Journal of General Virology (2000), **81**, 2891–2897. Printed in Great Britain SHORT COMMUNICATION

Molecular analysis of a serotype 8 human astrovirus genome

Martha Méndez-Toss, Pedro Romero-Guido, Maria Elena Munguía, Ernesto Méndez and Carlos F. Arias

Departamento de Genética y Fisiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apartado Postal 510-3, Colonia Miraval, Cuernavaca, Morelos 62250, Mexico

Human astroviruses are an important cause of gastroenteritis. As part of a molecular epidemiological study carried out in Mexico a human astrovirus isolate, Yuc-8, was adapted to grow in CaCo-2 cells, and its entire genome was sequenced. A 15 amino acid deletion in ORF1a, which has been associated with adaptation of astroviruses to grow in cells other than CaCo-2, was present in Yuc-8. Comparative sequence analysis of the Yuc-8 ORF2 with reported human astrovirus sequences revealed that this isolate belongs to genotype (serotype) 8. Two distinct domains in ORF2 were observed: an amino-terminal domain (residues 1 to 415), with identities higher than 81% among the strains analysed, and a carboxy-terminal domain (residues 416 to 782) with identities between 36 and 60%. Two non-superimposable phylogenetic trees were generated by separate analysis of these two domains, suggesting that a differential selective pressure is exerted along the structural polyprotein.

Human astroviruses are recognized as an important cause of infantile gastroenteritis around the world (Herrmann *et al.*, 1991). Astrovirus virions are formed by a non-enveloped protein capsid which surrounds a genome consisting of a positive-sense, single-stranded RNA molecule of 6·8 to 7·2 kb in length (Jiang *et al.*, 1993). The genomic RNA has three open reading frames (ORFs) designated 1a, 1b and 2. ORF1a and ORF1b code apparently for the nonstructural proteins of the virus; ORF1a contains viral serine protease and nuclear localization signal motifs; ORF1b contains motifs suggestive of an RNA-dependent RNA polymerase (Jiang *et al.*, 1993; Willcocks *et al.*, 1994*b*). On the other hand, ORF2 encodes the structural viral polypeptides (Sanchez-Fauquier *et al.*, 1994). The virus genome, which functions as the mRNA for the nonstructural proteins, is translated into two polyproteins. One comprises exclusively ORF1a, while the other includes both ORF1a and ORF1b (Willcocks *et al.*, 1999), and is generated as result of a translational frameshift occurring between these two ORFs (Lewis & Matsui, 1997; Marczinke *et al.*, 1994). It is thought that both nonstructural polyproteins are auto-processed by a viral protease encoded in ORF1a (Willcocks *et al.*, 1999). On the other hand, the structural proteins coded by ORF2 are translated from a subgenomic RNA synthesized during replication of the virus. A 86 to 90 kDa polyprotein is produced that is finally processed into three to five mature polypeptides (Bass & Qiu, 2000; Sanchez-Fauquier *et al.*, 1994). The subgenomic RNA is 3′-collinear with the genomic RNA and both are polyadenylated (Monroe *et al.*, 1991, 1993).

Based on the virus reactivity with polyclonal antibodies, human astroviruses (HAstVs) have been classified into seven serotypes (HAstV-1 to -7) (Herrmann *et al.*, 1988). Phylogenetic grouping of HAstV based on nucleotide sequence analysis of a limited region of ORF2 (Belliot *et al.*, 1997; Noel *et al.*, 1995) showed a good correlation between genotype and serotype, and recently the existence of an eighth HAstV serotype has been suggested based on genotyping of an HAstV strain isolated in UK (Belliot *et al.*, 1997). Only three additional strains of this genotype 8 astrovirus, isolated in Australia, Uganda and Ghaza, have been reported, which is why it is considered to be a rare serotype (Mustafa *et al.*, 2000). In a recent multicentric study carried out in Mexico (unpublished), it was found that six HAstV serotypes (1 to 4, 6 and 8) co-circulated in a single period of 6 months, with serotype 1 being the most frequent, as found in other studies from different geographical regions (Gaggero *et al.*, 1998; Noel & Cubitt, 1994; Palombo & Bishop, 1996; Shastri *et al.*, 1998). In one of the locations (Mérida City, Yucatan) included in the Mexican study, three of the eight HAstV strains detected belonged to genotype 8 (assignment made based on the sequence of the 3′-end of ORF2), suggesting that these strains might be more epidemiologically relevant than previously recognized. In this work, we adapted one of the genotype 8

Author for correspondence: Ernesto Méndez.
Fax +52 73 17 2388. e-mail ernesto@ibt.unam.mx

The GenBank accession number of the sequence reported in this paper is AF260508.

