New virus diseases: visible evolution

"... descent with modification ..." Charles Darwin

MC HORZINEK

Department of Infectious Diseases and Immunology, Veterinary Faculty, Utrecht University, PO Box 80.163, Yalebelaan 1, De Uithof, 3508 TD Utrecht, The Netherlands

Introduction

As Stephen Jay Gould (1991) has discovered, Darwin never used the word 'evolution' in his original writings, probably due to connotations he wanted to avoid. For the evolution of viruses, "descent with modifications" is certainly a more appropriate formula, devised some forty years before Martinus Willem Beijerinck at Delft elaborated the concept of a contagium vivum fluidum (a living, fluid infectious matter). It is these modifications that keep the virologists' minds occupied, the metamorphoses of mutation and recombination in the genome, the consequent differential effects on target cells, the changes in disease producing potential, the variation-versus-fixation phenomena in populations, but above all, the sudden emergence of new diseases.

When I had to decide about the topic for this prestigious lecture, I did not hesitate for a moment to single out virus evolution; my reasons, however, for this choice were biographical rather than esoteric. I remember cramming late-evening seminars at early International Virology Congresses where the topic was discussed in rather conjectural terms. As a young man, I considered this a futile mental exercise because so many "useful" things still had to be discovered in virology. A decade or so later, when everybody was using techniques in molecular genetics and when we employed them in characterising positive-stranded RNA viruses, we literally stumbled upon virus evolution. Today, I view the viral "descent with modifications" as the most exciting facet of biology. It is a personal bias certainly, but perhaps forgivable for a molecularly-inclined veterinarian and Darwinophile with a Roman Catholic upbringing.

This paper is intended to place recent information, preferentially that obtained at my laboratory in Utrecht, in the perspective of my chosen topic. I shall discuss toroviruses, which we have characterised for their own sake — viruses in search of serious potential, the variation-versus-fixation phenomena in populations, the trivial sense because small animal practitioners had never seen a similar enteritis and myocarditis in dogs before. More so, however, it has been baptised 'porcine respiratory coronavirus'.

Hence new syndromes are not necessarily caused by new viruses. There are famous examples of sudden disease outbreaks in species where neither the condition nor the infection had occurred before. On the veterinary stage, canine parvovirus (CPV) made its appearance in 1978, and although it has closely related peers in other carnivores (CPV and feline panleukopenia virus share more than 98% of their DNA sequence) it is new in the trivial sense because small animal practitioners had never seen a similar enteritis and myocarditis in dogs before. More significantly, serological evidence of the infection in dogs has not been obtained before the 1970s, which indicates that the virus, or a similar one, had not infected dogs before that date. A variant parvovirus of probably feline origin occupied another ecological niche, as it had done before in mink. Once in the dog, however, the evolution of CPV continues: after 1986, most isolates differed antigenically from their predecessors. The new type 2b largely replaced the previous type 2a, and the original CPV type 2 had been replaced between 1979 and 1981 by the type 2a strain. Types 2b and 2a differ by only two amino acid-changing nucleotide substitutions in the capsid protein genes (Parrish et al 1991).
The essence of these considerations is that quantitative genetic comparisons can neither explain nor predict new virus diseases: large deletions in an 'old' virus may be of no biological consequence, as we have recently observed in feline infectious peritonitis coronavirus, whereas single amino acid changes may have tremendous effects on the virulence of a virus.

What then about qualitative changes? Are there examples where viruses pick up 'new' genetic information from the cell or from another virus? In RNA virology, this has been found repeatedly. A famous example is bovine viral diarrhoea virus (BVDV), where insertions of ubiquitin-like sequences, of host cellular mRNA, and of short polynucleotide stretches have been demonstrated. These recombinations have vital consequences: they may determine the difference between non-cytopathogenic and cytopathogenic BVDV biotypes because both cooperate to determine the fatal outcome of the infection named mucosal disease. The recombinant virus is definitely involved in a new disease. It is not new in the 'hitherto unknown' sense of the word, rather it is different in terms of a (reproducibly) altered pathogenesis (Liess et al 1991).

