Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

IV, 6. Calicivirus RNA recombination

David O. Matson

Infectious Diseases Section, Center for Pediatric Research, Norfolk, Virginia 23510

Introduction

In this review, I will discuss evidence for the occurrence of RNA recombination in *Caliciviridae*, both within and outside the family. Constraints on recombination provided by the genomic diversity of caliciviruses (CVs), as well as implications of recombination on the natural diversity of CV strains and the clinical and biologic significance of RNA recombination, also will be considered. First, I will review some features of CVs that affect understanding of recombination.

Overview of genomic structure in CVs

The CV genome is a positive-sense, single-stranded, polyadenylated RNA molecule of about 7500 nucleotides in length. CVs fall into four genera that differ in their genomic organization (Green *et al.*, 2000a) (Fig. 1). Norwalk-like viruses (NLVs) have three open reading frames (ORFs). ORF1 encodes a polyprotein cleaved during replication into a set of nonstructural proteins, ORF2 encodes the capsid protein, and ORF3 encodes a protein that appears to be a minor structural protein (Sosnovtsev and Green, 2000). Where studied, CVs have been shown to synthesize a positive-sense subgenomic RNA that begins at the 5' of the capsid gene and that is co-terminus with the genome (Meyers *et al.*, 1991; Neill and Mengeling, 1988; Sosnovtser and Green, Section IV, Chapter 2 of this book). Vesiviruses differ from NLVs in having a longer genome that in some Vesiviruses (e.g., Pan-1; Rinehart-Kim *et al.*, 1999), but not others (e.g., feline CV; Carter *et al.*, 1992), includes a longer ORF1 with an additional predicted protein at its N-terminus. ORF2 of vesiviruses is longer than that of NLVs, with the extra nucleotides at the 5' end of ORF2. This extra sequence encodes a protein fragment that must be post-translationally cleaved to agree with experimental data of *Vesivirus* virion structure (Prasad *et al.*, 1994). ORF3 of vesiviruses is about one-half the size of that of NLVs (~120 amino acids vs. 250-275 amino acids, respectively). In lagoviruses and Sapporo-like viruses (SLVs), the genes that are in ORF1 and ORF2 of NLVs and vesiviruses are fused into one longer ORF1. A gene comparable to that of ORF3 of NLVs also is present. An ORF in another frame at the 5' end of the capsid gene occurs among SLVs, but not in all SLV strains (Liu *et al.*, 1999; Jiang *et al.*, 1997).

Fig. 1. Genome organizations of CV genera. The genome of Norwalk-like viruses (NLVs) has three open reading frames (ORFs), that 5' to 3' encode a nonstructural polypeptide (ORFI, white bars), the virion capsid gene (ORFII, grey bars), and a minor structural protein (ORFIII, black bars). The genomes of the other three genera differ from that of NLVs in the length of the ORFs, including an unique gene at the 5' end of ORF1 in at least one Vesivirus and a post-translationally cleaved N-terminus of the capsid protein in vesiviruses and Sapporo-like viruses (SLVs). Some sequence comparisons suggest that the longer ORF3 of NLV (and the comparable gene of SLVs) arose by intragenic recombination.

CV antigenic and genomic sequence diversity

The antigenic determinants (neutralization epitopes) that induce immunity against CVs presumably are located on the surface of the virion capsid. This capsid is composed of 180 copies of the capsid gene product, paired into 90 dimers (Prasad *et al.*, 1994; Prasad *et al.*, 1999). Despite the existence of just one capsid protein, CVs exhibit extensive antigenic diversity. In the best-characterized genus, *Vesivirus*, at least 40 distinct serotypes (neutralization types) exist, not including feline CVs and closely related strains, which among themselves are so diverse antigenically that definition of serotypes has been problematic (Lauritzen *et al.*, 1997; Hohdatsu *et al.*, 1999; Smith, 2000). The distinct *Vesivirus* serotypes are certainly determined by differences in nucleotide sequence of the capsid gene, resulting in differences in surface epitopes. The nucleotide differences sufficient to change the serotype are unknown, but likely to occur in a few distinct regions of the capsid gene (Neill, 1992; Rinehart-Kim *et al.*, 1999; Neill *et al.*, 2000).

When many capsid nucleotide sequences from different CV strains are simultaneously compared in phylogenetic analyses, the sequences within a genus fall into statistically significant clades (Berke *et al.*, 1997; Green *et al.*, 2000b). The biologic significance of these distinct clades is unknown. It is clear that such clades are related to differences in capsid gene sequences; sequence differences are less marked in the RNA polymerase gene: when RNA polymerase region sequences are analyzed in phylogenetic analyses, statistically significant differences similar to those observed among capsid gene sequences do not occur (Berke *et al.*, 1997). It is possible that separate capsid sequence clades within a genus indicate separate serotypes, but, even for *Vesivirus* capsid sequences, an insufficient number of strains have been analyzed to associate specific sequence differences with differences in serotype.

