Host-related nucleotide composition and codon usage as driving forces in the recent evolution of the *Astroviridae*

Formijn J. van Hemert, Ben Berkhout, Vladimir V. Lukashov *

Laboratory of Experimental Virology, Department of Medical Microbiology, Center for Infection and Immunity Amsterdam (CINIMA), Academic Medical Center, University of Amsterdam, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands

Abstract

The evolutionary history of the *Astroviridae* comprises the ancient separation between avian and mammalian astrovirus lineages followed by diversification among mammalian astroviruses. The latter process included several cross-species transmissions. We found that the recent, but not the ancient, evolution of astroviruses was associated with a switch in nucleotide composition and codon usage among non-human mammalian versus human/avian astroviruses. Virus and hosts phylogenies based on codon usage agreed with each other and matched the hosts' evolutionary emergence order. This recent switch in driving forces acting at the synonymous level points to the adaptation of codon usage by viruses to that of their hosts after cross-species transmissions. This is the first demonstration of nucleotide composition and codon usage being active driving forces during the recent evolutionary history of a virus group in the host–parasite system.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Astrovirus; Phylogenetic topology; (Non-)synonymous substitution; Codon preference; Virus/host species relation

Introduction

Astroviruses are small non-enveloped viruses isolated from a number of mammals and birds. Their ssRNA genome is 6.8–7.9 kb in length and contains three open reading frames: ORF1a and 1b, both coding for non-structural proteins, and ORF2, coding for structural proteins (Monroe et al., 1995). Taxonomically, astroviruses belong to the virus family of *Astroviridae* and are separated into two genera. The *Mamastrovirus* genus comprises all astroviruses of mammals: eight serotypes of human (HAstV-1 to HAstV-8 from *Homo sapiens*), feline (FAstV from *Felis catus*), porcine (PAstV from *Sus scrofa*), ovine (OAstV from *Ovis aries*), and mink (MAstV from *Mustela vison*) astroviruses. The *Avastrovirus* genus comprises all astroviruses of birds: two types of astroviruses from turkey (TAstV-1 and TAstV-2 from *Meleagris gallopavo*) and two types of avian nephritis viruses from chicken (ANV-1 and ANV-2 from *Gallus gallus*) (Lukashov and Goudsmit, 2002).

Astroviruses cause gastroenteritis in mammals, being the second most common cause of viral diarrhea among human populations in many countries, and a broader spectrum of diseases in birds, including enteritis, hepatitis, and nephritis (Matsui and Greenberg, 2001).

Our previous analysis (Lukashov and Goudsmit, 2002) of all available sequence information for astroviruses distinguished two deep-rooted groups within this family, one comprising mammalian astroviruses and the second comprising avian astroviruses. This finding led to the separation of the *Astroviridae* into two genera, whereas all astroviruses were originally classified as belonging to a single genus of *Astrovirus*. Within the group of mammalian astroviruses, ovine astrovirus and the recently sequenced mink astrovirus (Mittelholzer et al., 2003a) formed a subcluster distinct from a subcluster of human, feline, and porcine astroviruses (Lukashov and Goudsmit, 2002). This reconstruction of evolutionary events indicates that in the history of the *Astroviridae*, the separation between mammalian and avian virus lineages was followed by the consecutive divergence between the ovine/mink and human/feline/porcine virus lineages within the mammalian astroviruses, with the

* Corresponding author. Fax: +31 20 691 6531.
E-mail address: v.lukashov@amc.uva.nl (V.V. Lukashov).