When working on coronavirus transcription, we encountered more evidence for non-homologous recombination: an additional stretch of nucleotides was discovered in the peplomer gene sequence of the mouse hepatitis virus (MHV) strain A59 that is absent from other strains. The mRNA2 of the former contains two open reading frames (ORF), the second one is a pseudogene that lacks a translation initiation codon. The amino acid sequence of ORF2 is 30% identical to that of the haemagglutinin/acetyltransferase (HE) molecule of the negative-stranded influenza C virus, and a short stretch of nucleotides immediately upstream is 83% identical to a major histocompatibility complex (MHC) class I sequence (Luytjes et al 1988). The MHV-A59 strain does not possess esterase activity, which is understandable because ORF2 is not translated. On the other hand, bovine coronavirus does possess an esterase activity similar to that of influenza C virus. It has been localised on an additional envelope protein (E3) termed haemagglutinin, which is visible as a shorter class of spikes on the virion surface. As the HE gene is absent in coronaviruses from other antigenic clusters, recombination involving an influenza C-like virus and an ancestral coronavirus was postulated. A heterologous recombination event is required to explain the presence of the same genetic module in viruses of fundamentally different families (Luytjes et al 1988). Conceivably, a 'protocoronavirus' has cannibalised the orthomyxoviral enzyme information in a unique event long ago.

These two examples show that adoption of foreign genes and incorporation into the set of resident information occurs, and that adoption and incorporation may be pathogenetically important in RNA viruses. More important still, it may offer an evolutionary advantage. Viruses with this property adapt to altering environments and expand into new host species, perhaps causing new diseases. If the virus acquires genes coding for proteins with new structural properties, a new virus will arise. Taxonomically inclined virologists would give it a new name and assign it to a new species, genus, family.

Toroviruses — In Search of a Disease

In 1972, the late Dr Franz Steck and his colleagues isolated a virus from a horse with diarrhoea during routine diagnostic work at the Veterinary Faculty of Berne, Switzerland. The isolate displayed an unusual morphology but was not studied in detail until a similar virus was found in calves with diarrhoea in Breda, Iowa, USA, in 1979 (Woode et al 1982; Weiss et al 1983); Berne virus (BEV) and Breda virus (BRV) are antigenically related, and with BRV there appears to be a second 'serotype' (Woode et al 1985). Similar pleomorphic viruses were seen in the stools of children and adults with gastroenteritis (Beards et al 1984) and identified as BEV/BRV-related (Koopmans 1990). The presence of nucleocapsids in the form of a doughnut, a shape described in Latin by the word torus, led to the proposal of the name 'torovirus' (Horzinek and Weiss 1984).

Neutralising antibodies against toroviruses are widespread in horses, cattle, sheep, goats, and pigs (Weiss et al 1984; Brown et al 1987; Koopmans et al 1989). A more inclusive serological test is needed to determine the prevalence of more distantly related representatives in nature. Toroviruses have not been isolated from humans, cats or pigs, but physical evidence of their existence has been obtained. After experimental oral infection, BRV was found to cause neonatal calf diarrhoea (Woode et al 1985). Under field conditions, toroviruses may have a role in neonatal diarrhoea and respiratory syndromes and perhaps in generalised infections. The relative economic importance of these infections remains to be established.

Torovirions are pleomorphic and measure 120 to 140 nm in their largest diameter. Spherical, oval, elongated, and kidney-shaped particles are observed. Their two most conspicuous features are the spikes on the envelope, which resemble the peplomers of coronaviruses, and the tubular nucleocapsid of helical symmetry, which determines the shape of the virion (Weiss et al 1983).

From sedimentation studies, the BEV genome was estimated to be more than 20 kilobases (kb) in length. The genomic RNA is single-stranded, polyadenylated and infectious in a transfection assay (Snijder et al 1988). It probably contains six ORFs. As in coronaviruses, the first two from the 5'-end (ORF 1a and 1b) are translated from genomic RNA and constitute the viral replicase gene. The four remaining reading frames, of which ORF 2, 3 and 5 have been identified as structural genes, are expressed by the generation of a 3'-coterminal nested set of mRNAs (Snijder 1991).

From the genome of BRV, only the 269 nucleotides (nt) directly upstream of the poly(A) tail have been determined so far (Koopmans et al 1991). The sequence in this region is 93% identical to that of BEV, indicating that bovine and equine toroviruses are closely related, and labelled BRV genomic RNA fragments strongly hybridised to a broad range of BEV cDNA clones (Koopmans et al 1991). An exception may be the 5' part of the spike protein gene, which is highly variable also in coronaviruses (Spaan et al 1988).

