modification of the MHV M protein gene was achieved using fMHV targeted RNA recombination to map the interactions of the M and N proteins of MHV (Kuo and Masters, 2002). Interaction of the M protein with the S glycoprotein retains the S glycoprotein in the ERGIC (McBride et al., 2007) and a single amino acid residue in the SARS-CoV M protein has been identified as being necessary for interaction of the S and M proteins and assembly of virions (McBride and Machamer, 2010). The alpha- and gammacoronaviruses have M proteins glycosylated with N-linked sugars whereas betacoronaviruses such as MHV have M proteins that are O-glycosylated. Targeted RNA recombination was used to modify the MHV M protein so that it was either N-glycosylated or not glycosylated (de Haan et al., 2003a). Glycosylation of the M protein is not required for virus assembly, indicating that glycosylation of the M protein is involved in a virus–host interaction. Infection of cells in culture showed that the glycosylation status of the MHV M protein did not influence the growth kinetics of the viruses. However, rMHVs with an N-glycosylated M protein induced type I interferon (IFN) to a higher level when compared to viruses expressing M proteins either lacking glycosylation or with O-linked sugars. In vivo studies showed that the rMHVs differed in their ability to replicate in the livers but not in the brains of infected mice (de Haan et al., 2003a).

The coronavirus N protein

The N protein is a 50 kDa phosphorylated, highly basic structural protein that forms a helical nucleocapsid when bound to the coronavirus RNA genome within virus particles (Hogue and Machamer, 2008). The carboxy-terminal domain interacts with the M protein (Hurst et al., 2005) and is packaged into viral particles by the M protein (Narayanan et al., 2003; Hogue and Machamer, 2008). Targeted RNA recombination using fMHV identified that the carboxy termini of the MHV N protein was involved in the interaction of the M and N proteins (Kuo and Masters, 2002).
The coronavirus E protein

The 10 kDa E protein is an integral membrane protein and a minor component of the virus envelope containing a single hydrophobic domain flanked by two hydrophilic domains. Two different topologies for the orientation of the E protein in the virus membrane have been proposed with either one or two transmembrane (TM) domains (Hogue and Machamer, 2008). Recent studies have shown that the SARS-CoV E protein has a topological conformation indicative of a single TM domain with the amino-terminus orientated towards the lumen of intracellular membranes and the carboxy-terminus facing the cell cytoplasm (Nieto-Torres et al., 2011). The cytoplasmic tail of the E protein contains Golgi targeting motifs (Corse and Machamer, 2003). It is thought that interaction of the E and M proteins may cause membrane curvature (Fischer et al., 1998), promoting budding of virus particles indicating that the E protein plays a fundamental role in the generation of virus particles (Lim and Liu, 2001). The E protein forms cation-selective ion channels in the lipid envelope, enhancing membrane permeability (Wilson et al., 2004). A rTGEV lacking the E gene has been shown to require helper E protein for recovery of infectious TGEV, indicating that the E protein is essential for TGEV replication (Ortego et al., 2002). In contrast, recovery of a rMHV with a deleted E gene showed that the E protein is not essential for MHV replication (Kuo and Masters, 2003); however, the rMHV grew less efficiently. The rMHV acquired variant M genes encoding M proteins with truncated endodomains, which enhanced virus growth and were incorporated into virions, indicating a role for E in mediating interactions between TM domains of M monomers (Kuo and Masters, 2010). The MHV E gene was replaced by heterologous E genes from other coronaviruses (Kuo et al., 2007). Despite extensive sequence variability of the coronavirus E proteins and possible structural differences, rMHVs expressing the BCoV, SARS-CoV or IBV E proteins functioned as wild-type viruses, suggesting that the role of the E protein is not dependent on sequence-specific interactions with an M protein. Substitution of the MHV E protein with that of TGEV, however, required compensatory mutations.

The coronavirus S glycoprotein

The 180 kDa S glycoprotein is a type I membrane protein projecting 20 nm from the virus surface (Figure 2.8) (Delmas and Laude, 1990; Cavanagh, 1995; Hogue and Machamer, 2008) and is composed of two subunits; the amino-terminal S1 subunit responsible for binding to host cell receptors (Koch et al., 1990) and the carboxy-terminal S2 subunit responsible for cell–cell and virus–cell fusion (Luo and Weiss, 1998). The cell tropism of coronaviruses has been shown to be determined by the S glycoprotein (Kuo et al., 2000; Casais et al., 2003; Haijema et al., 2003).