(*a*)

```
Yuc-8     740 GDDVEFDYTEVINFDQAKPTPAPRTTKP-------KPCPEPKIEAQPLDLSQKKEKQP----------------EHEQQVAKPTKPQKIEPQPYSQTYGKA 818
HAstV-1   739 ....................I.......Q--------.I....EV.S............SEYEQQVVKSTKPQQL......V..I....S............ 832
HAstV-2   739 ............................-------......ET.T.............----------------......V.S.....N............ 817
HAstV-3   739 ........................S..QQANTSQ..R..LEA..............----------------..........M..N............ 824
```

in/del I **in/del II**

(*b*)

```
Yuc-8     MASKSDKQVTVEVNNN--GRSRSKSRARSQSRGRGRSVKITVNSHNKGRRQNGRNKYQSNQRVRKIVNKQLRKQGVTGPKPAICQTATATLGTIGSNTTGATEIEACILLNPVLVKDATG 118
HAstV-8   ................--.....................................................N.................G..................................... 118
HAstV-1   .....N.......S..--..N.............DK........R.RA...P..D.R..S....N.....................R.......V....S.T..................... 118
HAstV-2   .............N..--..N.....................................................................R.........T.......................... 118
HAstV-3   .............K..NN..N..R...........K........RSR.......D.........N..T.................R.............S.T..................... 120
HAstV-4   .....H.E.....SST--....R..............SN..GSP..........T.......................................................................... 118
HAstV-5   ....PS..........---.....R..P.......DK........R...........H.......N.....................R.......V....S.T....G.............. 117
HAstV-6   ...............T--..G.......................Q...........R.......N.....................K.............S.T..............E.... 118

Yuc-8     STQFGPVQALGAQYSMWKLKYLNVRLTSMVGASAVNGTVVRISLNPTSTPSSTSWSGLGARKHLDVTVGKNAVFKLKPSDLGGPRDGWWLTNTNDNASDTLGPSIEIHTLGQTMSSYQNT 238
HAstV-8   ...........S.............M..V....T..DAI..............N.....................I.................................D.....H......... 238
HAstV-1   .........................K..............L.V.................................T.....................................R.....K.E 238
HAstV-2   .........................K..............L.....................M.....R......R.....................................K.....K.E 238
HAstV-3   .........................K..................V.....................................A.....................V......M.....K.E 240
HAstV-4   .......PT......I.....................................................A.......................................................... 238
HAstV-5   .........................K......S....D....V.....................................................................R.......Q 237
HAstV-6   .......................................V.............................I.............................V......R.......Q 238

                                                                                                                    VR1                                  *      *
Yuc-8     QFTGGLFLVELSSAWCFTGYAANPNLVNLVKSTDKSVNVTFEGSAGTPLIMNVPEHSHFARTAVEHSSLSTSLSRAGGESSSDTVWQVLNTAVSAAELVTPPPFNWLVKGGWWFVKLIAG 358
HAstV-8   .......W...T.........G..............D..................................................................................... 358
HAstV-1   ...........A.E.....................NQ.P........S......S.G......VLAR.TTP.T.A...ERTT..................................... 358
HAstV-2   ...........A.E.....................HE........K........A.......M.EQ...I..TF.....DAT..................A.......I............ 358
HAstV-3   ...........A.E.....................Q........N......K...A........AR..VP.T.A.T.QNTT..................................... 360
HAstV-4   ................................AR..D.......................M.....................NCL..................................... 358
HAstV-5   ...........A.E..............M.....Q.....N....E..V.S..AT...V.AV.AR.T.P...A...ERTT..................................... 357
HAstV-6   ...........A.E.....................Q.............V....AA......VAQR.T.P..MA...ENTA..................................... 358
```