Evidence for recombination within CVs

With the description of statistically significant phylogenetic clades within CV genera, data were available to recognize strains that might be natural recombinants within CVs. Two examples are the well-characterized Argentine strain 320 (Arg320) and Snow Mountain virus (SMV), one of the prototype CVs, recognized to be recombinants when the RNA polymerase and capsid regions of these strains were characterized (Hardy *et al.*, 1997; Jiang *et al.*, 1999) (Fig. 2). At the time of publication, recombination was more certain for Arg320, because the sequence was derived from a single cDNA insert spanning the ~3.0 Kb at the 3' end of the genome, including the end of ORF1 and all of ORF2, ORF3, and the 3' non-coding region. In Arg320, the change of relative sequence identity occurred at the ORF1/capsid gene junction, indicating that the recombination occurred there. This site also was suggested (see below) to be the break-and-rejoin site for recombination between CVs and picornaviruses. For Arg320, the ORF1 sequence was closest to that of Lordsdale virus, among sequenced NLVs, and the capsid and ORF3 sequences were closest to those of Mexico virus. A similar change of relative sequence identity also occurred in SMV, when partial polymerase and capsid sequences were compared to reference Mexico and Melksham viruses. While SMV was likely also to be a recombinant virus, the capsid and RNA polymerase region amplicons of SMV were generated separately and that fact did not exclude the possibility of different sources of strains. Xi Jiang has confirmed the recombinant status of SMV by sequencing a single cDNA derived from a single RT-PCR amplicon (X Jiang, personal communication).

Potential origins of recombinants within CVs

Generation of recombinants within CVs requires biologic and molecular attributes of CVs. Outbreaks caused by multiple CV strains and co-infection by different HuCV strains occur (Matson *et al.*, 1995; Gray *et al.*, 1997; Reuter *et al.*, 2002). Infection of single cells simultaneously by two CVs implies absence of immune or molecular

Fig. 2. Sequence comparisons of Argentine CV strain 320 (Arg320; top) and Snow Mountain virus (SMV; bottom), each with two other NLVs. Arg320 is significantly closer to Lordsdale virus (LoV) than Mexico virus (MxV) in the RNA polymerase genome region and significantly closer to MxV than LoV in the capsid and ORF3 genes. Deletions/insertions in the capsid and ORF3 genes of Arg320 are shared with MxV, but not LoV. SMV is equally close to MxV and Melksham virus in the known polymerase region, but much closer to Melksham than MxV in the capsid region. ORF3 sequence is not available for SMV. (Box shadings as in Fig. 1. Data from Jiang *et al.*, 1999; Hardy ME *et al.*, 1997; and DO Matson, unpublished.)

Fig. 3. Sequence comparisons of the first 5' genomic 40 nt of a CV strain (ID="A" sequence for this Fig.) and of 40 nt near the 5' end of that strain's capsid gene (ID="B" sequence for this Fig.). Each A or B sequence is from a CV with a known complete genome sequence, is on a single line, and is repeated in two columns. In the left-hand column, each A or B sequence is compared with the first 40 nt of the Norwalk virus genome, i.e., the Norwalk virus A sequence (Jiang *et al.*, 1993). In the second column, the A or B sequence is compared with the first 40 nt (A sequence) of a prototype sequence for that genus. Within a genus, the A sequences are listed first and the B sequences given next. "-" indicates the nt at that site is identical to that in the comparison sequence. For EBHSV, the "*" indicates a residue I inserted into the 5' noncoding region sequence to improve alignment with the comparison sequence, RHDV Meyers. For the Norwalk-like (NLV) sequences, the two columns are identical. For each CV strain evaluated, A and B sequences of that strain showed sequence conservation. These comparisons demonstrate that A and B sequences of NLVs are highly conserved, with the strongest conservation being in the first 20 nt of the two regions. "A" sequences of NLV strains are much less conserved with homologous A or B sequences of strains in the other CV genera. The observation that the AUG (□) of the A and B regions for a single strain fall at the same position in the