recent events involving several cross-species transmissions. This order of evolutionary events within the family of *Astroviridae* was evident when nucleotide substitutions were analyzed for the full-length genomes as well as when nucleotide and non-synonymous substitutions were analyzed for each of the three ORFs. However, analysis of synonymous substitutions
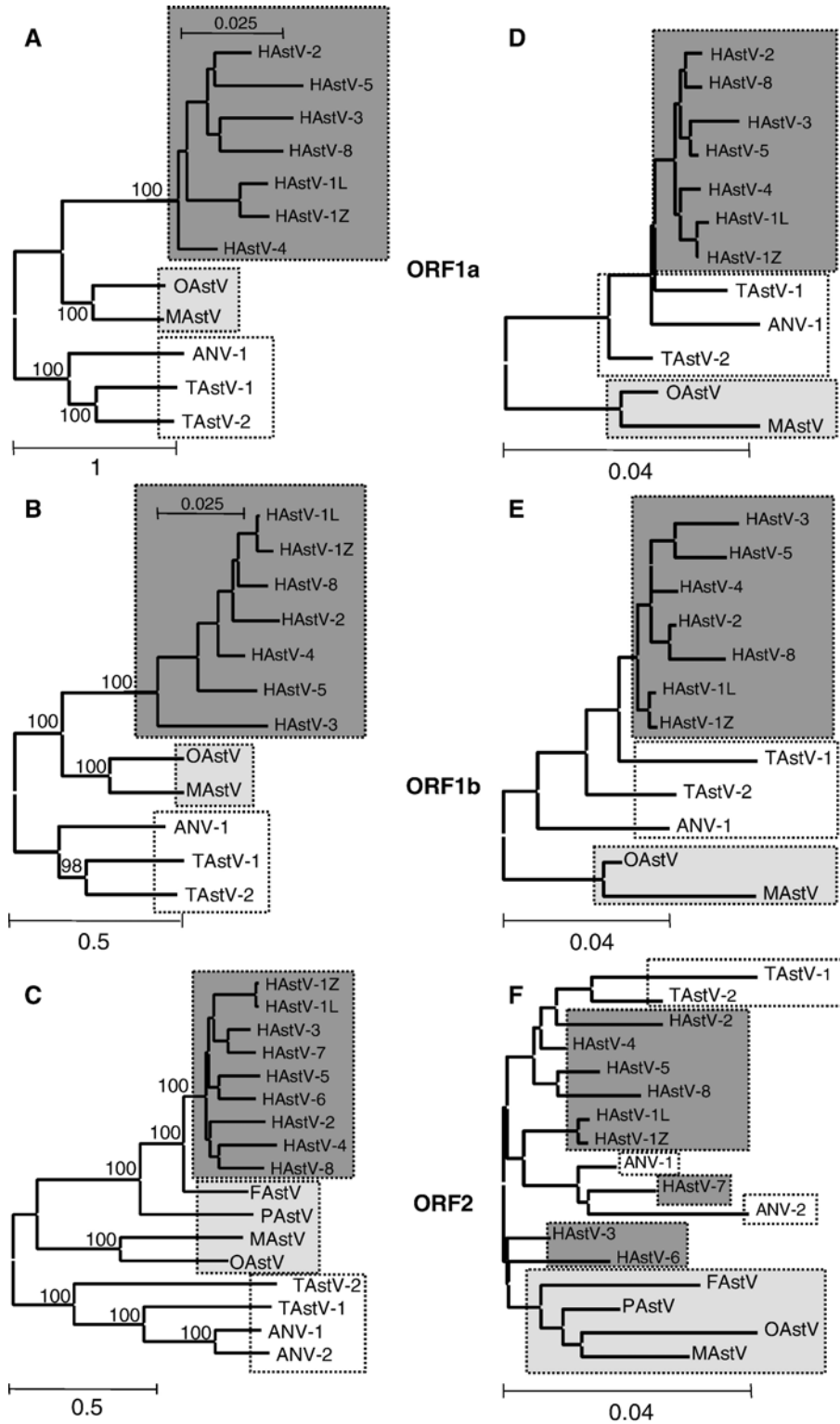

Fig. 1. Phylogenetic relationships among astroviruses based on amino acid replacements (A, B, C) and synonymous codon usage (D, E, F) in ORF1a (A, D), ORF1b (B, E), and ORF2 (C, F). Boxes indicate topological differences between the two phylogenies. The avian astroviruses are in white, the mammalian astroviruses are in gray (human astroviruses—dark gray, non-human astroviruses—light gray). The scale bars represent either the number of amino acid replacements per site or the relative difference in synonymous codon usage between the species. The bootstrap values are shown.

in all three ORFs of astrovirus genomes resulted in poorly structured trees, with avian and ovine viruses being virtually equidistant from each other and from the human viruses (Lukashov and Goudsmit, 2002). Such a loss of tree structure in analysis of synonymous substitutions has also been observed in other virus families and has been widely interpreted as the result of saturation of synonymous substitutions, since they are subject to reduced selection pressure compared to non-synonymous substitutions.

On the other hand, there are evolutionary forces that are specifically acting at synonymous positions, of which nucleotide composition and codon usage are considered to be central. However, no attempt has been undertaken so far to establish the role of these evolutionary forces in driving the evolution of a virus family. Partly this was due to the lack of data on the order of evolutionary events within a virus family. In the present study, we demonstrated that recent astrovirus evolution has been driven by evolutionary forces acting on nucleotide composition and codon usage, in the direction of codon usage of host species genes.

Results

Evolutionary relationships among astroviruses based on amino acid distances versus codon usage

Phylogenetic relationships among astroviruses based on the analysis of amino acid replacements (Figs. 1A, B, and C) was in agreement with our earlier findings based on the analysis of non-synonymous distances (Lukashov and Goudsmit, 2002). We observed an ancient separation between the mammalian and avian astroviruses, followed by a more recent separation of the ovine and mink virus lineages within mammalian viruses. For all three ORFs, the branching order of viruses was the same as previously demonstrated based on the analyses of nucleotide and nonsynonymous substitutions (Lukashov and Goudsmit, 2002).

However, analysis of codon usage among astroviruses resulted in identification of different virus groups (Figs. 1D, E, and F). In all three ORFs, human and avian viruses cluster together and separately from ovine and mink viruses. In ORF2, for which sequences of feline and porcine viruses were also available, these two viruses branched out close to the ovine/mink virus cluster. These topologies have been noticed previously in trees based on synonymous substitutions in astrovirus genes (Lukashov and Goudsmit, 2002).