Fig. 1. For legend see facing page.

```
           *  ↓  *                  *         VR2↓
Yuc-8      RARTGARRFYVYLSYQDALSNKPALCTGGVPASARQSNPVRTTLQFTQMNQPSLGHGATPMTFGRSIPEPGEQFRVLLTVGPPMAPNTANSQNWVNKTIVPP---ENQYTVKIGIDLEHY 475
HAstV-8    ...........P..............................................................................................---......R........ 475
HAstV-1    .T...S.S....P..............ST.GGM.TR...T...............EA.AAI.....A...EYK.V..F.S..S..AN.K.T....PLDA.---SGH.N...AK.VD.. 475
HAstV-2    .T...TKQ....P................TG-GVLRTTPV...............EHTA.I.SIVQD.SGEL.......SI.S..S.DR.V.L...LTA.GTNS.DNL...AH..G.. 477
HAstV-3    .....T.S....P..............SA.SGM.TRAA.M...............EMTA.L..A..S..DN.N.V..I.Q.L....LTK...L...T.S.---Q.HHV...AK.VNN. 477
HAstV-4    ...........P................S.YT........................A.L...V....D..K.IM...ALVQ..RSDT...LFT.VT..TG---HDAARV.WNTQ.. 475
HAstV-5    .T.N.T.S....P...............L.SGL.ATSA.T...............ENTA.L....ATA.DRLK.I....Q.VT..EN.K....G..NT.A-----EEVI...VNTQS. 472
HAstV-6    .V.N.N.S....A..............STSG.M.TRPA.M...............E..A.L.....TS..TLK......N.IS..ET.K.T......E..-----GAV.R..R.TQ.. 473

Yuc-8      TTMQGFTPVESVSWYTADFQPSDEPSPIPGLYARVNNTKKADVYGVQQFKSSHTNNRHQITSVFLVRVTTSFQVINYTSYFIRGAESGSNVSNLKIRDQTYHTPLQFTQGKWYLLTSTVM 595
HAstV-8    ........................................................................................V........................... 595
HAstV-1    L......SIA..D...I.....EA.A..K..QVL..IS......A.K..VTAQ...K..V..L...K...G...N..L...Y.ASAT.DATT..LV.GD..TAGIS....G.....N.SI 595
HAstV-2    LI....MHIKT.E...P.....RD.T..A.MSVM..I.......FMK...N.Y.........I..IKPLAD.K.QC.M...K.ESHDNDG.A..TV.SM.SPETIR.QV.E.......TL 597
HAstV-3    ........IS..D...T.....EA.T.....QVL..SS......AI..YLNNQ...KV.L..I...K.......N..L...Y.SYGT.TT.E...V.SD.TAQDVN.PV.-...M...AI 596
HAstV-4    L.I...LLID.LE.L.PNL.E.QE.PS..E.GVY.GIH...L..FM..YVNP....K..VS.I..IKPTEN.S.T..M...F.ESQ..Q..A.....P..WQQTVN.QR.....V.N.AI 595
HAstV-5    NV.N...MIS.ID..DEEM..LEV.A.MSE.LVMKGIN......AA..Y.N.IS..K......Y.I..KEN...T.HL...F.EKVDTTATELM...P.....TVN.V.DN........L 592
HAstV-6    C.LN...LITK.D.F.EE....E..A.VQ..MVLGD.H......AA..Y.NPI..DKQEV........NEG...T.HL...Y.NSVNTDA.E.I...SA.R..TVR.Y..S........L 593

Yuc-8      HDGPTSSGWVWMNQELTNNIAYRVDPGMMYLITPPPAASQLYFELHTVLPQARSEEPETYVD-----------APLPEEPPIEEEETDSDFESTEDENDEVDRFDLHPSS-ESDDDDVEN 703
HAstV-8    ...................V..I....TH...........R....A...V..........-----------.L..V..SLV.D...........DI...........-......... 703
HAstV-1    V..AMPP....N.V..KT.T..HM.K.LVH..M.L.EST.MCY.ML.SI.RS.ASGHGYES.----NIEYLDAPDSAD-QFK.DI...T.I......D..A....IIDT.--DEE.GN.T 708
HAstV-2    KENNLPE....DRV..KSDTP.YA.QALT.F.....VD..IL..GN.T..RIS.PPDNPSGRYMESHQQDCDSSDDEDDCENVS...ETED.ED...D..A......SPYSSEPE.SD.. 717
HAstV-3    FNA.APP..I.Q.V..L.DT..LI.Q...H..M....NT..L..MR.SISGS..MISSEEP.THEPGDEWCDALDASD-SRVLL....YED.ED...D..A......S.YGSEPE..D.. 715
HAstV-4    RN..PP.....D.I....ESI.YA.QVLAHF.N...QN.KI...V..TM..SSVRNIGLEE.-----------------Q----TDNWQEPDEDVQTST.ESDYETDSLEG....E.SNT 694
HAstV-5    .T.SLP.........HM..EN.I..Q..KH.M.T..VS......ML.S....KA.YLDQGDQ----AVVTYDS---.GQALFSA.............T.........L..-......... 704
HAstV-6    .T..PV...L......Q.DQ..II.Q.I.H......VS..I...MA.LV..T..GGG..GLE----LVMGLSDDEY.ISHVND....EYET..D...T........CCT.-D.E..-I.. 707

Yuc-8      DRATLLSTLLNQGISVERATRITNGAFP----TRAARVRRSVYNDLLVSGLSPGAAWSHACEQARRAGDNHDLQLSGSRDHAE 782
HAstV-8    G.R........................----P.TD.M............................................ 782
HAstV-1    ..V......V...MTMT.....ARR...----.LSD.IK.G..M......V...N.......E..K.VGETNPCT....G... 787
HAstV-2    N.V......I...MT.........KR...----.C.EKLK....M...A..A..SS...N..DE..NV.S.QLAK...D.G... 796
HAstV-3    N.V......I...MT.........KR...----.S.GKTK....M...A......N.......E..IM.I.QTPNV..D.G... 794
HAstV-4    C.ELVIN..V.....R....Y.GMS.Y.NVEWGSGEQSTSQHIQEISSDDVGAGAHY.CV..RKQQ------SLNQ...G... 771
HAstV-5    N.V......I...M......K..SN...----.....L..........V.D.D..........KV.......T....A... 783
HAstV-6    N.V......I...VT.D...M...R...----.PNYKP..EPS....APSDCLAT.R....NET--------C......G... 778
```