Replacement of the MHV S glycoprotein ectodomain with that of FIPV resulted in a recombinant virus, fMHV, with the tropism of FIPV rather than MHV (Kuo et al., 2000). This virus has subsequently been used to select additional recombinant viruses that have regained the ability to grow in murine cells, confirming that
2.7 REVERSE GENETICS SYSTEMS FOR THE MODIFICATION OF CORONAVIRUS GENOMES

Figure 2.8  IBV S glycoprotein structure. The IBV S glycoprotein is a class I fusion membrane protein comprising the amino-terminal S1 subunit and the carboxy-terminal S2 subunit. A signal sequence (SS) located at the amino-terminal end of S1 is cleaved during synthesis; the IBV RBD is located within the S1 sequence. The S2 subunit contains two heptad repeats (HR1 and HR2), a TM domain and a cytoplasmic tail (Cyto). The IBV S glycoprotein is cleaved by host cell proteases at a highly basic region between the S1 and S2 subunits.

cell tropism is determined by the S glycoprotein. This approach was also utilised to create the reverse virus, mFIPV, with the S glycoprotein ectodomain of MHV expressed by FIPV, which gained the ability to grow on murine cells and lost the ability to grow on feline cells (Haijema et al., 2003). The FIPV reverse genetics system based on mFIPV gave rise to potential live attenuated vaccines against FIPV (Haijema et al., 2004).

We used our IBV reverse genetics system to investigate the role of the S glycoprotein, the tropism and growth characteristics of different IBV strains and to develop potential vaccines that can be propagated in cell lines rather than embryonated chicken eggs. Casais et al. (Casais et al., 2003) showed that the cellular tropism of IBV Beaudette is conferred by the S glycoprotein by replacing the ectodomain of the Beaudette S glycoprotein with that of M41 within the background of Beau-R, creating BeauR-M41(S). IBV Beaudette is not only able to replicate in primary CK cells but is also able to replicate in Vero and BHK-21 cell lines, an African green monkey kidney and baby hamster kidney cell line respectively; in contrast, the IBV M41 strain is only able to replicate in primary CK cells. BeauR-M41(S) exhibited the in vitro cellular tropism of M41, demonstrating that the S glycoprotein is a determinant of host range in IBV. The rIBV did not replicate as well as M41 in the trachea and nose of infected birds and was apathogenic like Beau-R (Hodgson et al., 2004). However, chickens vaccinated with BeauR-M41(S) were protected against M41 whereas vaccination with Beaudette did not induce protection (Hodgson et al., 2004).
Further rIBVs were created in the background of Beaudette with different S genes. Two recombinants, BeauR-4/91(S) and BeauR-B1648(S), express the S glycoprotein ectodomain from either the UK/4/91 strain or the BE/B1648/87 strain respectively, which are pathogenic strains belonging to different IBV serogroups to Beaudette and M41. As with BeauR-M41(S), the S gene of 4/91 did not confer pathogenicity to BeauR-4/91(S) although the resultant virus had the tropism and \textit{in vitro} growth characteristics of the donor strain 4/91. Chickens vaccinated with BeauR-4/91(S) were protected from challenge with 4/91 (Armesto \textit{et al.}, 2011).

The potential for random recombination events within the IBV S glycoprotein was explored using the TDS system in which the S gene from the BE/B1648/87 nephropathogenic strain of IBV was recombined with the Beaudette S gene in the full-length IBV cDNA (Izadkhasti, 2006). Random recombination events occurred at different cross-over points between the two S genes. An rIBV possessing an S gene with the HR1 domain from BE/B1648/87 and the HR2 domain from Beaudette was recovered, whereas viruses containing chimaeric HR1 sequences were not viable, indicating the importance of the HR1 domain in the functionality of the spike. None of the resulting viruses were able to grow on CK or Vero cells, as with BE/B1648/87, although all possessed similar growth characteristics to the parental BE/B1648/87 in embryonated eggs. Viruses resulting from recombination within the S gene were shown to be viable, supporting the theory that IBV evolution is linked with recombination events (Cavanagh \textit{et al.}, 2007).