By metabolic labelling of BEV-infected cells, proteins of 19K, 22K, 37K, and 75K-100K were identified in [35S]methionine-labeled virions (Hozlitzke et al 1984). This is a pattern hitherto unknown for RNA viruses. After sequence analysis of the BEV structural genes and characterisation of their products, the 19K, 22K and 75K-100K polypeptides could be affiliated with ORF 2, 3 and 5, respectively (Den Boon et al 1991a; Snijder et al 1989, 1990b).

Viral replicase genes are essential in viral evolution because they provide an anchor in the stormy seas of genetic changes. Several domains in the replicase proteins are highly conserved among disparate groups of plant and animal RNA viruses (Strauss and Strauss 1988; Goldbach and Wellink 1988). Among positive-stranded RNA viruses, a number of successful replicase classes has been recognised, all of which are associated with a specific type of genome organisation and replication strategy. Superfamilies of picornavirus-like, alphavirus-like, flavivirus-like (Strauss and Strauss 1988; Goldbach and Wellink 1988) and coronavirus-like (CVL) (Snijder et al 1990a; Den Boon et al}
1991b; Snijder et al 1993) positive-stranded RNA viruses have been discriminated.

A comparison of the replication strategy and replicase properties of corona-, toro- and arteriviruses clearly distinguishes the CVL replicase module from its alpha-, picorna-, and flavivirus-like equivalents. The organisation (two ORFs) and expression (using ribosomal frameshifting) of the gene as well as the arrangement of conserved domains within its product are unique. Moreover, the basic CVL genome organisation (replicase gene-envelope protein-nucleocapsid protein genes) and expression strategy (using a 3'-coterminal nested set of 5 to 8 mRNAs) are unparalleled.

The concept of a ‘CVL replicase dictating a CVL replication strategy’ originated from the molecular characterisation of BEV and arose in particular from the detection of striking homologies between toro- and coronaviral replicase sequences (Snijder et al 1990a). Subsequently, the CVL group was expanded to include the arteriviruses, which also produce a 3'-coterminal nested set of mRNAs. The protein-coding regions of the genomes of equine arteritis virus (EAV) Den Boon et al 1991b), lactate dehydrogenase-elevating virus (LDV) (Godeny et al 1990; Kuo et al 1991; Godeny et al 1993; Chen et al 1993) and Lelystad virus (Meulenberget al 1993a; Conzelmann et al 1993) have now been sequenced fully. Although arteriviral replicase genes are considerably smaller (9.5-12 kb) than their toro- and coronaviral counterparts (17-19 kb), they also contain two ORFs, the downstream ones of which are expressed by ribosomal frameshifting. In addition, the conserved domains first identified in toro- and coronaviral ORF 1b products have been identified in the same relative positions in the arterivirus replicase.

The discovery of toroviruses was serendipitous—following a careful field observation (Woode et al 1982) and an incidental virus isolation (Weiss et al 1983). It was not easy to convince the community of virologists that toroviruses really were ‘new’, and the first Snowdon lecturer, Dr F Murphy, sent me many electron micrographs that showed very similar particles in mouse hepatitis material. The contentious title of my first structural analysis read “Berne virus is not coronavirus-like” and emphasised its structural uniqueness (Horzinek et al 1984). However, some years later we had to publish that BEV “polymerase is expressed by ribosomal frameshifting and contains sequence motifs which indicate that toro- and coronaviruses are evolutionarily related” (Snijder et al 1990a).

**Porcine Reproductive and Respiratory Syndrome — In Search of an Agent**

The disease that is currently referred to as porcine reproductive and respiratory syndrome (PRRS) had been known in the USA as mystery swine disease since 1987. It mainly affects sows and piglets in breeding herds, causing abortion and stillbirth in the former, and anorexia and respiratory distress in pigs of all ages. When it appeared in Europe in 1990 as a ‘new’ disease, it was immediately taken up as an object of study by researchers in industry, government laboratories and academia.

