JIDC | THE JOURNAL OF INFECTION IN DEVELOPING COUNTRIES

Original Article

Molecular characterization of three enteroviral strains isolated in Kuwait from young children with serious conditions

Wassim Chehadeh, Sanaa Abdulkader Moalim Ali, Syeda Mubeen Maimoona

Department of Microbiology, Faculty of Medicine, Kuwait University, Kuwait

Abstract

Introduction: Human enteroviruses are single stranded RNA viruses associated with many serious diseases such as encephalitis and myocarditis. They consist of up to 100 immunologically and genetically distinct types. Three enteroviral isolates, 2104, 3936 and 3988, were previously isolated from patients with neurological disorders or sepsis-like illness. In this study, the molecular characterization of the isolates was investigated.

Methodology: A full genome sequencing was performed by Sanger method, followed by phylogenetic and bootscanning analyses. A detailed analysis of genetic differences between the clinical and prototype isolates were investigated by mapping polymorphisms at nucleotide and amino acid levels, and by comparing RNA secondary structure in the noncoding regions.

Results: Based on the phylogenetic analysis of the VP1 gene and complete genome, 2104 was typed as coxsackievirus B1, 3936 as coxsackievirus B5, and 3988 as echovirus 7. Similarity and bootscan plots provided support for intra- and intertypic recombination with crossover points occurring mainly along the nonstructural coding regions of the isolates. A sequence divergence of 12 to 14% was detected in the 5'-noncoding region between the clinical isolates and their corresponding prototype strains. Synonymous and nonsynonymous substitutions could be also mapped to different coding regions of the isolates, including those coding for the Puff, Knob and the hydrophobic pocket of the capsid. Examination of relative frequencies of synonymous and nonsynonymous substitutions in different coding regions of enteroviral isolates showed no evidence for selective pressure.

Conclusion: The results provided a better understanding of the genetic variations, evolution and adaptation of enteroviruses in Kuwait.

Key words: enterovirus; isolates; phylogeny; recombination; Kuwait

J Infect Dev Ctries 2017; 11(8):626-639. doi:10.3855/jidc.9065

(Received 28 June 2016 – Accepted 17 November 2016)

Introduction

Enteroviruses (EVs) can cause many different diseases affecting a variety of target organs causing neurological, respiratory, and cardiovascular diseases [1]. They belong to the *Picornaviridae* family. The non-segmented positive-strand RNA genome is about 7,500 nucleotides long, encoding the four structural proteins, VP4, VP2, VP3 and VP1 (forming the P1 region), and the seven non-structural 2A to 2C (P2 region) and 3A to 3D (P3 region) viral proteins. The open reading frame (ORF) is flanked by the 5' and 3' untranslated regions (5'UTR and 3'UTR). The 5'UTR contains the internal ribosome entry site (IRES) responsible for cap-independent initiation of translation, as well as other secondary structural elements responsible for genome replication. The 3'UTR plays a role in initiating the synthesis of negative-strand RNA [2].

Based on the degree of their genetic relatedness, human EVs are classified into four species, EV-A to D. Human EV-A includes 19 types: coxsackievirus (CV) A2-8, CVA10, CVA12, CVA14, CVA16, EV-A71, EV-A76, EV-A89-91, EV-A114, EV-A119, and EV-A120. Human EV-B includes 58 types: CVA9, CVB1-6, EV-B69, EV-B73-75, EV-B77 -88, EV-B93, EV-B97, EV-B98, EV-B100, EV-B101, EV-B106, EV-B107, echovirus (E) 1-7, E9, E11-21, E24-27, and E29-33. Human EV-C includes 23 types: poliovirus (PV) 1-3, CVA1, CVA11, CVA13, CVA17, CVA19-22, and CVA24. Human EV-D includes 4 types EV-D68, EV-D70, EV-D94, and EV-D111. The three species of rhinoviruses (human rhinovirus A, B, and C) have been officially included in the Enterovirus genus [3].