Fig. 1. (*a*) Alignment of a fragment of the ORF1a sequence of Yuc-8 with reported sequences of human astroviruses of serotypes 1, 2 and 3, around the variable region which contains short in/del regions (representing inserted or deleted sequences, respectively, among the HAstV strains analysed). In/del regions I and II are marked. Accession numbers of the compared sequences are: Z25571 (HAstV-1), L13745 (HAstV-2) and AF141381 (HAstV-3). (*b*) Alignment of the complete Yuc-8 ORF2 amino acid sequence with sequences of HAstV of several serotypes. Variable regions VR1 and VR2 in the amino-terminal 415 amino acids of the polyprotein are underlined. The arrows show the arginine residues which are cleaved to yield VP79, VP29 and VP26, according to Bass & Qiu (2000), and Sanchez-Fauquier *et al.* (1994). Asterisks above the sequence denote the conserved basic amino acid residues susceptible to trypsin (see text). Alignment was made by ClustalW analysis (http://www.ebi.ac.uk/clustalw/) of sequences with the following accession numbers: L23513 (HAstV-1), A45695 (HAstV-2), AF141381 (HAstV-3), Z33883 (HAstV-4), U15136 (HAstV-5), Z46658 (HAstV-6) and Z66541 (HAstV-8). HAstV-7 (accession number L38508) was not included in this analysis since only 118 of the 782 amino acid residues of the complete ORF2 have been reported. In (*a*) and (*b*), dots indicate amino acid sequence identity, and only amino acid changes are marked; a dash denotes a deletion. The single-letter amino acid code is used.

Table 1. Percentage amino acid identity between human astrovirus structural proteins of different serotypes

Numbers are percentage identity acording to a ClustalW analysis. For simplicity, strains of the same serotype were named A, B, C and D. Accession numbers for sequences used to generate the data in this table are Z25771, S68561 and L23513 for strains of HAstV-1; L41395, L06802 and L13745 for strains of HAstV-2; AF141381 and AF117209 for strains of HAstV-3; Z33883 and U15136 for strains of HAstV-4 and HAstV-5, respectively; Z46658, AB031031, AB031030 and AB013618 for HAstV-6 and Z66541 for HAstV-8. The Yuc-8 sequence was determined in this work. Bold letters denote identities between proteins of viruses belonging to the same serotype.

		Amino acid percentage identity															
Serotype / Strain		Yuc-8	8	1			2			3		4	5	6			
				A	B	C	A	B	C	A	B			A	B	C	D
Yuc-8		**100**															
HAstV-8		**94**	**100**														
HAstV-1	A	70.9	68.5	**100**													
	B	70.9	68.5	**100**	**100**												
	C	70.8	68.5	**97.6**	**97.3**	**100**											
HAstV-2	A	69.6	67.8	68.3	68.3	69.3	**100**										
	B	69.5	67.6	68.1	68.1	69.1	**99.9**	**100**									
	C	69.6	67.8	68.3	68.3	69.3	**100**	**99.9**	**100**								
HAstV-3	A	71.6	69.3	75.3	75.3	75.7	71.8	71.7	71.8	**100**							
	B	72.1	69.4	75.2	75.2	75.6	71.7	71.6	71.7	**98.2**	**100**						
HAstV-4		69.2	67.3	61.2	61.2	60.7	64.1	64	64.1	59.6	60	**100**					
HAstV-5		73.3	70.6	69.3	69.3	69.5	65	64.9	65	70.4	70.4	61.7	**100**				
HAstV-6	A	72	70	70.2	70.2	70.3	65.8	65.7	65.8	70.6	71.9	62.6	74	**100**			
	B	71.7	70.3	71.4	71.4	71.4	66.7	66.6	66.7	70.9	72.3	62.6	75.2	**95.5**	**100**		
	C	71.6	70.3	71.3	71.3	71.3	66.6	66.5	66.6	70.8	72.2	62.3	75.1	**95.4**	**99.6**	**100**	
	D	71.7	70.3	71.4	71.4	71.4	66.7	66.6	66.6	70.9	72.3	62.4	75.2	**95.5**	**99.7**	**99.9**	**100**

HAstV strains (named Yuc-8) to grow in CaCo-2 cells, and determined its complete genomic nucleotide sequence, the first for an HAstV-8 strain.