Nucleotide composition

The relative proportions of the four nucleotides at the three codon positions revealed an evolutionary phenomenon that was specific for ovine and mink (and, in ORF2, also for feline and porcine) astroviruses: the strongly decreased content of A at the 3rd codon positions in all three ORFs is accompanied by the increased content of G and C (Fig. 2). A strong bias against A was observed at the 3rd codon positions of OAstV and MAstV: 17.4% and 14.3%, 18.0% and 18.7%, 17.5% and 18.2%, for
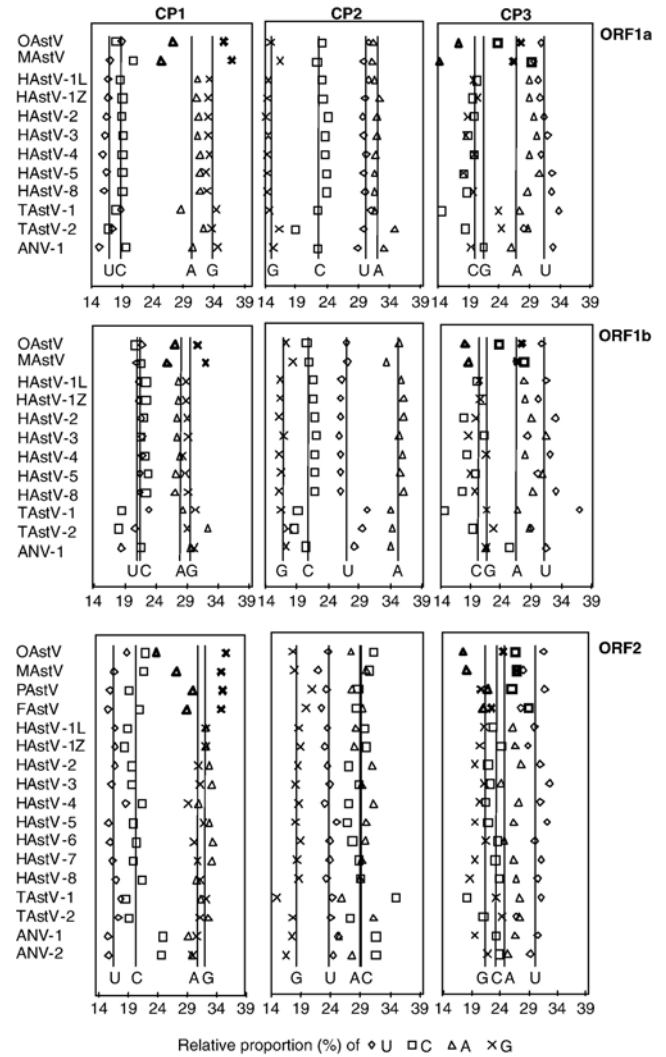

Fig. 2. Nucleotide composition of the codon positions 1, 2, and 3 (CP1, CP2, and CP3, respectively) in astrovirus ORF1a, ORF1b, and ORF2. Relative proportions of the nucleotides are given on the X-axis. Vertical lines represent the average proportion of the respective nucleotide. Symbols in bold indicate the deviant nucleotide preferences characteristic for the non-human mammalian viruses.

ORF1a, 1b, and 2, respectively. This was in contrast to the mean content of A at the 3rd codon positions of human and avian astroviruses, which was above the theoretically unbiased value of 25% and averaged 28.9% (range, 26.0–30.4%), 28.2% (range, 21.6–31.6%), and 26.4% (range, 24.1–27.6%), for ORF1a, 1b, and 2, respectively. FAstV and PAstV, for which only ORF2 sequences were available for the analysis, demonstrated a less pronounced bias against A at the 3rd codon positions with the contents of A being 21.0% and 21.8%, respectively. This is still outside the range of the human and avian viruses. The difference in the mean contents of A at the 3rd codon positions between non-human mammalian and other viruses was statistically significant ($p < 0.01$).

This evolutionary bias against A of the ovine and mink astroviruses was accompanied by a preference for G and C at the 3rd codon positions. The GC-content at the 3rd codon

positions of the ovine and mink astroviruses was markedly higher (51.5% and 55.9%, 51.2% and 54.6%, 51.0% and 53.8%, for ORF1a, 1b, and 2, respectively) than the GC-content of human and avian astroviruses, which varied in narrow intervals of 36.6–42.9%, 36.3–46.9%, and 41.3–45.5%, for ORF1a, 1b, and 2, respectively. This difference between ovine/mink and human/avian astroviruses was even more pronounced for the synonymous 3rd positions (excluding AUG and UGG codons, Fig. 3). The GC-content of FAstV and PAstV at the 3rd codon positions of ORF2 were also outside the range of human and avian astroviruses: 51.3% and 46.5%, respectively. The difference in the mean GC contents at the 3rd codon positions between non-human mammalian and other viruses was statistically significant ($p < 0.01$).
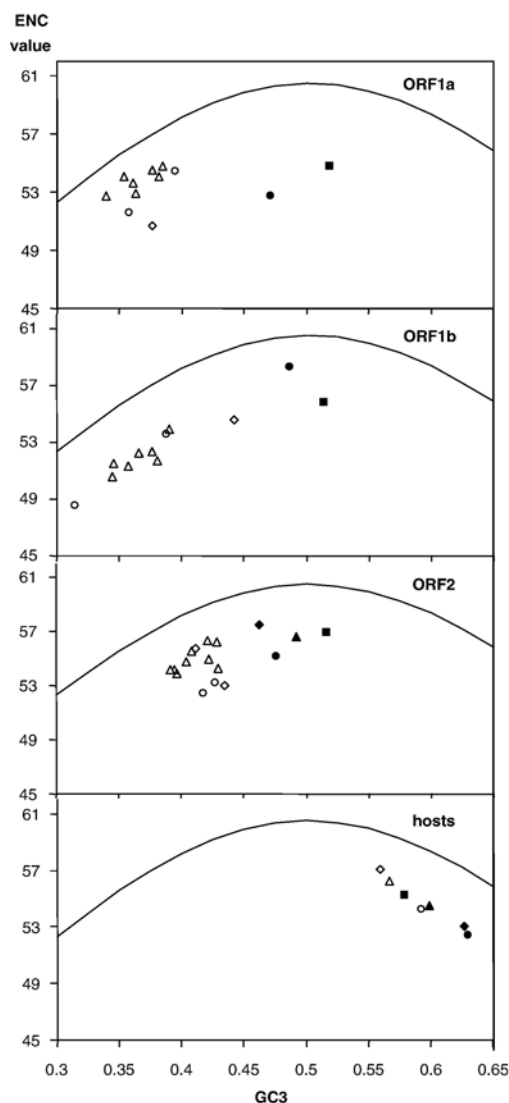