In Kuwait, EVs are highly prevalent in patients with aseptic meningitis or encephalitis, and in patients suffering from febrile illness. Genotyping of detected viral sequences in clinical samples showed the presence of echoviruses 4, 7, 9, 11 and 30, coxsackieviruses B2-6, and other serotypes [4-6]. Since only a limited number of complete sequences of clinical EV isolates have been reported so far, compared to complete

nucleotide sequences available for all of the prototype EV strains, a full-genome analysis of clinical EVs might contribute to foster better understanding of the genetic variations, evolution and adaptation of the circulating EV strains, and their role in EV outbreaks and virulence.

In this study, the complete nucleotide sequences of three clinical enteroviral isolates were compared to the corresponding prototype strains. A detailed analysis of genetic differences between the clinical and prototype isolates were investigated by mapping polymorphisms at nucleotide and amino acid levels, and by comparing RNA secondary structure in the noncoding regions. To shed light on the evolution of the three clinical EV isolates, the phylogenetic relationship of the clinical isolates to other human enteroviruses were determined, a recombination analysis was performed, and the type of evolutionary selection pressure acting at the molecular level was identified.

Methodology

Viruses

Three EV strains isolated on Vero cell line from three different patients were used in this study. The 2104 strain was isolated in 2008 from the cerebrospinal fluid of a 1-year-old female with respiratory distress and aseptic meningitis. The 3936 strain was isolated in 2009 from the stool of a 4-year-old male with aseptic meningitis. Finally, the 3988 strain was isolated in 2009 from the blood of a neonate with sepsis-like illness.

Nucleotide sequencing of enteroviruses

Full genome sequencing of the enteroviral strains was carried out as described previously [7]. Briefly, amplification of enteroviral genome was done by RT-PCR using "primer-walking" strategy. The Qiagen One-Step RT-PCR Kit (Qiagen GmbH, Hilden, Germany) was used in all amplification reactions. The amplified PCR products were analyzed on 2% agarose gel electrophoresis. The PCR product was then purified using the Wizard SV GEL and PCR Clean-Up System kit (Promega Corporation, Madison, USA), and the nucleotide sequences of both DNA strands were determined by direct double-strand DNA cycle sequencing in two directions using the ABI PRISM BigDye Terminator Cycle Sequencing v3.1 kit (Applied Biosystems, Foster City, USA). Post sequencing PCR purification was performed to remove unbound fluorescent dye deoxy terminators using BigDye XTerminator Purification kit (Applied Biosystems, Foster City, USA). The samples were then denatured for 2 minutes at 95˚C, and immediately chilled on ice, and loaded on an ABI 3130xl Genetic Analyzer (Applied Biosystems, Foster City, USA). DNA sequences were subjected to electrophoresis on a 50-cm capillary array using POP6 polymer (Applied Biosystems, Foster City, USA) as a separation medium, and analyzed using the Sequencing Analysis Software v 5.3.1 (Applied Biosystems, Foster City, USA).

The derived sequences of all segments were assembled by using The Molecular Evolutionary Genetics Analysis (MEGA) software version 4.02 [8]. A homology search was conducted using the BLAST server at the National Center for Biotechnology Information (National Library of Medicine, Bethesda, USA). The resulting DNA sequences were aligned with complete prototype HEV genome sequences available in the GenBank database. Alignment analysis was done using the ClustalW method in the MEGA software version 4.02 [8]. Highly variable regions were identified in the coding and non-coding sequences. The complete genome sequence of EV sequences from this study has been deposited in GenBank under accession number KP260537, KP233830 and KP202389.

Prediction of RNA secondary structures

Putative RNA secondary structures of the 5'UTR and 3'UTR of the clinical enteroviral isolates were constructed using RNAstructure software version 5.7 [9], and the lowest folding free energy levels for the optimal structures were compared to those of corresponding prototypes.

Evolutionary divergence analysis

To determine whether the mutations supposed to be detected in the enteroviral isolates are the consequence of selection pressure, the average number of synonymous substitutions per synonymous site (dS) and the average number of nonsynonymous substitutions per nonsynonymous site (dN), for each coding region were calculated for all available EV genomes using the Nei-Gojobori (p-distance) method in MEGA software [10]. Identical genes and genes with evidence of recombination were excluded from the analysis. The dN/dS ratio was used as an index to assess positive selection; dN/dS >1 indicates positive (diversifying) selection, dN/dS = 1 indicates neutral selection, whereas dN/dS <1 means negative (purifying) selection [11].