For isolation of the virus, a stool sample was diluted 1:2 in PBS, extracted with Freon, and filtered through a 0·22 μm membrane. This material was treated with trypsin (10 μg/ml) for 1 h at 37 °C, and inoculated into a CaCo-2 cell monolayer. After 60 min, the inoculum was removed, fresh Eagle's minimum essential medium was added, and the cells were incubated at 37 °C for 3 days. The virus was harvested by three freeze–thaw cycles, and was passaged again in the same cells, as described above. After seven passages, the presence of the virus was confirmed by immune electron microscopy with a hyperimmune serum to HAstV-1 (Herrmann *et al.*, 1990); by ELISA (IDEIA Astrovirus, Dako); by immunocytochemistry with monoclonal antibody 8G4, which recognizes HAstV-1 to -7 (Bass & Upadhyayula, 1997); and by RT–PCR with oligonucleotides Mon244 and Mon245 (Noel *et al.*, 1995). Total RNA from Yuc-8-infected cells was obtained with Trizol (Gibco-BRL), and used as template to amplify the astrovirus genome by RT–PCR. Reverse transcriptase SuperScript II (Gibco-BRL) and Vent DNA polymerase (New England Biolabs) were used for the RT–PCR reactions. Oligonucleotides synthesized were initially based on the previously reported sequence for HAstV-1 (accession no. Z25771), and subsequently based on the sequence obtained from Yuc-8. The amplified DNA fragments were sequenced with an ABI Prism DNA automatic sequencer, model 377-18 (Perkin Elmer). The sequence of the 5′ non-translated region (NTR) was determined from a PCR fragment obtained with an upstream 20-mer oligonucleotide corresponding to the 20 5′-end nucleotides which are conserved among astrovirus serotypes 1, 2 and 3. Nucleotide sequence of the 3′ NTR was determined from a DNA fragment obtained by RT–PCR with oligo(dT) and oligonucleotide Beg (Saito *et al.*, 1995) as primers. At least three PCR fragments of a given region, amplified independently, were used for determining the consensus sequence.

The full-length genomic RNA of HAstV Yuc-8 consists of 6759 bases, followed by a poly(A) tract. It has a 5′ NTR and a 3′ NTR of 83 and 85 nucleotides, respectively, and it is organized in three sequential open reading frames corresponding to ORFs 1a, 1b and 2 (Jiang *et al.*, 1993; Willcocks *et al.*, 1994*b*). The encoded polyproteins have the characteristic motifs described above.