Fig. 3. The effective number of codons (ENC) in astroviruses and their hosts. ENC-values are plotted against the proportion of G+C in the 3rd synonymous codon position (GC3) for the three ORFs of astroviruses as well as for nuclear genes of their hosts (lower panel). The bell-shaped curve marks the unbiased synonymous codon usage as a function of GC3. Symbols used for the viruses and their hosts are: human—open triangles, turkey—open circles, ANV—open diamonds, mink—closed squares, sheep—closed circles, cat—closed triangles, and pig—closed diamonds.

Additionally, the non-human mammalian astroviruses were characterized by a decreased contents of A accompanied by the increased content of G at the 1st codon positions in all three reading frames, compared to human and avian astroviruses ($p < 0.01$, Fig. 2). Again, the feline and porcine viruses in ORF2 demonstrated the same phenomenon, but to an extent lesser than the ovine and mink astroviruses (Fig. 2). These patterns of nucleotide composition in ovine/mink compared to human/avian astroviruses were observed at the 1st and 3rd codon positions, which contain synonymous sites, and were absent at the 2nd position, which does not contain synonymous sites (Fig. 2).

Beside this evolutionary phenomenon that was observed for the ovine and mink, but not for the human and avian viruses, we identified trends that were common for all astroviruses in all ORFs: the increased AG content accompanied by the decreased UC content at the 1st codon position, the increased A content accompanied by the decreased G content at the 2nd codon position, and the increased U content at the 3rd codon position (Fig. 2). ORF-specific differences in nucleotide content were mainly confined to the 2nd codon position, where the content of A was the highest in ORF1b and the content of C was equal to the content of A in ORF2 (Fig. 2). The mean nucleotide content of all codon positions in astrovirus genes (25.4% U, 22.0% C, 29.3% A, and 23.3% G) was close to the unbiased distribution.

Nucleotide composition in relation to codon usage in astroviruses

To establish whether the nucleotide composition (Fig. 2) or other factors (translational selection) in the ovine and mink astroviruses can explain their codon usage (Figs. 1D–F), we analyzed the effective number of codons (ENC) in relation to the GC-content at synonymous sites at the 3rd codon positions (GC3, Fig. 3) (Wright, 1990).

This analysis revealed that the ENC-values were close to the upper possible limit for all astroviruses, indicating weak (if any) translational selection. The ENC-values varied between 48.6 and 58.3. In all three ORFs, the ovine and mink (and, in ORF2, also the feline and porcine) astroviruses clustered together and separately from the human and avian astroviruses. Yet, the ENC-values were similar in the two clusters: the projections of both clusters to the *Y*-axis were overlapping (Fig. 3). At the same time, the projections of both clusters to the *X*-axis did not overlap, indicating that the characteristics of codon usage in the ovine and mink (and, for ORF2, also in the feline and porcine) astroviruses are largely determined by a different nucleotide composition, in particular by higher GC3-values.

This point is further supported by the analysis of individual viruses. For instance, ANV-1 was an extreme member of the human/avian cluster based on the ENC-value of ORF1b (Fig. 3). This finding was in agreement with both the clustering of this virus in the phylogenetic analysis of ORF1b based on codon usage (Fig. 1E) and the nucleotide composition of ORF1b of this virus (Fig. 2).

For human astroviruses, the mean ENC-values were the highest in ORF2 followed by ORF1a and ORF1b—54.95, 53.83, and 51.96, respectively (Fig. 3).

Nucleotide composition and codon usage in the host–parasite system

To learn whether the nucleotide composition and codon usage of astroviruses are related to those in their host species, we first established the evolutionary relationships among the host species based on their codon usage (Fig. 4B). The data on codon usage in nuclear genes of the host species were obtained from the Codon Usage Database (Nakamura et al., 2000).