To investigate which subunit of the IBV S glycoprotein is involved in conferring the ability of Beaudette to grow on Vero cells, two rIBVs with the genomic background of Beaudette with either the S1 or S2 subunit from M41 were generated; BeauR-M41(S1) has the background of Beau-R with the S1 subunit from M41, BeauR-M41(S2) has the S2 subunit from M41. Both rIBVs replicated to similar titres as the parent viruses on CK cells. Growth of BeauR-M41(S2) on Vero cells was poor like M41, in contrast, BeauR-M41(S1) was found to replicate on Vero cells with similar growth kinetics as Beau-R albeit at a lower titre (Figure 2.9) (Bickerton, 2010). However, after seven passages on Vero cells, BeauR-M41(S1) replicated to a similar titre as observed for Beau-R (Figure 2.9). As the S1 subunit is responsible for receptor binding, it was anticipated that this subunit would be responsible for the ability of Beau-R to replicate in Vero and BHK-21 cells rather than the S2 subunit, which is involved in virus-to-cell and cell-to-cell fusion. The extended host range of a variant of MHV has previously been mapped to the S1 subunit (Schickli \textit{et al.}, 2004). Surprisingly, we found that the ability of IBV Beaudette to infect and replicate in Vero and BHK-21 cells was conferred by the S2 subunit, not the S1 subunit (Bickerton, 2010) that contains the receptor binding domain (RBD).

\subsection*{2.7.2 Modification of coronavirus accessory protein genes}

All coronaviruses contain group-specific genes located in the 3′ end of the genome interspersed among the structural protein genes. These group-specific genes vary in
2.7 REVERSE GENETICS SYSTEMS FOR THE MODIFICATION OF CORONAVIRUS GENOMES

Number and location between different coronaviruses and express proteins with potentially diverse functions. Reverse genetics systems have enabled manipulation of these genes to elucidate their role(s) in replication. IBV contains two group-specific genes, 3 and 5, encoding accessory proteins 3a and 3b, and 5a and 5b, respectively (Figure 2.2). By generating a series of rIBVs, Hodgson et al. (Hodgson et al., 2006) were able to show that the 3a and 3b proteins were dispensable for replication of IBV. Through mutation of the translation initiation codons of one or both proteins, it was demonstrated that, despite the loss of expression of these proteins, the growth kinetics of the viruses remained similar to that of wild-type parental virus when grown in primary CK cells or 11-day-old embryonated eggs (Hodgson et al., 2006).

A similar approach was taken by Casais et al. (Casais et al., 2005) to investigate the requirement of IBV gene 5. Recombinant IBVs were generated in which either the 5a or 5b, or both 5a and 5b, translation initiation codons were modified to prevent expression of the proteins. Further recombinants were generated in which either the first, or both, of the IBV Beaudette gene 5 TRSs were scrambled to prevent transcription of the gene 5 subgenomic mRNA. As with gene 3, each of the recombinant viruses showed growth kinetics similar to that of wild-type parental both in vitro in CK cells, in ovo in 11-day-old embryonated eggs, and also ex vivo in tracheal organ cultures (Casais et al., 2005). In an alternative approach Youn et al. (Youn et al., 2005) also demonstrated that the IBV 5a protein was not essential for replication by replacing the 5a sequence with that of enhanced green fluorescent protein (EGFP). A rIBV that expressed EGFP instead of the 5a protein was successfully rescued and grew on Vero cells, albeit to a 10-fold lower titre compared with the wild-type

Figure 2.9 Growth kinetics of rIBVs on Vero cells. Vero cells were infected with (a) Beau-R, M41-CK, BeauR-M41(S), BeauR-M41(S1) and BeauR-M41(S2); and (b) Beau-R and BeauR-M41(S1) P7-Vero. Supernatants were harvested at 1, 12, 24, 48, 72 and 96 hours post-infection and titrated on CK cells. Three replicates of each growth curve were performed and the averages taken. Error bars indicate standard error of the mean.
virus. These results indicate that the proteins from IBV genes 3 and 5 are not essential for replication in vitro but may have a role in vivo and are therefore classified as accessory proteins.