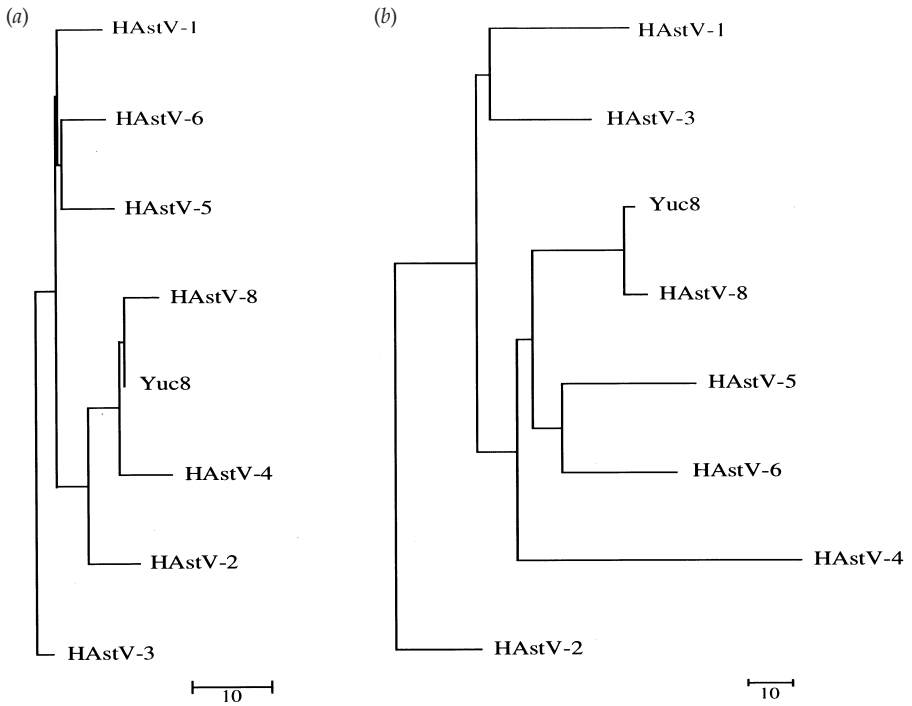

Fig. 2. Dendrograms showing the genetic relatedness of the ORF2 amino acid sequence from astrovirus serotypes 1 to 6 and 8. (*a*) Conserved domain of ORF2 comprising amino acids 1 to 415. (*b*) Hypervariable domain of ORF2 comprising amino acids 416 to 782 (Yuc-8 numbering). The scale bars in (*a*) and (*b*) represent 10 amino acid substitutions per each 100 residues. The dendrograms were generated with the GrowTree program of the Genetics Computer Group sequence analysis package (http://gcg.ceingebi.unam.mx/gcg-bin/seqweb.cgi), using the Kimura distance correction method and the neighbour-joining construction method. Accession numbers of the compared sequences are: L23513 (HAstV-1), A45695 (HAstV-2), AF141381 (HAstV-3), Z33883 (HAstV-4), U15136 (HAstV-5), Z46658 (HAstV-6) and Z66541 (HAstV-8). HAstV-7 was not included in this analysis since its complete sequence has not been reported.

Comparison of the non-structural polyprotein region of astrovirus Yuc-8 with the corresponding sequence of viruses belonging to serotypes 1, 2 and 3 showed a high level of conservation among these strains (higher than 93% identity at amino acid level). A small region of high variability was found around amino acids 767 and 790 in ORF1a (numbering according to the Yuc-8 sequence), where 7 and 15 residues were missing in Yuc-8 and HAstV-2, as compared to HAstV-3 and HAstV-1, respectively (Fig. 1*a*, regions in/del I and in/del II, respectively). The absence of the 15 amino acid residues at position 790 (region in/del II) has been associated with the adaptation of astroviruses to HEK and LLCMK2 cells, since viruses grown in these cells, but not in CaCo-2 cells, lack that region (Willcocks *et al.*, 1994*a*). Yuc-8 is also missing those 15 residues despite its having been adapted to grow in CaCo-2 cells, which suggests that the in/del region II, if it plays a role, is not the only factor involved in adaptation of human astroviruses to a specific cell line. A more detailed analysis of field and culture-adapted astrovirus strains is needed to resolve this issue.

ORF2 of Yuc-8 encodes a polyprotein of 782 amino acids, with an estimated molecular mass of 86·5 kDa. Comparison of the predicted full-length ORF2 amino acid sequence from viruses of different serotypes and geographical regions indicated a good correlation between genotype and serotype. Within one given serotype where more than one ORF2 complete sequence has been reported, the strains were more than 94 % identical (Table 1). Yuc-8 was thus confirmed as a genotype (serotype) 8 virus, since it was more closely related to an isolate of HAstV-8 from the United Kingdom (94% identical) than to HAstV serotypes 1 to 6 (between 69 and 73% identity) (Table 1, Fig. 2). Also, comparison of the Yuc-8 sequence with the partial sequence (amino acid residues 82 to 200) reported for HAstV-7 revealed an identity of 90% between these two strains, the lowest value for any pairwise comparison when that specific region was considered (not shown).

Knowledge about processing of the ORF2 structural polyprotein is limited. It has been reported that this polyprotein is first cleaved intracellularly at the conserved lysine residue at position 71, to yield a protein of 79 kDa; in the absence of trypsin in the culture medium this seems to be the end product, which can be incorporated into the mature virions (Bass & Qiu, 2000). On the other hand, products of 34, 29 and 26 kDa (VP34, VP29 and VP26) are consistently found in trypsin-treated purified viruses (Bass & Qiu, 2000; Sanchez-Fauquier *et al.*, 1994). VP29 and VP26 were shown by N-terminal sequence analysis to be generated by cleavage of the

polyprotein of HAstV-2 at arginine residues 361 and 395, respectively (numbers based on the Yuc-8 sequence); thus, these two proteins probably share their carboxy terminus (Sanchez-Fauquier *et al*., 1994). Arg-361 (but not Arg-395) is totally conserved among all astroviruses sequenced to date, as are other trypsin-susceptible cleavage sites present at amino acid residues 347, 354, 359, 365 and 380, which could potentially serve as processing sites (Fig. 1*b*). This trypsin cleavage region is surrounded by highly conserved regions, which could play a structural role to make the cleavage sites accessible for processing. In rotaviruses, the trypsin cleavage region on the VP4 protein, which is responsible for the enhancement of virus infectivity, is also surrounded by highly conserved regions (Lopez *et al*., 1986).