In the phylogenetic tree based on codon usage, the branching order of the host species was identical to that of their astroviruses (except for the MAstV, Fig. 4, B versus A). Also, the branching order of the host species in this codon tree displayed a topology similar to their emergence during evolution (Fig. 4, B versus C). The avian/mammalian divergence occurred 310 million years ago. Among the mammalian lineages in our study, the human lineage (*Omnivora*) was the first to branch out during the late Mesozoic period (Falkowski et al., 2005). The separation of the *Carnivora* into *Aeluroidea* (cat-like animals) and *Arctoidea* (bear-like animals, among which the *Mustelidae*—mink) was the third evolutionary event, and the separation between sheep and pig (both *Herbivora*) was the most recent event. Humans appeared to be the most closely related to birds in their codon usage (Fig. 4B), followed by mink and cat, with pig and sheep being the species most distantly related to birds. Remarkably, the topology of the evolutionary tree based on codon usage in astroviral genes was similar to that of their host species, and both phylogenies were congruent to the established order of the evolutionary emergence of the hosts (Fig. 4).

Our analysis of the ENC-values in relation to the GC3-values for the astroviral host species demonstrated that astroviruses and their hosts have identical ENC-values but differ with respect to the values for GC3. Humans and birds have the lowest value for GC3 (0.567) among the hosts, in agreement with the lowest GC3-value (0.40) of astroviruses isolated from those species.

Also, pig and sheep have the highest GC3-values (0.627), in line with increased GC3-values (0.50) of PAstV and OAstV (Fig. 3).

Discussion

The central role of nucleotide composition and codon usage as driving forces of evolution, virus evolution in particular, including their implications for therapy and vaccine development, is well recognized (Weiss et al., 2005). However, unlike for bacteria, yeasts, *Drosophila*, and mammals, for which relevant studies are available (Kreitman and Comeron, 1999; Moriyama and Powell, 1998; Santos et al., 2004), little is known about the power and origin of these forces in viruses. Previous studies aimed at peculiarities of nucleotide composition and codon usage that are either general for large groups of viruses, like all (human) RNA viruses (Jenkins and Holmes, 2003), all plant viruses (Dreher and Miller, 2006), segmented versus non-segmented viruses (Haseloff et al., 1984) or specific for individual virus species, like SARS coronavirus (Gu et al., 2004) and H5N1 influenza A virus (Zhou et al., 2005). These studies did not address the role of codon and nucleotide preferences as driving forces in the evolution of particular virus families, primarily due to the lack of data on their evolutionary history. For instance, the differences in codon and nucleotide preferences among the *Retroviridae* were described (van Hemert and Berkhout, 1995; Berkhout et al., 2002). Yet, pinpointing these differences to specific evolutionary events is obstructed by the lack of a reliable evolutionary history of this virus family.

The present study is the first one aimed at studying nucleotide composition and codon usage as driving forces shaping up the evolution of a virus family. To summarize our findings, several conclusions can be drawn:

(i) analysis of codon usage revealed a deviance for non-human mammalian viruses from the evolutionary history
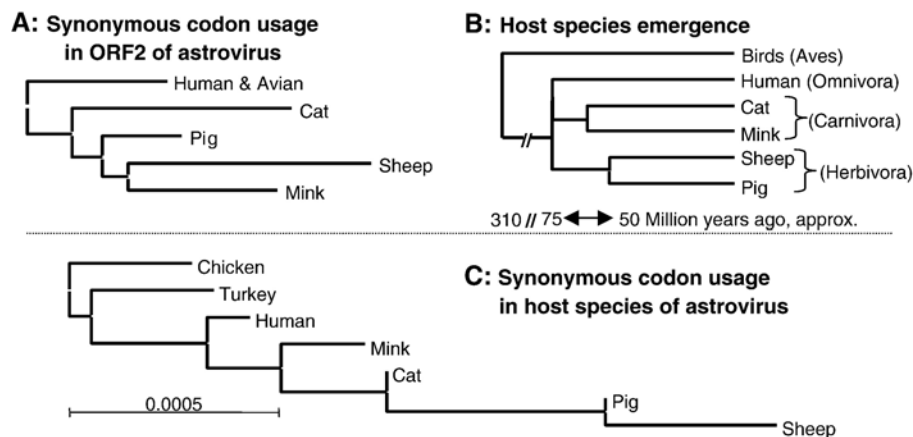

Fig. 4. Correlation between codon usage by astroviruses and their hosts, in relation to the time scale of the evolutionary emergence of host species. Panel A is a subtree of Fig. 1F representing synonymous codon usage in ORF2 of astroviruses. A neighbor-joining tree of synonymous codon usage in astrovirus host species (panel B) is constructed on the basis of codon frequencies reported in the Codon Usage Database (Nakamura et al., 2000). Chicken was defined as an outgroup. The scale bar indicates the relative difference in codon usage between the species. Host species emergence is plotted with evolutionary time (panel C, consensus topology adapted from Falkowski et al., 2005).

of the *Astroviridae*, indicating a recent switch in driving forces acting at the synonymous level;

(ii) this switch in the non-human mammalian viruses was associated with the accumulation of G and C at synonymous codon positions, rather than with translational selection;

(iii) evolutionary relationships among host species on the basis of codon usage matched the order of separation between hosts' lineages;

(iv) evolutionary relationships among astroviruses established on amino acid replacements (or non-synonymous nucleotide substitutions; Lukashov and Goudsmit, 2002) were the opposite to the order of separation of hosts' lineages; evolutionary relationships among astroviruses on the basis of codon usage matched the order of separation of hosts' lineages.