Phylogenetic and recombination analysis

Phylogenetic analysis was performed for the full enteroviral genome, and for different genetic regions separately.

Table 1. Enteroviruses and accession numbers of genome sequences used in this study.

Enterovirus Type	Strain	Country	Isolation Year	Accession number
CV-A9	Griggs	USA	Early 1950s	D00627
CV-B1	Conn-5	USA	1948	M16560
CV-B1	Chi07	USA	2007	KJ849619
CV-B1	2104/08	Kuwait	2008	KP260537
CV-B1	MSH/KM9/2009	China	2009	JN596588
CV-B1	1167438_pmMC	Switzerland	2010	JN797615
CV-B1	CVB1SD2011CHN	China	2011	JX976769
CV-B1	CVB1Nm	USA	Unknown	EU147493
CV-B2	Ohio-1	USA	1947	AF085363
CV-B3	Nancy	USA	1949	M16560
CV-B4	JVB	USA	1951	X05690
CV-B4	E2	USA	1979	AF311939
CV-B5	Faulkner	USA	1952	AF114383
CV-B5	1954/85/UK	UK	1985	X67706
CV-B5	2000/CSF/KOR	South Korea	2000	AY875692
CV-B5	SH1	China	2008	GU376747
CV-B5	A210/KM/09	China	2009	JX843811
CV-B5	CVB5/SD/09	China	2009	JX276378
CV-B5	2009-148-4	China	2009	KP266578
CV-B5	2009-154-2	China	2009	KP266577
CV-B5	3936/09	Kuwait	2009	KP233830
CV-B5	COXB5/Henan/2010	China	2010	HQ998851
CV-B5	CVB5/CC10/10	China	2010	JN580070
CV-B5	2001-01-101	China	2010	KP266576
CV-B5	2001-01-11	China	2010	KP266575
CV-B5	19CSF	China	2011	JX017381
CV-B5	20CSF	China	2011	JX017380
CV-B5	03001N	China	2011	JX017383
CV-B5	17Y	China	2011	JX017382
CV-B5	CV-B5/P727/2013	China	2013	KP289438
CV-B6	Schmitt	Philippines	1953	AF105342
E1	Farouk	Egypt	1951	AF029859
E2	Cornelis	USA	1951	AY302545
E3	Morrisey	USA	1951	AY302553
E4	Pesacek	USA	1951	AY302557
E5	Noyce	USA	1954	AF083069
E6	D'Amori	USA	1955	AY302558
E7	Wallace	USA	1953	AY302559
E7	UMMC	Malaysia	2000	AY036578
E7	2001-31	China	2001	KP266570
E7	3988/09	Kuwait	2009	KP202389
E7	DH22G	China	2012	KJ765699
E9	DM	Finland	2000	AF524867
E9	Barty	USA	1956	AF524866
E9	Hill	USA	1953	X92886
E11	Gregory	USA	1953	X80059
E11	HUN-1108	Hungary	1989	AJ577589
E11	FIN-0666	Finland	1989	AJ577590
E11	ROU-9191	Romania	1991	AJ577594
E12	Travis	Philippines	1953	X79047
E13	Del Carmen	Philippines	1953	AY302539
E14	Tow	USA	1954	AY302540
E15	CH 96-51	USA	1951	AY302541
E16	Harrington	USA	1951	AY302542

Table 1. (continued) Enteroviruses and accession numbers of genome sequences used in this study.