Alignment of the ORF2 sequences of seven HAstV serotypes (Fig. 1*b*) showed the existence of two distinct domains, which were first observed by Willcocks *et al*. (1995) who analysed three serotypes only. One conserved amino-terminal domain spans amino acid residues 1 to 415, with identities higher than 81% among the different serotypes. Two short variable regions between residues 292 to 319 (VR1) and 386 to 399 (VR2), not reported previously, were observed in this domain (Fig. 1*b*). The highly variable second domain (identities of 36 to 60% among serotypes) starts at amino acid 416 and extends to the end of the protein, but in contrast to the report of Willcocks *et al*. (1995), the conservation at the end of the ORF2 was not observed among the serotypes (Fig. 1*b*).

Phylogenetic analysis of ORF2 showed a completely different genetic relatedness among serotypes 1 to 8 when either of the two domains described above was analysed. The phylogenetic tree generated by comparison of the conserved domain (Fig. 2*a*) was very similar to the trees reported previously by Belliot *et al*. (1997) and Noel *et al*. (1995), who analysed a region of 137 amino acids (residues 73 to 210 of Yuc-8). Serotypes 2, 4 and 8 were found to be closely grouped, while serotypes 1, 3, 5 and 6 were more distantly related to this group (Fig. 2*a*). On the other hand, the dendrogram generated by comparison of the variable domain of ORF2 (residues 416 to 782) was quite different. Serotypes 4 and 8 were among the least related (Fig. 2*b*), which resulted in a phylogenetic tree similar to that reported in another analysis including part of the 3′-end of the astrovirus genome (Monceyron *et al*., 1997). It is likely that the variable carboxy-terminal domain of the ORF2 polyprotein is subjected to immunological pressure, which probably contributes to the variability observed. In fact, neutralizing antibodies which recognize VP29 and VP26 have been identified (Bass & Upadhyayula, 1997). Furthermore, it is of interest that the genetic relatedness observed among the various HAstV serotypes, when the variable region was analysed, differs so much from the relationships found in the conserved amino-terminal half of the protein. The selective pressure that created these marked deviations from the mutation rate of astroviruses could operate differentially along the ORF2 polyprotein, most likely as the result of a sum of factors including, among others, the intrinsic structural constraints of the virus particle and the host immune response of the particular populations infected. On the other hand, the non-superimposable dendrograms shown in Fig. 2 could also be the result of intraserotypic astrovirus recombination, suggested as previously (Belliot *et al*., 1997; Jonassen *et al*., 1998). Whatever the reason for the observed variations, it is evident that the genomic region chosen to compare different HAstV strains should take into account the differential variability observed in the astrovirus genome, particularly that of ORF2.

We acknowledge the technical assistance of Paul Gaytan, Eugenio López and René Hernández for the synthesis of oligonucleotides and DNA sequencing. We thank Dorsey Bass (Stanford University) for providing monoclonal antibody 8G4, used in this work. This work was partially supported by grants 75197-527106 from the Howard Hughes Medical Institute, MENSE31739. from the National Council for Science and Technology–Mexico, and IN200999 from DGAPA–UNAM.

References

Bass, D. & Qiu, S. (2000). Proteolytic processing of the astrovirus capsid. *Journal of Virology* **74**, 1810–1814.

Bass, D. & Upadhyayula, U. (1997). Characterization of human serotype 1 astrovirus-neutralizing epitopes. *Journal of Virology* **71**, 8666–8671.

Belliot, G., Laveran, H. & Monroe, S. (1997). Detection and genetic differentiation of human astroviruses: phylogenetic grouping varies by coding region. *Archives of Virology* **142**, 1323–1334.

Gaggero, A., O'Ryan, M., Noel, J., Glass, R., Monroe, S., Mamani, N., Prado, V. & Avendano, L. (1998). Prevalence of astrovirus infection among Chilean children with acute gastroenteritis. *Journal of Clinical Microbiology* **36**, 3691–3693.

Herrmann, J., Hudson, R., Perron-Henry, D., Kurtz, J. & Blacklow, N. (1988). Antigenic characterization of cell-cultivated astrovirus serotypes and development of astrovirus-specific monoclonal antibodies. *Journal of Infectious Diseases* **158**, 182–185.

Herrmann, J., Cubitt, W., Hudson, R., Perron-Henry, D., Oshiro, L. & Blacklow, N. (1990). Immunological characterization of the Marin County strain of astrovirus. *Archives of Virology* **110**, 213–220.

Herrmann, J., Taylor, D., Echeverria, P. & Blacklow, N. (1991). Astroviruses as a cause of gastroenteritis in children. *New England Journal of Medicine* **324**, 1757–1760.