The disease came with a bad reputation from the USA, where several groups had tried to elucidate its infectious etiology. Various porcine viruses were isolated, but none fulfilled Koch’s postulates. Early suspects were encephalomyocarditis virus, porcine parvo- and enteroviruses, orthomyxoviruses, paramyxoviruses, pseudorabies virus, pestiviruses and chlamydia.

The mystery of the swine disease was finally solved when Lelystad virus (LV) was isolated in the Netherlands and shown to cause the condition (Terpstra et al 1991). LV replicated in *vitro* exclusively in porcine lung alveolar macrophages, but only to low titres (Wensvoort et al 1992). When I was shown the first electron microscopic images, my excitement was great. The 50 to 65 nm particles with a 30 to 35 nm core (Wensvoort et al 1992) could be either alpha- or arteriviruses — and the former had been excluded by serology. Then other data surfaced: the genome is single-stranded and polyadenylated, about 15 000 nt in length and contains eight ORFs. The replication strategy is corona-, toro- and arterivirus-like, with a fan of six subgenomic RNA molecules occurring in infected cells that have 3'-coterminal sequences. Similar sequence motifs in the polymerase gene, a pseudoknot structure, and amino acid sequence homologies with LDV and, to a lesser extent, EAV identified it as a member of the proposed family Arteriviridae. However, this virus causing a new disease is new to the scientific community, its origin is unknown and its relationship to other arteriviruses is distant.

**Torovirus Evolution**

Features shared by members of the coronavirus-like superfamily include the basic genome organisation, the production of 3'-coterminal nested sets of mRNAs, and the presence of homologous replicase domains. Noticeable differences are dissimilar N proteins and nucleocapsid architecture, the probable absence of a common 5' leader sequence in the BEV mRNAs, the much smaller genome size and the absence of a large spike glycoprotein in arteriviruses.

Analysis of the BEV genome has revealed vestiges of all those processes that make RNA virus evolution such a fascinating field. Firstly, there are the unambiguous primary protein sequence homologies in the replicase. They are the most convincing evidence for common ancestry and for the divergent evolution, which has resulted in the present-day replicase genes of toro-, corona- and arteriviruses. Secondly, the common ancestry probably connects the BEV and coronaviral S and M proteins. The primary sequence similarities of these proteins are no longer detectable, but their predicted structural characteristics and their linkage to homologous replicase genes are highly suggestive. The observations made for the S and M proteins are in agreement with the general perception that structural proteins evolve at a higher rate than nonstructural ones. Thirdly, toroviruses underline the role of recombination in RNA virus evolution in general, and in the evolution of corona- and toroviruses in particular (Snijder et al 1991). Fourthly, CVL genomes may show the results of a process for which the name ‘modular evolution’ has been coined (Zimmern 1987; Strauss and Strauss 1988; Goldbach and Wellink 1988): the exchange by recombination of complete genes or gene sets (modules). Modular evolution provides a plausible explanation for the presence of a CVL replicase and for the definitely coronavirus-unlike set of structural proteins in the arterivirus group (Vries et al 1992). However, the BEV N protein is difficult to explain: its size, sequence and quaternary structure embracing the genomic RNA are features neither of corona- nor of arteriviruses.

Enlarging on modular evolution, RNA viruses conspicuously show arrays of conserved sequence motifs in some enzymes, irrespective of the extensive variations in other proteins. Proteins with such motifs include RNA-dependent RNA polymerase, putative RNA helicase, chymotrypsin-like and papain-like proteases, and methyltransferases. The genes for these proteins form partially conserved modules in large subsets of viruses. Koonin and Dolja (1993) discussed the attractive concept of the virus genome as an evolutionarily stable ‘core’ of house-keeping genes accompanied by a much more flexible ‘shell’ consisting of genes for virion components. These authors considered the shuffling of ‘shell’ genes between distant groups of viruses a major factor of virus evolution.
Taxonomy and Conclusion

Intrigued by their unique morphological and physicochemical characteristics, I initially proposed the *Toroviridae* as a new family of enveloped RNA viruses (Horzinek and Weiss 1984). However, our more recent analyses of the genetic information and replication strategy of the prototype Berne virus (Snijder et al 1988, 1990a, 1990c) have revealed that toroviruses are not that unique; they are clearly related to the *Coronaviridae* and, albeit more distantly, to the arterviruses (Den Boon et al 1991b). This information has led to their reclassification as toroviruses as a new genus within the *Coronaviridae* family (Pringle 1992) and to the introduction of the unofficial term 'coronavirus-like superfamily' to indicate the evolutionary ties between the three virus groups of toro-, arteri- and coronaviruses.