MHV contains four group-specific genes, 2a and HE located between ORF 1b and the S gene, 4 and 5a located between the S and E genes (Figure 2.2). Not all of these genes are expressed in all strains of MHV. Targeted RNA recombination was used to generate a series of rMHVs with group-specific gene deletions to investigate their requirements (de Haan et al., 2002). The rMHVs had either genes 2a and HE or genes 4 and 5a deleted or all four genes deleted. In vitro characterisation of the rMHVs showed they had similar growth kinetics to wild-type parental virus, although deletion of genes 4 and 5a alone, or in combination with the deletion of gene 2a and HE, resulted in approximately 10-fold lower titres than wild-type virus or the gene 2a/HE deletion rMHV. However, inoculation of mice with each of the rMHVs identified that they were highly attenuated, concluding that the MHV group-specific genes, although not essential in vitro, have a role to play in virulence in vivo.

TGEV contains two group-specific genes, 3a/b and gene 7. Ortego et al. (Ortego et al., 2003) investigated whether gene 7 is dispensable for TGEV replication. The growth kinetics of the rTGEV, in which gene 7 had been deleted, were similar to those shown by the parental virus, viral titres were also similar at 24 hours post-infection indicating that TGEV gene 7 is not essential for replication in cell culture. When used in vivo to infect piglets, the rTGEV showed approximately 100-fold reduction in virus titres compared to parental virus with an increase in the survival of piglets. This again highlights the potential role for the group-specific genes in virus virulence within the natural host.

SARS-CoV encodes a large number of group-specific genes. As with other coronaviruses, these genes, 3a/b, 6, 7a/b, 8a/b, and 9b, are located among the structural genes and the functions of most are still unknown. Yount et al. (Yount et al., 2005) systematically deleted five of the eight group-specific genes and assessed the replication and gene expression of the resulting recombinants in vitro and in vivo in mice. Deletion of the SARS-CoV group-specific genes individually, or in combination, resulted in little effect on the replication of the virus in vitro, supporting the previous results that coronavirus group-specific genes are not essential for growth in cell culture. Freundt et al. (Freundt et al., 2010) showed that the deletion of gene 3a resulted in a reduction of cell death at 48 hours post-infection when compared to the wild-type virus. Gene 3a was also found to be necessary for the formation of intracellular vesicles, a prominent feature of cells from SARS-CoV-infected patients, and fragmentation of the Golgi apparatus. Again this implies that the group-specific proteins have a role to play in vivo that may be difficult to determine in cell culture alone.

Although the group-specific genes are not conserved across the different coronaviruses, it seems clear that they are dispensable for replication in cell culture but are often implicated in roles specific to the host, particularly with regards to virulence and pathogenesis. This non-essential nature of these genes has led to research
to investigate whether the deletion, or replacement, of these genes can result in attenuated vaccines or in the use of coronaviruses as gene delivery systems.

### 2.7.3 Modification of the coronavirus replicase gene

The replicase gene is over 20kb in size and the availability of coronavirus reverse genetics systems has allowed the modification of several replicase nsps to investigate their role in the virus life cycle, in pathogenicity, host–virus interactions and to identify sequence changes involved in attenuation.

Nsp1 of MHV and SARS-CoV has been implicated in reducing cellular gene expression as a result of mRNA degradation (Kamitani et al., 2006; Zust et al., 2007). Deletion of the carboxy-terminal part of MHV nsp1, K_{124}L_{241}, was found to be dispensable for replication in cell culture but was required for efficient proteolytic cleavage of nsp1 and for optimal replication (Brockway and Denison, 2005). Further work showed that deletions proximal to residue K_{124} did not result in recovery of infectious virus, although, point mutagenesis in the amino-terminal region did result in infectious virus albeit with altered replication and RNA synthesis; indicating a potential role for nsp1 in RNA synthesis and virus replication. Zust et al. (Zust et al., 2007) deleted nucleotides 829–927 from the carboxy-terminal region of the MHV nsp1 and rescued a virus that had growth kinetics in cell culture that were indistinguishable to those showed by parental wild-type virus. The resulting recombinant, MHV-nsp1Δ99, was attenuated in vivo, however, replication and spread of the rMHV were restored in type I IFN receptor-deficient (IFNAR) mice when compared to wild-type mice; leading the authors to hypothesise that MHV nsp1 interferes with the type I IFN system.