Genus	Genomic	RNA Sequence	RNA Sequence
Host/Strain	Region	Hu/NV Jiang For Comparison	First Strain of Genus For Comparison
Norwalk-like			
A Hu/NV Jiang	1-40	GUGAUGUGAUGGCGUCAAAGACGCGUCCUACUGCU	GUGAUGUGAUGGCGUCAAAGACGCGUCCUACUGCU
A Hu/NV Schreier	1-40	-----G-----AAC	-----G-----AAC
A Hu/Southampton	1-40	-----G-----AA-	-----G-----AA-
A Hu/Chiba	1-40	-----G-----AG--A-	-----G-----AG--A-
A Hu/Lordsdale	1-40	-----A-----U--C--CUUCCG--G-C--	-----A-----U--C--CUUCCG--G-C--
A Hu/Camberwell	1-40	-----A-----U--C--CUUCCG--G-C--	-----A-----U--C--CUUCCG--G-C--
A Hu/GII Clarke	1-40	-----A-----U--C--CUUCCG--G-C--	-----A-----U--C--CUUCCG--G-C--
A Hu/GII Maryland	1-40	-----A-----U--C--CUUCCG--G-C--	-----A-----U--C--CUUCCG--G-C--
A Hu/Hawaii	1-40	-----A-----U--C--CUUCCG--G-C--	-----A-----U--C--CUUCCG--G-C--
A Hu/Jena	1-40	-----A--CUUU--A--G--U--UGGA--UC--A--GUUA	-----A--CUUU--A--G--U--UGGA--UC--A--GUUA
B Hu/NV Jiang	5354-93	-----A-----U--G--CUACAU--A--CC--UC	-----A-----U--G--CUACAU--A--CC--UC
B Hu/Chiba	5342-61	-----A-----U--G--CUACA--A--GC--A	-----A-----U--G--CUACA--A--GC--A
B Hu/NV Schreier	5339-78	-----A-----U--G--C--CCA--AU--GC--	-----A-----U--G--C--CCA--AU--GC--
B Hu/Southampton	5351-90	-----A-----U--G--C--CC--AA--GC--	-----A-----U--G--C--CC--AA--GC--
B Hu/Lordsdale	5081-120	-----A-----G--U--C--AAC--AU--A-	-----A-----G--U--C--AAC--AU--A-
B Hu/Camberwell	5081-120	-----A-----G--U--C--AAC--AU--A-	-----A-----G--U--C--AAC--AU--A-
B Hu/GII Clarke	5081-120	-----A-----G--U--C--AAC--AU--A-	-----A-----G--U--C--AAC--AU--A-
B Hu/GII Maryland	5081-120	-----A-----G--U--C--AAC--AU--A-	-----A-----G--U--C--AAC--AU--A-
B Hu/Hawaii	5081-120	-----A-----G--U--C--CC--AU--AA-	-----A-----G--U--C--CC--AU--AA-
B Hu/Jena	5017-86	---A---A---A--UGAC--UGACAAGCA--GUGC--GGAAG	---A---A---A--UGAC--U--ACAAGG--UG--GC--GAA-
Sapporo-like			
A Hu/Manchester	1-40	---C--G--U--GAUG---UUCCAAGCCAUCAAGCCAAUAG	GUGAUGGUGAUGGUGUUUCC AAGCCAAUCAAAGCCAAUAG
A Po/Saif	1-40	---C--C--GUG--CUAAUUGCCGUGCCGUGCCUADUGGCC	---C--U--U--CCCAA--C--CUU--G--GCCUADUGGCC
B Hu/Manchester	5158-60	A--AGUGUCC--GAUG--AGGGC--AUG--CUCCAA--CCAGAGC	A--G--GUU--G--G--GGG--UGGC--C--C--GAGC
B Po/Saif	5131-48	---UUC--UGAUG--AG--CGCCUGCC--CAACCCGUGCGGUG	---U--C--UGAU--GA--CGC--UGCC--AC--CGUU--GG--UG
Lagovirus			
A RHDV Meyers	1-40	---A--U--AUG--CG--CJ--UGUCG--C--U--A--UGG--ALGA	GUGAAGGUUUGCGCGGCUAUGUGCGGCCCUACUGGCAUGA
A RHDV Czech	1-40	---A--U--AUG--CG--CJ--UGUCG--C--U--A--UGG--ALGA	-----U-----U-----
A RHDV Iowa 2000	1-40	---A--U--AUG--CG--U--UGUCG--C--U--A--UGG--ALGA	-----U-----U-----
A RHDV Rossi	1-40	---A--U--AUG--CG--CJ--UGUCG--C--U--A--UGG--ALGA	-----U-----U-----
A RHDV SD	1-40	---A--U--AUG--CG--CJ--UGUCG--C--U--A--UGG--ALGA	-----U-----U-----
A EBHSV	1-40*	---AUU--AGG--GU--CGCUC--CGCC--U--G--GGC--GC	-----U--GC-----C--UGC--G--G
B RHDV Meyers	5296-335	---U--AUG--AG--G--GCC--CGAG--GCCGCAAG	---U-----A--GC--AAG--C--UGAG--GCCGCAAG
B RHDV Czech	5296-335	---U--AUG--AG--G--GCC--CGAG--GCCGCAAG	---U-----A--GC--AAG--C--UGAG--GCCGCAAG
B RHDV Iowa 2000	5296-335	---U--AUG--AG--G--GCC--CGAG--GCCGCAAG	---U-----A--GC--AAG--C--AGAG--GCCGCAAG
B RHDV Rossi	5296-335	---U--AUG--AG--G--GCC--CGAG--GCCGCAAG	---U-----A--GC--AAG--C--UACG--GCCGCAAG
B RHDV SD	5296-335	---U--AUG--AG--C--GCC--CACAG--GCCGCAAG	---U-----A--GC--AAG--C--AGAG--GCCGCAAG
B EBHSV	5274-313	---U--AUG--AG--GJ--GCCU--GGC--GA--GC--CCUG	---U-----A--G--A--C--U--GGC--GAC--C--CCUG
Vesivirus			
A Fe/F4	1-40	---A---A---AAU--U--A--A---UUCUCAAAC--UG--GCUUC	GUAAAGAAACUUGAGACAUUGCCUAAACUGAGCUUC
A Fe/F9	1-40	---A---A---AAU--U--A--A---UUCUCAAAC--UG--GCUUC	-----U-----U-----
A Fe/CFI	1-40	---A---A---AAU--U--A--A---UUCUCAAAC--UG--GCUUC	-----U-----U-----
A Fe/F65	1-40	---A---A---AAU--U--A--A---UUCUCAAAC--UG--GCUUC	-----U-----U-----
A PP/Pan-1	1-40	---A---A---GA--UUGAG--U--UG--CUCAAAACGCGCUCAAAA	---U--G--A--UGAG--U--G--G--CUCAAAA
A Po/A48	1-40	---A---A---GA--UUGAG--U--UG--CUCAAAACG--UG--AAAA	---U--G--A--UGAG--U--G--G--CUCAAAA
A Ca/Canine	1-40	---U---U---GA--AU--GC--UGUGCC--UGG--UC--UGAGCGCG	---U--G--G--A--GCUUAGCCA--GCU--CGAGCG
B Fe/F4	5297-338	---UUC--A--U--U--A--UGUGCU--AA--C--G--GCUAACG	---GUUC--G--G--G--GCUA--CGUG--GCUAACG
B Fe/F9	5297-338	---UUC--A--U--U--A--UGUGCU--AA--C--G--GCUAACG	---GUUC--G--G--G--GCUA--CGUG--GCUAACG
B Fe/CFI	5294-335	---UUC--A--U--U--A--UGUGCU--AA--C--G--GCUAACG	---GUUC--G--G--G--GCUA--CGUG--GCUAACG
B Fe/F65	5297-336	---UUCU--AAU--U--A--UGUGCU--AA--C--G--GCUAACG	---UUCU--AAU--U--A--UGUGCU--AA--C--G--GCUAACG
B PP/Pan-1	5649-88	UGUUGAGAGAUUA--CA--U--U--CUACUAC--AC--GCU	UGUUGAG--A--CC--U--G--ACU--AC--CGC--U
B Po/A48	5652-91	UGUUGAG--AAU--AA--CA--C--U--CUACUAC--A--GCU	UGUUGAG--AAAC--G--G--ACU--AA--G--GCU
B Ca/Canine	5787-826	---UU--GA--U--A--CJ--UGGCU--CUA--CUUG--ACUCA	---UUUU--G--A--CJ--UGGCU--CUU--A--CA