A meaningful classification of the members of the CVL superfamily clearly requires four hierarchical levels. As mentioned above, the corona- and torovirus species have now been domiciled in two genera. The obvious evolutionary link of this superfamily clearly requires four hierarchical levels. As mentioned above, the corona- and torovirus species have now been domiciled in two genera. The obvious evolutionary link of this superfamily would be most accurately reflected by accommodating the present genus arterivirus in a family of its own and by establishing an order (to replace the 'superfamily') comprising the *Coronaviridae* and *Arteriviridae* families.

The above implies that the importance of traditional taxonomic criteria, such as virion and nucleocapsid structure, will diminish; characteristics like genome organisation, replication strategy and sequence homologies will become more important. Thus molecular virology has provided the basis for reassigning EAV, formerly considered a member of the *Togaviridae* family. The continuous nature of sequence variation will pose taxonomic and nomenclative problems emphasising the impossible reconciliation of 'fuzzy' phenomena and 'crisp' terms. Which percentage of identical amino acids is required to define a genus or a family? How should the results of genetic recombination, the exchange of entire modules of information be integrated into virus systematics? On the other hand, these two issues confront virologists most cogently with the genetic basis of the problem. As virus evolution is governed by heredity, the processes of mutation, recombination, and selection will have to be accommodated in future virus taxonomy.

Acknowledgments

During the last twenty years, many students and postdoctoral fellows have contributed to coronavirus, torovirus and equine arteritis virus research performed at the Virology Division, Veterinary Faculty in Utrecht, The Netherlands. In acknowledging their conceptual input I would like to dedicate this article to Ben van der Zeejst, Willy Spaan and Peter Rottier who were in charge of the molecular virology group at my unit in this period.

Special thanks go to Marion Koopmans and Eric Snijder, once torovirolologists *sensu stricto*. In addition I should like to acknowledge my (former and actual) collaborators Mario van Berlo, Johan den Boon, Peter Bredenbeek, Lisette Cornelissen, Joke Ederveen, Raoul de Groot, Arnold Herreweghe, Annelies Kromman, Willem Luytjes and Twan de Vries.

I would also like to thank Marianne Weiss and Gerald Wood for many fruitful discussions in the early phases of torovirus research, and the authors of the following references for sharing unpublished data and preprints of their publications: Chen et al (1993), Godeny et al (1993), Herold et al (1993) and Meulenberg et al (1993b). More detailed reviews on toroviruses are in press (Koopmans and Horzinek *Adv Virus Res*; Snijder and Horzinek *J Gen Virol*).