The combined data indicate that during the evolutionary history of astroviruses several cross-species transmissions did occur, most probably of a porcine-like virus to cats and then to humans. As the result of these cross-species transmissions, the astrovirus phylogeny based on the analysis of non-synonymous substitutions does not match the phylogeny of the virus hosts. After the cross-species transmissions, the viruses adapted their codon usage to that of their hosts. As the result, the virus phylogeny based on codon usage matches the hosts' phylogeny. This indicates that the driving force associated with codon usage was operational in the relatively recent diversification among mammalian astroviruses rather than in the relatively old separation between the mammalian and avian astrovirus lineages. From the point of view of classical evolution, since the relatively low GC-content was observed in both lineages (avian and mammalian), while the relatively high GC-content was observed only in the mammalian lineage, the first feature has to be considered ancestral. The influence of these distinct evolutionary forces on nucleotide composition and codon usage largely explains the topological discrepancies between phylogenetic trees based on either synonymous substitutions (or codon usage) or non-synonymous substitutions (amino acid replacements) in the *Astroviridae* (Lukashov and Goudsmit, 2002).

Separate infections of sheep and pig by an astrovirus of avian origin may be examples of ancient cross-species transmissions. Also, minks are related to cats, both being members of the *Carnivora*. In astroviral trees, however, MAstV maps near OAstV. This may point to a more recent zoonotic transfer of an astrovirus from sheep to mink, both being farm animals and, hence, MastV may be in the process of adapting its codon usage in a mink-like direction.

Codon usage in a gene may be related to the efficiency of translation. We noticed a decrease of mean ENC-values in the order ORF2 > ORF1a > ORF1b for the human astroviruses (Fig. 3, see also Simmonds, 2006). Extent of codon bias in relation to the level of expression of virus genes in a host–parasite system has been described previously (van Hemert and Berkhout, 1995; Berkhout et al., 2002). A preference for a subset of the available codons influences either positively (Bennetzen and Hall, 1982;

Sharp and Li, 1987) or negatively the rate of translation by competition for iso-accepting tRNA species. Negative effects on the rate of protein synthesis have been demonstrated for poliovirus by deoptimization of synonymous codons (Burns et al., 2006; Mueller et al., 2006). The reduction of virus particles produced by these mutants was limited to one order of magnitude, whereas viral infectivity decreased by several orders of magnitude. Synonymous substitutions, which do not change amino acids, nevertheless do affect the infectivity of polioviruses. Also, *in vitro* translation of codon-deoptimized virus genes is severely affected only in non-dialyzed extracts without exogenous additions and in the presence of competing cellular mRNAs (Mueller et al., 2006), pointing to a crucial role in the shutoff of cellular protein synthesis induced *in vivo* by the virus (Sanchez et al., 2003). Astrovirus triggers an apoptotic pathway in host cells (Guix et al., 2004). In MCF7 cells, the mRNA proportion associated with polysomes is reduced to 3% within 4 h after induction of apoptosis (Bushell et al., 2006). In addition, RNA editing by adenosine deamination cannot be excluded as a source of sequence diversification in viruses (Bass, 2002). Species-specific characteristics of ADAR enzymes relevant for virus hosts have not yet been documented.

In conclusion, our study is the first to address which evolutionary events within a virus family are driven by forces operational at the level of synonymous substitutions: nucleotide composition, translational selection, and codon usage. We demonstrated that the driving force associated with codon usage was operational in the relatively recent (diversification among mammalian astroviruses) rather than in relatively ancient evolutionary history (separation between the mammalian and avian astrovirus lineages). The switch in codon usage during diversification of mammalian astroviruses (human versus non-human astroviruses) was related to their different nucleotide composition, rather than caused by translational selection, and pointed towards the codon usage of the host species.

Materials and methods

Sequences

Sequences of astroviruses were obtained from GenBank and annotated previously (Lukashov and Goudsmit, 2002). In addition, recently obtained sequences of mink (Mittelholzer et al., 2003b) and human serotype 4 and 5 (Silva et al., 2006) astroviruses were used in the analysis. Sequences of all three ORFs were available for HAstV-1L, 1Z, 2, 3, 4, 5, and 8 (GenBank accession numbers L23513, Z25771, L13745, AF141381, DQ070852, DQ028633, and Z66541, respectively), OAstV (Y15937), MAstV (AY179509), TAstV-1 (Y15936) and TAstV-2 (AF206663), and ANV-1 (AB033998). HAstV-1L differs from HAstV-1Z by the absence of the oligopeptide deleted upon cell culture adaptation of the virus (Willcocks et al., 1994). Sequences of ORF2 were available for HAstV-6 (AB031031) and 7 (Y08632), FAstV (AF056197), PAstV (AB037272), and ANV-2 (AB046864). With respect to ORF2 of human astroviruses, their representation in Figs. 1–4 is limited to one randomly selected sequence per serotype. Other

sequences displayed similar patterns (data available upon request).