homologous regions strengthens the conclusion that the existence of B regions is meaningful for CVs. These comparisons demonstrate that if sequence identity of A and B sequences is needed for (some types of) RNA recombination to occur, with the available sequence information, recombination would be more favored between NLV strains, between *Lagovirus* strains, or between feline *Vesivirus* strains, than between strains from different hosts within a CV genus or between strains representing different genera. Genbank accession numbers for the analyzed sequences are: Hu/NV Jiang (M87661), Hu/NV Schreier (AF093797), Hu/Southampton (L07418), Hu/Chiba (AB042808), Hu/Lordsdale (NC:001674), Hu/Camberwell (NC:002614), Hu/GII Clarke (X86557), Hu GII Maryland (AY032605), Hu/Hawaii (U07611), Bo/Jena (AJ011099), Hu/Manchester (X86560), Po/Saif (AF182760), RHDV Meyers (M67473), RHDV Czech (U54983), RHDV Iowa 2000 (AF258618), RHDV Rossi (X87607), RHDV SD (Z29514), EBHSV (NC:002615), Fe/F4 (D31836), Fe/F9 (M8679), Fe/CFI (U13992), Fe/F65 (AF109465), PP/Pan-1 (AF091736), Po/A48 (U76874), and Ca/Canine (AB070225). Abbreviations used include: Hu = human, NV = Norwalk virus, GII = genogroup II, Bo = bovine, Po = Porcine, RHDV = rabbit hemorrhagic disease virus, EBHSV = European brown hare syndrome virus, Fe = feline, PP = *Pan paniscus*, and Ca = canine.