References

- Beards GM, Hall C, Green J, Flewett TH, Lamouliatte F and DuPasquier P (1984) *Lancet* ii:1050
- Den Boon JA, Snijder EJ, Krijnsse Locker J, Horzinek MC and Rottier PJM (1991a) *Virology* 182:655
- Den Boon JA, Snijder EJ, Chisimide BD, De Vries AAF, Horzinek MC and Spaan WJM (1991b) *J Virol* 65:2910
- Brown DWG, Beards GM and Flewett TH (1987) *J Clin Microbiol* 25:637
- Chen Z, Kuo L, Rowland RRR, Even C, Faasberg LS and Plagemann PGW (1993) *J Gen Virol* 74:643
- Conselman KK, Visser N, Van Woensel P and Thiel HJ (1993) *Virology* 193:329
- De Vries AAF, Chisimide EH, Horzinek MC and Rottier PDM (1992) *J Virol* 66:6294
- Distelbosch B, Wunner WH et al (1983) *Proc Natl Acad Sci USA* 80:70
- Godeny EK, Speichudt DW and Brinton MA (1990) *Virology* 177:768
- Godeny EK, Chen L., Kumar SN, Medhvun SL, Koonin EY and Brinton MA (1993) *Virology* 194:585
- Goldbach R and We protection (1998) *Invertebr Virol* 29:260
- Pringle CJ (1991) *Ever Since Darwin*, Penguin, London
- Horzinek J, Raabe T, Schelle-Prinz B and Siddell SG (1993) *Virology* 198:680
- Horzinek MC and Weiss M (1984) *Zentralbl Veterinarmed Reihe B* 31:049
- Horzinek MC, Weiss M and Ederveen J (1984) *J Virol* 63:645
- Koonin EY and Dolja VV (1993) *CRC Rev Biochem Mol Biol* (in press)
- Koopmans M (1990) *Diagnosis and Epidemiology of Torovirus Infections in Cattle*, PhD thesis, Utrecht University
- Koopmans M, Van den Bocm U, Woode GN and Horzinek MC (1989) *Vet Microbio* 19:233
- Koopmans M, Snijder EJ and Horzinek MC (1991) *J Clin Microbiol* 29:493
- Koo LL, Harty JT, Ericson L, Palmer GA and Plagemann PGW (1991) *J Virol* 65:5118
- Liest L, Moenning V, Pohlenz J and Trautwine G (1991) *Arch Virol Suppl* 3
- Luytjes W (1989) *Surface Proteins of Marine and Bovine Coronaviruses*, PhD thesis, Utrecht University
- Luytjes W, Bredeneck PJ, Nooten AFH, Horzinek MC and Spaan WJM (1988) *Virology* 166:415
- McNell D and Freiberger P (1993) *Fuzzy Logic*, Bookman Press, Melbourne
- Meulenberg,UL, Hult MM, De Meijer EJ et al (1993a) *Virology* 192:62
- Meulenberg JJM, De Meijer E and Moormann RJM (1993b) *J Gen Virol* (in press)
- Page KW, Mawdii KL and Britton P (1991) *J Gen Virol* 72:579
- Parnish CR, Aquadro CF, Strasheim ML, Evermann JF, Sgro JY and Mohammed HO (1991) *J Virol* 65:6544
- Pringle CJ (1992) *ASM News* 58:475
- Raschaert D, Duarte M and Laude H (1990) *J Gen Virol* 71:2599
- Snijder E (1991) *Bernevirus, Replication and Evolution of the Torovirus Prototype*, PhD thesis, Utrecht University
- Snijder EJ, Ederveen J, Spaan WJM, Weiss M and Horzinek MC (1988) *J Gen Virol* 69:2135
- Snijder EJ, Den Boon JA, Spaan WJM, Verjans GMGM and Horzinek MC (1989) *J Gen Virol* 70:3363
- Snijder EJ, Den Boon JA, Bredenbeek PJ, Horzinek MC, Rijnbrand R and Spaan WJM (1990a) *Nucleic Acids Res* 18:4535
- Snijder EJ, Den Boon JA, Spaan WJM, Weiss M and Horzinek MC (1990b) *Virology* 178:355
- Snijder EJ, Horzinek MC and Spaan WJM (1990c) *J Virol* 64:331
- Snijder EJ, Den Boon JA, Horzinek MC and Spaan WJM (1991) *Virology* 180:448
- Snijder EJ, Horzinek MC and Spaan WJM (1993) In *Coronaviruses: Molecular Biology and Pathogenesis*, edited by Laude H and Vautheroer JF, Plenum, New York, (in press)
- Spaan WJM, Cavanagh D and Horzinek MC (1988) *J Gen Virol* 69:2939
- Strauss JH and Strauss EG (1988) *Ann Rev Microbio* 42:657
- Terscha C, Wensvoorn G and Pol JMA (1991) *Vet Quarte* 13:131
- Weiss M, Steck F and Horzinek MC (1983) *J Gen Virol* 64:1849
- Weiss M, Steck F, Kaderli R and Horzinek MC (1984) *Vet Microbio* 9:523
- Wensvoorn G, de Kuyver EP, Pol JMA, Wagenaar F, Moorman RJM et al (1992) *Vet Microbio* 33:185
- Wesley RD, Woods R&D and Chong AK (1991) *J Virol* 65:3369
- Wood GN, Reed DE, Runnels PL, Herrig MA and Hill HT (1982) *Vet Microbio* 7:221
- Wood GE, Saif LJ, Quresida M et al (1985) *Am J Vet Res* 46:1003
- Zimmerman D (1987) In *RNA Genetics*, edited by Holland JI, Domingo E and Ahliqui P, CRC Press, Boca Raton, vol 2, p 211