Phylogenetic analysis

Amino acid sequences were aligned by using ClustalW (Thompson et al., 1994). Phylogenetic analyses were performed by using MEGA3 software (Kumar et al., 1994). Neighbor-joining trees were constructed based on amino acid replacements (Poisson correction, pairwise gap exclusion). Statistical robustness of phylogenetic trees was assessed by bootstrap resampling (500 replications).

Nucleotide composition and codon usage

Nucleotide composition at the 1st, 2nd, and 3rd codon positions as well as codon usage was calculated using MEGA3 software. Nucleotide composition of synonymous sites was calculated for the 3rd position of all codons except AUG (Met) and UGG (Trp), which do not have synonymous codons as they are encoded by a single codon each.

Codon usage data for the nuclear genes of the host species were obtained from the Codon Usage Database (Nakamura et al., 2000). The data were available for the following numbers of codons (number of genes in parentheses): 36,349,745 (84,949) for *H. sapiens*, 111,549 (289) for *F. catus*, 798,957 (2,028) for *S. scrofa*, 194,104 (612) for *O. aries*, 10,534 (19) for *M. vison*, 23,451 (79) for *M. gallopavo*, and 2,541,006 (5579) for *G. gallus*.

For both the astroviruses and their host species, codon usage data were converted into distance matrices according to Long and Gillespie (1991) by using a Delphi-based interface. Neighbor-joining trees based on codon usage distance matrices were built with MEGA3 software.

While codon usage is associated with nucleotide composition, this association is not absolute: the absence of nucleotide bias at the 3rd codon position (the contents of U, C, A, G are 25% each) is not equivalent to the unbiased codon usage. Next to the base content, translational selection can influence the codon usage in a gene. Several methods exist to interpret data in a codon usage table (Comeron and Aguade, 1998). A single value can be calculated either for the use of a "preferred set of codons" or for the divergence from completely random codon usage. The measurement of the effective number of codons (ENC) belongs to the second method and produces a single ENC-value for each gene that is independent of its length and amino acid composition (Wright, 1990). An ENC-value can vary between the value of 61, which indicates a completely random, unbiased usage of all 61 existing sense codons, and the value of 20, which indicates the absolute preference, maximally possible bias, when only a single codon is used for each of the 20 amino acids. The upper limit of the ENC-value of 61 holds for the case when all four nucleotides are present in equal proportions (25% each). However, a bias in nucleotide composition limits the randomness of codon choice. For instance, if G and C were absent at the 3rd codon position at all (GC-content is equal to 0), there would be only 30 sense codons with A or T at the 3rd position, and the maximal ENC-value is 30. On the other hand, with only G and C present at the 3rd codon position (GC-content is equal to 1), the maximum of the ENC-value is 31.

Codon usage was characterized according to Wright (1990) by means of characterizing the ENC-values of astrovirus genes versus their GC-content at the 3rd synonymous (excluding AUG (Met) and UGG (Trp)) codon positions (GC3-values)—the "Nc-plot". A continuous line indicates theoretical ENC-values with random codon usage as a function of GC3. Deviation from this line in the direction of lower ENC-values points to translational selection acting in favor of a preferred set of codons, as has been described for highly expressed genes in yeast (Bennetzen and Hall, 1982) and *Escherichia coli* (Sharp and Li, 1987). In the coding region of 50 viral RNA genomes, the ENC-values ranged from 38.9 to 58.3 and the GC3-values from 0.21 to 0.80 (Jenkins and Holmes, 2003). A value for GC-content at the 3rd codon position includes all codons while a value for GC3 is confined to the 3rd position of synonymous codons being all codons except AUG for Met and UGG for Trp.

The statistical significance of nucleotide composition differences was assessed by *T*-test (SPSS software, SPSS Inc., Chicago, IL).

Acknowledgments

References

Bass, B.L., 2002. RNA editing by adenosine deaminases that act on RNA. Annu. Rev. Biochem. 71, 817–846.

Bennetzen, J.L., Hall, B.D., 1982. Codon selection in yeast. J. Biol. Chem. 257, 3026–3031.

Berkhout, B., Grigoriev, A., Bakker, M., Lukashov, V.V., 2002. Codon and amino acid usage in retroviral genomes is consistent with virus-specific nucleotide pressure. AIDS Res. Hum. Retroviruses 18, 133–141.

Burns, C.C., Shaw, J., Campagnoli, R., Jorba, J., Vincent, A., Quay, J., Kew, O., 2006. Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the capsid region. J. Virol. 80, 3259–3272.

Bushell, M., Stoneley, M., Kong, Y.W., Hamilton, T.L., Spriggs, K.A., Dobbyn, H.C., Qin, X., Sarnow, P., Willis, A.E., 2006. Polypyrimidine tract binding protein regulates IRES-mediated gene expression during apoptosis. Mol. Cell 23, 401–412.

Comeron, J.M., Aguade, M., 1998. An evaluation of measures of synonymous codon usage bias. J. Mol. Evol. 47, 268–274.

Dreher, T.W., Miller, W.A., 2006. Translational control in positive strand RNA plant viruses. Virology 344, 185–197.

Falkowski, P.G., Katz, M.E., Milligan, A.J., Fennel, K., Cramer, B.S., Aubry, M.P., Berner, R.A., Novacek, M.J., Zapol, W.M., 2005. The rise of oxygen over the past 205 million years and the evolution of large placental mammals. Science 309, 2202–2204.