interference. CV RNA also must have an attribute that permits/favors recombination, such as a site where errors in procession of RNA polymerase can occur. The subgenomic RNA is the most likely molecule to participate in recombination as noted above. The highly conserved 5' end sequence of the genome and at the 5' end of the capsid gene in NLVs is an obvious common target for CV RNA polymerase, for genomic and subgenomic RNA synthesis.

The sequence data indicated that recombination in strain Arg320 occurred at the ORF1/capsid gene junction where high sequence identity exists between the putative parent clades. The genomic sequence of 10 NLV strains has been determined. A comparison of sequence at the 5' genomic sequence with sequence near the 5' end of ORF2 shows a high degree of sequence identity (Fig. 3). No other region of an NLV strain genome shares this degree of identity, even closely, with the 5' end of the genome. In addition, sequence identity comparable to that shown between the NLV 5' end and near the 5' end of the NLV capsid may occur among CVs of a single genus from a single host, once enough strains are sequenced.

The "copy choice" model has been preferred for recombination of single-stranded RNA viruses, including picornaviruses and coronaviruses (Kirkegaard and Baltimore, 1986; Makino *et al.*, 1986; Lai and Cavanagh, 1997; Nagy and Simon, 1997). In the copy-choice model, recombination occurs during RNA replication when the viral RNA polymerase switches templates from the RNA derived from one strain (donor template) to the RNA derived from a second strain (acceptor template), at a highly conserved genome region, without releasing the nascent strand (Lai and Cavanagh, 1997). Models for RNA virus recombination have utilized two terminologies to describe the degree that features of the donor and acceptor templates are shared: homologous, aberrant homologous, and non-homologous types (Lai and Cavanagh, 1997) or sequence similarity-essential, similarity-assisted, and similarity-nonessential (Nagy and Simon, 1997). The putative parent clades of intragenic CV recombinants have a long region of identical sequence and predicted stable hairpin structures at the proposed recombination site, which supports the classification of these recombinants as homologous or (at least) similarity-assisted. Interaction like that between genomic and subgenomic CV RNAs could occur by the same mechanisms for two genomic 5' ends, but the outcome of such recombination events would be hard to predict. Furthermore, if a virion contains genomic and subgenomic RNAs, recombination could occur in a generation after initial co-infection.

Evidence for recombination of CVs with other virus families

Evidence for recombination of CVs depends upon sequence comparisons. Upon sequencing a portion of a feline CV strain F9, Neill (1990) observed that (what later was designated) ORF1 contained significant sequence identity with picornaviruses. This significant identity was concentrated around certain amino acid motifs within ORF1 that are homologous to those within the non-structural region of picornaviruses, encoding, in order, 2C, 3C, and 3D genes. The order of these motifs and the approxi-

mate number of nucleotides between them were the same in both virus families (Fig. 4).

The capsid gene of CVs also is homologous to the VP1 to VP4 capsid proteins of picornaviruses to the extent of a shared PPG amino acid motif in a relatively conserved 5' portion of the capsid gene(s), formation of capsomeres having polypeptide β -pleated sheets as a core structural element, and formation of a spherical virion capsid by the protein(s) (Prasad *et al.*, 1999). These findings led to the hypothesis that at some point in time CVs and picornaviruses were/are "recombination partners" (Dinulos and Matson, 1994).

Fig. 4. Genome organization of CVs and picornaviruses showing the switch in order of the nonstructural and structural genes. In CVs, the nonstructural genes are encoded in ORF1; several of the genes were recognized because of the presence of significant sequence identity and consensus motifs like those known for picornaviruses (2C, 3C, and 3D). The capsid gene of CVs is encoded in ORF2 (NLVs and vesiviruses; *cf.* Fig. 1), which lies 3' to the nonstructural genes, and is marked by a "PPG" motif that signals a relatively conserved region between the families. In picornaviruses, this order of nonstructural-structural "gene cassettes" is reversed, with the order of motifs within the nonstructural peptide the same, and about the same number of nucleotides from each other in the genome. (Box shadings as in Fig. 1.)