Enterovirus Type	Strain	Country	Isolation Year	Accession number
E17	CHHE-29	Mexico	unknown	AY302543
E18	Metcalf	USA	1955	AF317694
E19	Burke	USA	1955	AY302544
E20	JV-1	USA	1956	AY302546
E21	Farina	USA	1950	AY302547
E24	DeCamp	USA	1956	AY302548
E25	JV-4	USA	1957	AY302549
E26	Coronel	Philippines	1953	AY302550
E27	Bacon	Philippines	1953	AY302551
E29	JV-10	USA	1958	AY302552
E30	Bastianni	USA	1958	AF162711
E31	Caldwell	USA	1955	AY302554
E32	PR-10	Puerto Rico	1961	AY302555
E33	Toluca-3	Mexico	1959	AY302556
EV-B69	Toluca-1	Mexico	1959	AY302560
EV-B73	088	China	2004	KF874626
EV-B73	CA55-1988	USA	1955	AF241359
EV-B74	CA75-10213	USA	1975	AY556057
EV-B74	Rikaze-136	China	2010	JQ397329
EV-B75	OK85-10362	USA	1985	AY556070
EV-B75	102	China	1997	KF874627
EV-B77	CF496-99	France	1999	AJ493062
EV-B77	TX97-10394	USA	1997	AY843302
EV-B79	CA79-10384	USA	1979	AY843297
EV-B79	NH95-0601	Japan	1995	AB426610
EV-B80	CA67-10387	USA	1967	AY843298
EV-B80	HZ01/SD/CHN/2004	China	2004	JX644073
EV-B81	CA68-10389	USA	1968	AY843299
EV-B81	99298c	China	1999	KJ755190
EV-B81	99279	China	1999	KJ755189
EV-B82	CA74-10390	USA	1974	AY843300
EV-B83	CA76-10392	USA	1976	AY843301
EV-B84	CIV03-10603	Côte d'Ivoire	2003	DQ902712
EV-B85	BAN00-10353	Bangladesh	2000	AY843303
EV-B85	HT-LYKH202F	China	2011	JX898908
EV-B85	HTPS-MKLH04F	China	2011	JX898907
EV-B85	HTPS-MJH21F	China	2011	JX898906
EV-B85	HTYT-ARLH403F	China	2011	JX898905
EV-B85	HTYT-ARL-AFP02F	China	2011	JX898909
EV-B86	BAN00-10354	Bangladesh	2000	AY843304
EV-B87	BAN01-10396	Bangladesh	2001	AY843305
EV-B87	LY02/SD/CHN/2000	China	2000	KC292019
EV-B88	BAN01-10398	Bangladesh	2001	AY843306
EV-B93	38-03	Congo	2000	KM273013
EV-B97	BAN99-10355	Bangladesh	1999	AY843307
EV-B97	99188	China	1999	GU550508
EV-B97	DT94-0227	Japan	1994	AB426611
EV-B98	T92-1499	Japan	1992	AB426608
EV-B100	BAN2000-10500	Bangladesh	2000	DQ902713
EV-B101	CIV02-10361	Côte d'Ivoire	2002	AY843308
EV-B106	148/YN/CHN/12	China	2012	KF990476
EV-B107	TN94-0349	Japan	1994	AB426609

CV, coxsackievirus; E, echovirus; EV, enterovirus

The Phylogenetic trees were reconstructed using neighbor-joining method [12], with evolutionary distances computed using the Kimura 2-parameter method [13]. A bootstrap test with 1,000 replicates was used to estimate the confidence of branching patterns in the trees [14]. The resulting alignments were analyzed for evidence of recombination using the Simplot software package version 3.5.1 (Johns Hopkins University, Baltimore, USA). Similarity plots were built with a window size of 200 nt, incremental steps of 20 bases, and a transition-transversion ratio of 2. To assess potential recombinational relationships, aligned sequences were subsequently analyzed by using the bootscanning method implemented in SimPlot. Phylogenetic trees were generated for each 200-nucleotide window in 20-nucleotide increments, by using the neighbor-joining method, with Kimura two-parameter model, and a transition-transversion ratio of 2. The bootstrap values for all possible sequence comparisons were plotted as a function of genome position.