Deletion of nsp2 in both MHV and SARS-CoV resulted in the recovery of infectious virus (Graham et al., 2005). The authors reported that RTCs in cells infected with the MHVΔnsp2 virus were morphologically identical to those infected with wild-type parental virus, even though they lacked nsp2 protein; nsp2 co-expressed using a retroviral system was recruited to the replication complexes in the MHVΔnsp2 infected cells. Overall the results showed that nsp2 is not required for replication in cell culture.

The nsp3 replicase protein is the largest replicase component and a multifunctional protein that encodes several enzymatically active domains, including two papain-like proteases and ADP-ribose-1’-monophosphatase (ADRP; previously known as the coronavirus X domain). Generation of an ADRP-deficient HCoV-229E virus showed that this conserved coronavirus enzyme is not required for replication in cell culture with no detectable differences in RNA synthesis or virus growth when compared to parental virus (Putics et al., 2005). A rMHV encoding an inactivated ADRP replicated, although with reduced titres, in the livers of infected mice but did not induce liver disease (Eriksson et al., 2008). The authors observed reduced IL-6 production in the spleens and livers of mice infected with the ADRP-deficient MHV and postulated that ADRP exacerbates MHV-induced liver
pathology through the induction of inflammatory cytokines. Recombinant HCoV-229E and SARS-CoV viruses expressing inactivated ADRPs have been shown to have an increased sensitivity to IFN-α when compared to the parental viruses indicating that ADRP may play a role in virus escape from host innate immune responses induced by a coronavirus infection (Kuri et al., 2011).

The nsp4 replicase protein is a membrane protein, with four predicted TM-spanning domains, involved in the assembly of RTCs on DMVs by anchoring the RTC to the membranes (Gorbalenya et al., 1989). Various deletions of the TM domains 1–3 from MHV nsp4 either did not result in recovery of infectious virus, resulted in viruses that were impaired for growth or those that demonstrated wild-type replication (Sparks et al., 2007). Indicating that nsp4 is required for MHV replication and that TM domains 1–3 are essential but TM4 and the carboxy-terminal amino acids K398–T492 are dispensable for replication. Amino acid substitution N258T of MHV nsp4 resulted in a rMHV with a ts phenotype that caused a reduction in DMV formation and partial localisation of nsp3 and nsp4 to mitochondria at the non-permissive temperature, indicating a role of nsp4 in DMV assembly (Clementz et al., 2008).

The nsp12-16 replicase proteins, encoded by the ORF1b region of gene 1, have a variety of enzymatic activities associated with RNA synthesis and processing (Ziebuhr, 2008); in which nsp12 has the RNA-dependent-RNA polymerase activity, nsp13 helicase activity, nsp14 3′-to-5′ exoribonuclease (ExoN) activity, nsp15 nidoviral endoribonuclease (NendoU) (this enzyme is unique to and conserved in all nidoviruses) activity and nsp16 2′-O-ribosyltransferase (MT) activity. A SARS-CoV-based replicon system was used to show that the RNA-processing enzymes ExoN, NendoU and 2′-O-MT are essential for RNA synthesis (Almazan et al., 2006). Inactivation of the 2′-O-MT domain in the MHV and HCoV-229E nsp16 replicase proteins resulted in recovery of infectious recombinant viruses that induced an increased expression of type I IFN in virus-infected cells and were highly sensitive to type I IFN (Zust et al., 2011). The authors reported that induction of type I IFN by the 2′-O-MT-deficient viruses was dependent on the cytoplasmic RNA sensor Mda5. They proposed, due to the link between Mda5-mediated sensing of viral RNA and 2′-O-methylation of viral mRNAs, that RNA modification by 2′-O-methylation provides a molecular signature allowing discrimination between self and non-self mRNA in the context of innate immune responses by the cell. Interestingly, the 2′-O-MT-deficient MHV was apathogenic in wild-type mice but both replication and virus spread were restored in IFNAR mice or in mice lacking Toll-like receptor 7 and Mda5 (Zust et al., 2011).