In a recent report, Gibbs and Weiller (1999) suggested from sequence analyses that CVs (RNA genome, mostly in vertebrates) may have recombined with *Nanovirus* (ss DNA genome, plant virus) to generate (a) *Circovirus(es)* (Fig. 5). The sequence data, biologic information about host distribution of the putative parents, and pathways of DNA and RNA synthesis suggested to the authors that a *Nanovirus* crossed the invertebrate—vertebrate barrier, where, in a mixed infection, a 2C region cDNA of a CV recombined with the *Nanovirus*. A cDNA from a CV RNA limited to the 2C region was inserted into the *Nanovirus* DNA. The steps would not need to be contemporaneous, but would include: 1) *Nanovirus* infecting the vertebrate, 2) cDNA production from CV RNA, 3) excision of the 2C region, and 4) fusion of the 2C region cDNA with the ligated

Fig. 5. Suggested origin of *Circovirus(es)* from recombination of *Nanovirus(es)* with a fragment of a CV genome. In this model, *Nanovirus* crossed the invertebrate-vertebrate barrier and infected a cell containing a CV 2C gene cDNA, which was inserted into the *Nanovirus* genome to derive a *Circovirus*. Therefore, reverse transcription of the CV RNA genome must have occurred and a fragment of the CV genome have been excised, either before or after fusion with the *Nanovirus* DNA genome. (Drawn from data in Gibbs and Weiller, 1999.)

Nanovirus DNA. That this set of steps occurred is suggested by significant sequence identities of two regions of circoviruses, one including the Rep (ligation initiation) gene of nanoviruses and the other 2C-like sequences closest to those of CVs. However, a reverse-transcriptase initiation site is not known in the CV genome. The possibility that recombination occurred in invertebrates is not excluded, given the existence of viruses in insects with close sequence identity to CVs (Govan *et al.*, 2000).

Genomic recombination as a mechanism for generating CV diversity

The differences in genome organization among CV genera imply different constraints on how RNA recombination might have occurred. For example, if the different genera are derived from a single “parental” genomic structure, different events must have occurred to generate the diversity of genome structures exhibited by the different CV genera—even within genera—for some genes are absent and others present. Alternatively, if, as discussed above, CVs are “recombination partners” with (an)other virus family(ies), then “convergent evolution” might explain the shared genomic features of CVs, despite multiple “parental” genomic structures.

Recombinants extend our knowledge of the genetic diversity within CVs. They also place constraints on methods that “genotype” CVs. If the RNA recombination of Arg320 and SMV is a common phenomenon, genotyping would be more difficult. For example, many reports of CV genotyping have been based upon sequence of the RNA polymerase region, due to its relatively high sequence conservation and relative ease of designing RT-PCR primers. In contrast, fewer capsid genes have been characterized (See also Jiang, Section IV, Chapter 4 of this book). The viral capsid protein is responsible for virion antigenicity and probably for inducing immunity. Genotyping of CVs based upon the RNA polymerase sequences clearly is not the best choice if recombination at the ORF1/capsid gene is common. In addition, it remains unclear whether additional recombination sites exist. Recombination in NLVs at the ORF1-ORF2 junction has been described upon the characterization of this genomic region for relatively few strains. Thus, one might discover other types of recombinants as more strains are characterized.

Clinical and biologic implications of recombination

One recombinant NLV, Arg320, was first recovered from ill children and adults in Argentina, the United States and the Netherlands (Jiang *et al.*, 1999; Jiang *et al.*, 2000; M Koopmans, personal communication). SMV (Morens *et al.*, 1979) and many very similar strains have been recovered from outbreaks of gastroenteritis worldwide. Many SMV-like NLVs have been characterized at the genomic level only in the RNA polymerase genome region. Each of these strains is a potential SMV-like recombinant, like the prototype, awaiting sufficient characterization of the capsid sequence to draw this conclusion. The possible widespread occurrence of recombinants in symptomatic persons suggests their ready infectivity in the host(s), their easy transmissibility, furthermore that recombination does not necessarily ablate virulence, and that recombinants are genetically and ecologically stable. Perhaps the most striking feature of Arg320 and SMV is that they and their associated illness were otherwise unremarkable. Their recombinant status was recognized only because their genomes were initially characterized in both the RNA polymerase region and capsid regions. Also, the two potential parental strains for each of Arg320 and SMV are within the range of genetic diversity of strains currently co-circulating. Therefore, the recombination event could have occurred recently, but not necessarily during the infection of the child from whom Arg320 was recovered. On the other hand, it would not be difficult to imagine that many CV strains currently co-circulating could have derived from remote recombination events in the past.

Recombination may permit CVs to escape host immunity quickly, analogous to antigenic shifts in influenza viruses, but by a different molecular mechanism. Recombination during calicivirus replication may be common or rare; it is possible to envision the generation of many non-viable or attenuated recombinants. The ORF1 polyprotein genes could persist in virus with a new capsid selected for by the host's immunity. Viable recombinants could be a model for laboratory manipulation of cap-

sids (e.g., Neill *et al.*, 2000), including study of packaging constraints and antigenicity. In the two natural recombinants described above, ORF3, encoding a minor structural protein, segregated with the capsid gene in the recombinants. Whether recombination can occur with an ORF3 derived from another strain is unknown.