Jiang, B., Monroe, S., Koonin, E., Stine, S. & Glass, R. (1993). RNA sequence of astrovirus: distinctive genomic organization and a putative retrovirus-like ribosomal frameshifting signal that directs the viral replicase synthesis. *Proceedings of the National Academy of Sciences, USA* **90**, 10539–10543.

Jonassen, C., Jonassen, T. & Grinde, B. (1998). A common RNA motif in the 3′ end of the genomes of astroviruses, avian infectious bronchitis virus and an equine rhinovirus. *Journal of General Virology* **79**, 715–718.

Lewis, T. & Matsui, S. (1997). Studies of the astrovirus signal that induces (−1) ribosomal frameshifting. *Advances in Experimental Medicine Biology* **412**, 323–330.

Lopez, S., Arias, C., Mendez, E. & Espejo, R. (1986). Conservation in rotaviruses of the protein region containing the two sites associated with trypsin enhancement of infectivity. *Virology*. **154**, 224–7.

Marczinke, B., Bloys, A., Brown, T., Willcocks, M., Carter, M. & Brierley, I. (1994). The human astrovirus RNA-dependent RNA polymerase

coding region is expressed by ribosomal frameshifting. *Journal of Virology* **68**, 5588–5595.

Monceyron, C., Grinde, B. & Jonassen, T. (1997). Molecular characterization of the 3′-end of the astrovirus genome. *Archives of Virology* **142**, 699–706.

Monroe, S., Stine, S., Gorelkin, L., Herrmann, J., Blacklow, N. & Glass, R. (1991). Temporal synthesis of proteins and RNAs during human astrovirus infection of cultured cells. *Journal of Virology* **65**, 641–648.

Monroe, S., Jiang, B., Stine, S., Koopmans, M. & Glass, R. (1993). Subgenomic RNA sequence of human astrovirus supports classification of Astroviridae as a new family of RNA viruses. *Journal of Virology* **67**, 3611–3614.

Mustafa, H., Palombo, E. & Bishop, R. (2000). Epidemiology of astrovirus Infection in young children hospitalized with acute gastroenteritis in Melbourne, Australia, over a period of four consecutive years, 1995 to 1998. *Journal of Clinical Microbiology* **38**, 1058–1062.

Noel, J. & Cubitt, D. (1994). Identification of astrovirus serotypes from children treated at the Hospitals for Sick Children, London 1981–93. *Epidemiology and Infection* **113**, 153–159.

Noel, J., Lee, T., Kurtz, J., Glass, R. & Monroe, S. (1995). Typing of human astroviruses from clinical isolates by enzyme immunoassay and nucleotide sequencing. *Journal of Clinical Microbiology* **33**, 797–801.

Palombo, E. & Bishop, R. (1996). Annual incidence, serotype distribution, and genetic diversity of human astrovirus isolates from hospitalized children in Melbourne, Australia. *Journal of Clinical Microbiology* **34**, 1750–1753.

Saito, K., Ushijima, H., Nishio, O., Oseto, M., Motohiro, H., Ueda, Y., Takagi, M., Nakaya, S., Ando, T., Glass, R. and others (1995). Detection of astroviruses from stool samples in Japan using reverse transcription and polymerase chain reaction amplification. *Microbiology and Immunology* **39**, 825–828.

Sanchez-Fauquier, A., Carrascosa, A., Carrascosa, J., Otero, A., Glass, R., Lopez, J., Martin, C. S. & Melero, J. (1994). Characterization of a human astrovirus serotype 2 structural protein (VP26) that contains an epitope involved in virus neutralization. *Virology* **201**, 312–320.

Shastri, S., Doane, A., Gonzales, J., Upadhyayula, U. & Bass, D. (1998). Prevalence of astroviruses in a children's hospital. *Journal of Clinical Microbiology* **36**, 2571–2574.

Willcocks, M., Ashton, N., Kurtz, J., Cubitt, W. & Carter, M. (1994*a*). Cell culture adaptation of astrovirus involves a deletion. *Journal of Virology* **68**, 6057–6058.

Willcocks, M., Brown, T., Madeley, C. & Carter, M. (1994*b*). The complete sequence of a human astrovirus. *Journal of General Virology* **75**, 1785–1788.

Willcocks, M., Kurtz, J., Lee, T. W. & Carter, M. (1995). Prevalence of human astrovirus serotype 4: capsid protein sequence and comparison with other strains. *Epidemiology and infection* **114**, 385–394.

Willcocks, M., Boxall, A. & Carter, M. (1999). Processing and intracellular location of human astrovirus non-structural proteins. *Journal of General Virology* **80**, 2607–2611.

Received 9 May 2000; Accepted 8 August 2000