In order to study the pathogenicity determinants in IBV, the VV-based IBV reverse genetics system was used to replace the genes downstream of the replicase gene in the apathogenic IBV strain Beaudette with those from the pathogenic strain M41. A GPT plasmid containing the distal part of the Beau-R replicase gene fused to the M41 sequence from the S gene to the poly(A) tail was used as a donor sequence. A Beau-R-based cDNA consisting of the complete replicase gene, followed by part of the S gene fused to the N gene and 3′-UTR, was used as a receiver
sequence (Armesto et al., 2009). The resultant chimaeric virus, rBeauR-Rep-M41-Struct, was tested for pathogenicity in chickens. Observations of clinical signs, ciliary activity and analyses for viral RNA in the trachea of the infected chickens, showed that rBeauR-Rep-M41-Struct was not pathogenic, suggesting that determinants of pathogenicity reside in the replicase gene (Armesto et al., 2009). Therefore, to determine which region or regions of the Beaudette replicase gene are involved in loss of pathogenicity, the replicase gene of rBeauR-Rep-M41-Struct was sequentially replaced with the replicase gene sequence from M41 using the TDS system. Several rVVs, containing chimaeric Beau-R/M41 replicase sequences, were generated and viable rIBVs with the chimaeric replicase sequences are being used for virus pathogenicity studies.

2.8 Using coronavirus reverse genetics systems for gene delivery

The advent of reverse genetics systems for coronaviruses has also allowed the study of their potential to act as vectors for the expression and delivery of foreign genes. The large size of the genome means in principle the virus could accommodate large inserts while the specific tropism afforded by the S glycoproteins means that vectors could be targeted to the desired tissues. Due to the nature of coronavirus transcription, provided a TRS signal is also included, foreign genes can in principle be inserted into the coronavirus genome and expressed as a new sub-genomic mRNA. The non-essential phenotype of the group-specific genes means that targeted replacement of these genes may offer an alternative to insertion in order to express a foreign gene from TRS signals already available.

Although any of the coronavirus accessory genes could be replaced with a foreign gene, or new genes inserted at any point in the genome distal to gene 1, studies carried out have shown that the stability of such recombinant viruses may depend on a number of factors including the spatial position of the foreign gene and the virus backbone, as well as the intrinsic properties of the genes themselves. These aspects of stability were studied, using a combination of FIPV or MHV expressing the firefly or Renilla luciferase genes (de Haan et al., 2003b; de Haan et al., 2005). Expression of a foreign gene in MHV was found to be greater the closer to the 3’ end of the genome it was inserted (de Haan et al., 2003b) and that Renilla luciferase was maintained more stably than the firefly luciferase (de Haan et al., 2005). The stability of the firefly gene was also dependent on genome position with higher instability seen at more 3’ positions, indicating that not only the gene but the position in which it is inserted can affect the stability of the resulting viruses. The virus background can also make a difference; firefly luciferase expression was found to be more stable following replacement of gene 3abc in FIPV than when inserted between the E and M genes of MHV (de Haan et al., 2005), indicating that replacement of non-essential genes, rather than insertion between viral genes, may result in recombinants with greater stability.
IBV has been proposed as a vaccine vector to protect against IB and other viral infections of chickens based on the insertion of a foreign gene. We have investigated the possibility of such vectors through the generation of a number of rIBVs expressing the reporter genes EGFP or Renilla luciferase using the Beau-R molecular clone. Recombinants in which gene 5 has been replaced show stability of expression to at least passage 8 in primary CK cells, and this is increased to at least passage 10 if the foreign gene has been codon optimised to that of IBV (unpublished data). Shen et al. (Shen et al., 2009) also generated a number of rIBVs expressing EGFP or firefly luciferase from their P65 Vero cell-adapted Beaudette strain. The greatest genetic stability was identified following replacement of gene 3ab with firefly luciferase, which maintained high levels of expression up to passage 15. This supports previous work (de Haan et al., 2005) proposing that greater stability was found with the replacement of non-essential genes. Shen et al. (Shen et al., 2009) also demonstrated the ability of IBV to express a number of other viral and host proteins including the SARS-CoV gene 6, the dengue virus 1 core protein and the eukaryotic translation initiation factor eIF3f. All the genes were found to be expressed although the dengue virus core protein was found to be the least stable, only reaching passage 3 in Vero cells, reiterating the fact that the inserted gene itself can have an effect on the stability of the virus.