Summary

RNA recombination apparently contributed to the evolution of CVs. Nucleic acid sequence homology or identity and similar RNA secondary structure of CVs and non-CVs may provide a locus for recombination within CVs or with non-CVs should co-infections of the same cell occur. Natural recombinants have been demonstrated among other enteric viruses, including *Picornaviridae* (Kirkegaard and Baltimore, 1986; Furione *et al.*, 1993), *Astroviridae* (Walter *et al.*, 2001), and possibly rotaviruses (e.g., Desselberger, 1996; Suzuki *et al.*, 1998), augmenting the natural diversity of these pathogens and complicating viral gastroenteritis prevention strategies based upon traditional vaccines. Such is the case for CVs and *Astroviridae*, whose recombinant strains may be a common portion of naturally circulating strains. The taxonomic — and perhaps biologic — limits of recombination are defined by the suggested recombination of *Nanovirus* and CV, viruses from hosts of different biologic orders; the relationship of picornaviruses and CVs, viruses in different families, as recombination partners; and the intra-generic recombination between different clades of NLVs.

Acknowledgements

Thank you to my friend Xi Jiang for a special 10-year collaboration. I thank Tamas Berke for continued insights.

References

- Berke T, Golding B, Jiang X, Cubitt WD, Wolfaardt M, Smith AW, Matson DO (1997). A phylogenetic analysis of the caliciviruses. *J. Med. Virol.* **52**: 419-424.
- Carter MJ, Milton ID, Meanger J, Bennett M, Gaskell RM, Turner PC (1992). The complete nucleotide sequence of a feline calicivirus. *Virology* **190**: 443-448.
- Desselberger U (1996). Genome rearrangements of rotaviruses. *Adv. Virus Res.* **46**: 69-95.
- Dinulos MB, Matson DO (1994). Recent developments with human caliciviruses. *Pediatr. Infect. Dis. J.* **13**: 998-1003.
- Furione M, Guillot S, Otelea D, Balanant J, Candrea A, Crainic R (1993). Polioviruses with natural recombinant genomes isolated from vaccine-associated paralytic poliomyelitis. *Virology* **196**: 199-208.
- Gibbs MJ, Weiller GF (1999). Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. *Proc. Natl.*

- Acad. Sci. U.S.A.* **96**: 8022-8027.
- Govan VA, Leat N, Allsopp M, Davison S (2000). Analysis of the complete genome sequence of acute bee paralysis virus shows that it belongs to the novel group of insect-infecting RNA viruses. *Virology* **277**: 457-463.
- Gray JJ, Green J, Cunliffe C, Gallimore C, Lee JV, Neal K, Brown DW (1997). Mixed genogroup SRSV infection among a party of canoeists exposed to contaminated recreational water. *J. Med. Virol.* **52**: 425-429.
- Green J, Vinjé J, Gallimore CI, Koopmans M, Hale A, Brown DW (2000). Capsid protein diversity among Norwalk-like viruses. *Virus Genes* **20**: 227-236.
- Green KY, Ando T, Balayan MS, Clarke IN, Estes MK, Matson DO, Nakata S, Neill JD, Studdert MJ, Theil H-J (2000). Caliciviridae. In: van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB, eds. *Virus Taxonomy: Classification and Nomenclature of Viruses*. Seventh Report of the International Committee on Taxonomy of Viruses, pp. 725-735. San Diego: Academic Press.
- Hardy ME, Kramer SF, Treanor JJ, Estes MK (1997). Human calicivirus genogroup II capsid sequence diversity revealed by analyses of the prototype Snow Mountain agent. *Arch. Virol.* **142**: 1469-1479.
- Hohdatsu T, Sato K, Tajima T, Koyama H (1999). Neutralizing features of commercially available feline calicivirus (FCV) vaccine immune sera against FCV field isolates. *J. Vet. Med. Sci.* **61**: 299-301.
- Jiang X, Berke T, Zhong W, Dai X, Dunham D, Pickering LK, Matson DO (1997). Sapporo-like human caliciviruses are genetically and antigenically diverse. *Arch. Virol.* **142**: 1813-1827.
- Jiang X, Wang M, Wang K, Estes MK (1993). Sequence and genomic organization of Norwalk virus. *Virology* **195**: 51-61.
- Jiang X, Espul C, Zhong WM, Cuello H, Matson DO (1999). Characterization of a novel human calicivirus that may be a naturally occurring recombinant. *Arch. Virol.* **144**: 2377-2387.
- Jiang X, Wilton N, Zhong WM, Farkas T, Huang PW, Barrett E, Guerrero M, Ruiz-Palacios G, Green KY, Green J, Hale AD, Estes MK, Pickering LK, Matson DO (2000). Diagnosis of human caliciviruses by use of enzyme immunoassays. *J. Infect. Dis.* **181**: S349-S359.
- Kirkegaard K, Baltimore D (1986). The mechanism of RNA recombination in poliovirus. *Cell* **47**: 433-443.
- Lai MM, Cavanagh D (1997). The molecular biology of coronaviruses. *Adv. Virus Res.* **48**: 1-100.
- Lauritzen A, Jarrett O, Sabara M (1997). Serological analysis of feline calicivirus isolates from the United States and United Kingdom. *Vet. Microbiol.* **56**: 55-63.
- Liu BL, Lambden PR, Gunther H, Otto P, Elschner M, Clarke IN (1999). Molecular characterization of a bovine enteric calicivirus: relationship to the Norwalk-like viruses. *J. Virol.* **73**: 819-825.
- Makino S, Keck JG, Stohlman SA, Lai MMC (1986). High-frequency RNA recombination of murine coronaviruses. *J. Virol.* **57**: 729-737.