Gu, W., Zhou, T., Ma, J., Sun, X., Lu, Z., 2004. Analysis of synonymous codon usage in SARS Coronavirus and other viruses in the Nidovirales. Virus Res. 101, 155–161.

Guix, S., Bosch, A., Ribes, E., Dora, M.L., Pinto, R.M., 2004. Apoptosis in astrovirus-infected CaCo-2 cells. Virology 319, 249–261.

Haseloff, J., Goelet, P., Zimmern, D., Ahlquist, P., Dasgupta, R., Kaesberg, P., 1984. Striking similarities in amino acid sequence among nonstructural proteins encoded by RNA viruses that have dissimilar genomic organization. Proc. Natl. Acad. Sci. U.S.A. 81, 4358–4362.

Jenkins, G.M., Holmes, E.C., 2003. The extent of codon usage bias in human RNA viruses and its evolutionary origin. Virus Res. 92, 1–7.

Kreitman, M., Comeron, J.M., 1999. Coding sequence evolution. Curr. Opin. Genet. Dev. 9, 637–641.

Kumar, S., Tamura, K., Nei, M., 1994. Molecular evolutionary genetics analysis (MEGA). Comput. Appl. Biosci. 10, 189–191.

Long, M., Gillespie, J.H., 1991. Codon usage divergence of homologous vertebrate genes and codon usage clock. J. Mol. Evol. 32, 6–15.

Lukashov, V.V., Goudsmit, J., 2002. Evolutionary relationships among *Astroviridae*. J. Gen. Virol. 83, 1397–1405.

Matsui, S.M., Greenberg, H.B., 2001. Astroviruses. In: Knipe, D.M., Howley, P.M., et al. (Eds.), Fields Virology. Lippincott-Raven Publishers, Philadelphia, pp. 875–893.

Mittelholzer, C., Englund, L., Hedlund, K.O., Dietz, H.H., Svensson, L., 2003a. Detection and sequence analysis of Danish and Swedish strains of mink astrovirus. J. Clin. Microbiol. 41, 5192–5194.

Mittelholzer, C., Hedlund, K.O., Englund, L., Dietz, H.H., Svensson, L., 2003b. Molecular characterization of a novel astrovirus associated with disease in mink. J. Gen. Virol. 84, 3087–3094.

Monroe, S.S., Carter, M.J., Herrmann, J.E., Kurtz, J.B., Matsui, S.M., 1995. Family Astroviridae. In: Murphy, F.A., Fauquet, C.M., Bishop, D.L., Ghabrial, S.A., Jarvis, A.W., Martelli, G.P., Mayo, M.A., Summers, M.D. (Eds.), Virus taxonomy: sixth report of the International Committee on Taxonomy of Viruses. Springer-Verlag, Vienna, Austria, pp. 364–367.

Moriyama, E.N., Powell, J.R., 1998. Gene length and codon usage bias in *Drosophila melanogaster*, *Saccharomyces cerevisiae* and *Escherichia coli*. Nucleic Acids Res. 26, 3188–3193.

Mueller, S., Papamichail, D., Coleman, J.R., Skiena, S., Wimmer, E., 2006. Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. J. Virol. 80, 9687–9696.

Nakamura, Y., Gojobori, T., Ikemura, T., 2000. Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res. 28, 292.

Sanchez, G., Bosch, A., Pinto, R.M., 2003. Genome variability and capsid structural constraints of hepatitis A virus. J. Virol. 77, 452–459.

Santos, M.A., Moura, G., Massey, S.E., Tuite, M.F., 2004. Driving change: the evolution of alternative genetic codes. Trends Genet. 20, 95–102.

Sharp, P.M., Li, W.H., 1987. The Codon Adaptation Index—A measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 15, 1281–1295.

Silva, P.A., Cardoso, D.D., Schreier, E., 2006. Molecular characterization of human astroviruses isolated in Brazil, including the complete sequences of astrovirus genotypes 4 and 5. Arch. Virol. 151, 1405–1417.

Simmonds, P., 2006. Recombination and selection in the evolution of picornaviruses and other mammalian positive-stranded RNA viruses. J. Virol. 80, 11124–11140.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

van Hemert, F.J., Berkhout, B., 1995. The tendency of lentiviral open reading frames to become A-rich: constraints imposed by viral genome organization and cellular tRNA availability. J. Mol. Evol. 41, 132–140.

Weiss, R., Hammerl, P., Hartl, A., Hochreiter, R., Leitner, W.W., Scheiblhofer, S., Thalhamer, J., 2005. Design of protective and therapeutic DNA vaccines for the treatment of allergic diseases. Curr. Drug Targets. Inflamm. Allergy 4, 585–597.

Willcocks, M.M., Ashton, N., Kurtz, J.B., Cubitt, W.D., Carter, M.J., 1994. Cell culture adaptation of astrovirus involves a deletion. J. Virol. 68, 6057–6058.

Wright, F., 1990. The 'effective number of codons' used in a gene. Gene 87, 23–29.

Zhou, T., Gu, W., Ma, J., Sun, X., Lu, Z., 2005. Analysis of synonymous codon usage in H5N1 virus and other influenza A viruses. Biosystems 81, 77–86.