An rTGEV in which EGFP replaced the non-essential gene 3a was found to be highly stable; the virus continued to express similar levels of EGFP after 20 passages (Sola et al., 2003). When studied in vivo, in swine, the insertion of EGFP was found to have an effect on the behaviour of the virus; growth in enteric tissues was found to decrease 10 to 100-fold and the mean day of death observed in piglets was delayed by 1 or 1.5 days. The study of such recombinants on their ability to induce immunity is required in order to assess their usefulness as vaccine vectors. This study addressed the issue by immunising pregnant sows with the recombinant virus expressing EGFP and assessing TGEV- and EGFP-specific antibody levels in both sows and progeny. Significant levels of both TGEV- and EGFP-specific antibodies were found in the serum of piglets, demonstrating that this vector had elicited lactogenic immunity, thus establishing that vaccine vectors based on coronaviruses are viable options as they are able to illicit immune responses in vivo.

In addition to utilising viruses as vectors for delivering foreign antigens for potential vaccines, reverse genetics has also allowed the modification of viruses for virotherapy, that is, a process in which viruses can be modified to target and destroy cancer cells. The potential of coronaviruses to be utilised as virotherapy vectors has been investigated. MHV has been retargeted to a non-native receptor through the use of a soluble adapter protein to initiate an infection of cells normally non-permissive for MHV (Verheije et al., 2006). The adapter protein consisted of the N-terminal D1 domain of the MHV cellular receptor mCEACAM1a, which binds the S glycoprotein of MHV. The mCEACAM1a D1 domain was linked to a six-amino acid histidine (His) tag via three alanine residues and a myc tag. The His tag binds to a sPvHis receptor expressed on target cells resulting in infection of the target cell with MHV. Additional adapter proteins were designed with a hinge
region that would facilitate dimerisation of the adapter protein mimicking the natural mCEACAM1a receptor. The mechanism of cell entry by MHV when mediated by the adapter protein was shown to be via S-mediated membrane fusion as used by MHV when naturally infecting murine cells.

Utilising knowledge previously gained from the deletion of the MHV group-specific genes (de Haan et al., 2002), the adapter proteins were inserted into the MHV genome replacing genes 2a and HE, resulting in the generation of an rMHV that could express its own adapter protein for targeting, and therefore could be independently propagated on target cells. However, only viruses expressing adapter proteins containing the hinge region were able to infect target cells expressing the sFvHis receptor and establish a multi-cycle infection resulting in the killing of cells. By generating a new adapter protein, coupling the mCEACAM1a receptor to the epidermal growth factor (EGF) protein and inserting this into the MHV genome, it was possible to target MHV to EGF receptor-expressing human cells (Verheije et al., 2009). The EGF receptor is abundantly expressed on a majority of tumours and it was shown that targeting MHV to glioblastoma cells expressing this receptor resulted in efficient killing of the cancer cells. While further work is required in this area, these studies demonstrate the potential for exploiting coronaviruses in the treatment of cancers.

Acknowledgements

The authors would like to acknowledge the support of the following organisations: Department of the Environment, Food and Rural Affairs (Defra), Biotechnology and Biological Sciences Research Council (BBSRC), Pfizer Animal Health, Intervet Schering-Plough, British Egg Marketing Board Research (BEMB) and Education Trust, British Poultry Council (BPC) and the British Egg Industry Council (BEIC).

References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


Senanayake, S. D. and Brian, D. A. 1999. Translation from the 5′ untranslated region (UTR) of mRNA 1 is repressed, but that from the 5′ UTR of mRNA 7 is stimulated in coronavirus-infected cells. *Journal of Virology* **73**: 8003–8009.


CH02  CORONAVIRUS REVERSE GENETICS


REFERENCES

by dendritic cell transfer through DC-SIGN. *Journal of Virology* 78: 5642–5650.