- Matson DO, Mitchell DK, Van R, Jiang X, Pickering LK (1995). Enteric viral pathogens as causes of outbreaks of diarrhea among children attending day care centers during one year of observation. *Pediatr. Res.* **37**:826.
- Morens DM, Zweighaft RM, Vernon TM, Gary GW, Eslien JJ, Wood BT, Holman RC, Dolin R (1979). A waterborne outbreak of gastroenteritis with secondary person-to-person spread. Association with a viral agent. *Lancet* **1**:964-966.
- Meyers G, Wirblich C, Thiel HJ (1991). Genomic and subgenomic RNAs of rabbit hemorrhagic disease virus are both protein-linked and packaged into particles. *Virology* **184**:677-686.
- Nagy PD, Simon AE (1997). New insights into the mechanisms of RNA recombination. *Virology* **234**: 1-9.
- Neill JD (1990). Nucleotide sequence of a region of the feline calicivirus genome which encodes picornavirus-like RNA-dependent RNA polymerase, cysteine protease and 2C polypeptides. *Virus Res.* **17**: 145-160.
- Neill JD (1992). Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus: identification of conserved and non-conserved amino acid sequences among calicivirus capsid proteins. *Virus Res.* **24**: 211-222.
- Neill JD, Mengeling WL (1988). Further characterization of the virus-specific RNAs in feline calicivirus infected cells. *Virus Res.* **11**:59-72.
- Neill JD, Sosnovtsev SV, Green KY (2000). Recovery and altered neutralization specificities of chimeric viruses containing capsid protein domain exchanges from antigenically distinct strains of feline calicivirus. *J. Virol.* **74**: 1079-1084.
- Prasad BV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK (1999). X-ray crystallographic structure of the Norwalk virus capsid. *Science* **286**: 287-290.
- Prasad V, Matson DO, Smith AW (1994). Three-dimensional structure of the primate calicivirus. *J. Mol. Biol.* **240**: 256-264.
- Reuter G, Farkas T, Berke T, Jiang X, Szucs G, Matson DO (2003). Molecular epidemiology of human calicivirus gastroenteritis outbreaks in Hungary, 1998 to 2000. *J. Med. Virol.*, in press.
- Rinehart-Kim J, Zhong W-M, Jiang X, Smith AW, Matson DO (1999). Complete nucleotide sequence and genomic organization of a primate calicivirus, Pan-1. *Arch. Virol.* **144**: 199-208.
- Smith AW (2000). Virus cycles in aquatic mammals, poikilotherms, and invertebrates. In: Hurst CJ, ed. *Viral Ecology*, pp. 447-491. San Diego: Academic Press.
- Sosnovtsev SV, Green KY (2000). Identification and genomic mapping of the ORF3 and VPg proteins in feline calicivirus virions. *Virology* **277**: 193-203.
- Suzuki Y, Gojobori T, Nakagomi O (1998). Intragenic recombinations in rotaviruses. *FEBS Letters* **427**: 183-187.
- Walter JE, Briggs J, Guerrero ML, Matson DO, Pickering LK, Ruiz-Palacios G, Berke T, Mitchell DK (2001). Molecular characterization of a novel recombinant strain of human astrovirus associated with gastroenteritis in children. *Arch. Virol.* **146**: 2357-2367.
- Webster RG, Laver WG, Air GM, Schild GC (1982). Molecular mechanisms of variation in influenza viruses. *Nature* **296**: 115-121.