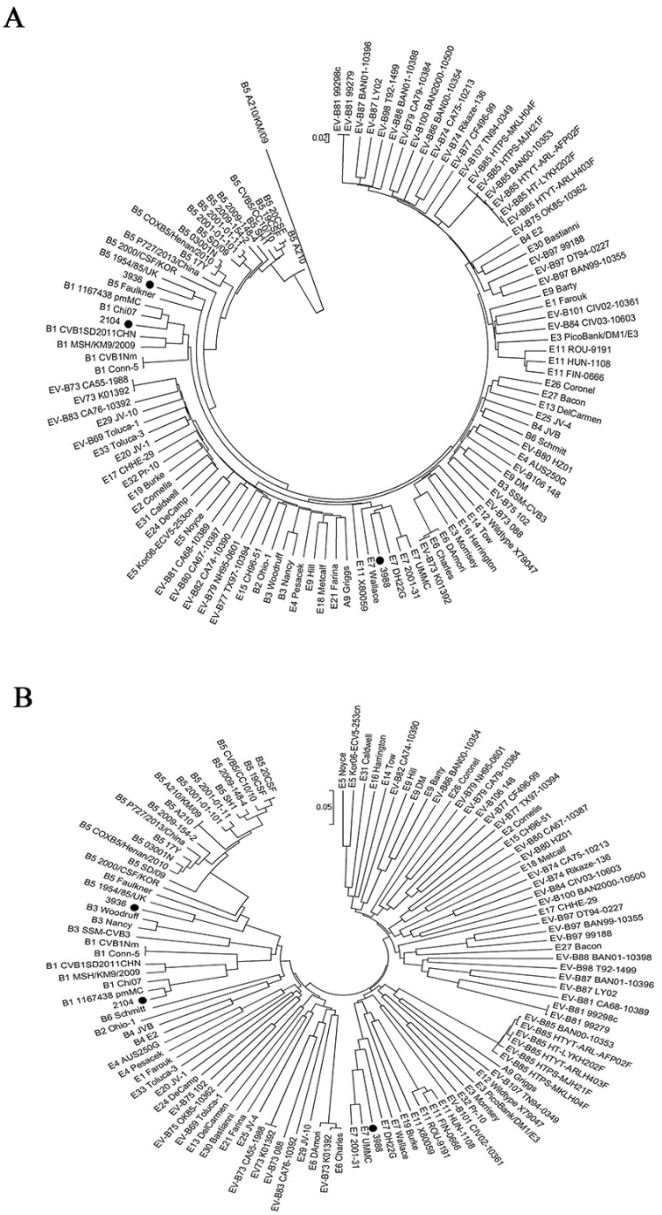
Results

Phylogenetic analysis

Based on the similarity of the VP1 nucleotide sequence of the three EV isolates with that of other strains available in GenBank database, 2104 was identified as CV-B1 (best BLAST hit, 99%), 3936 as CV-B5 (best BLAST hit, 95%), and 3988 as echovirus 7 (best BLAST hit, 88%). The phylogenetic analysis was then conducted for the complete nucleotide sequences of HEV-B species including all available full genome sequences for CV-B1, CV-B5 and echovirus 7 (Table 1). The phylogenetic dendrogram obtained with the full EV genome (Figure 1A) and the structural coding P1 region (Figure 1B) confirmed the typing results. The 2104 isolate forms one cluster with the CV-B1 Chi07 and 1167438_pmMC isolates (bootstrap value, 100%, whole genome) or with the CV-B1 1167438_pmMC isolate (bootstrap value, 91%, P1 region). The 3936 isolate is grouped with the CV-B5 1954/85/UK isolate (bootstrap value, 100%, whole genome and P1 region), whereas the 3988 isolate forms one cluster with the echovirus 7 DH22G isolate (bootstrap value, 87%, whole genome) or the echovirus UMMC and 2001-31 isolates (bootstrap value, 100%, P1 region).

The results of the phylogenetic analysis of the non-coding and nonstructural coding regions (P2 and P3) were not consistent (Table 2). The three EV isolates could not be typed using P3 region, whereas using P2 region, only the 2104 isolate could be typed as CV-B1 (bootstrap value >70%). Furthermore, the phylogenetic analysis conducted using 5'UTR was not reliable for the 3988 isolate (Table 2). This conflicting tree topologies

Figure 1. Phylogenetic analysis based on the full-genome (A) and P1 region (B) of human Enteroviruses B. The evolutionary history was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The black circles next to the taxa represent enteroviral sequences from this study. Each enterovirus reference sequence is labeled with its corresponding type.



in the 5' UTR, P2 and P3 genomic regions may be due to a previous recombination event resulting in chimeric genome sequences.

Recombination analysis

A similarity plot analysis of the 2104 strain with genomic sequences of other HEV-B strains listed in Table 1 was performed. The 5'UTR and P1 regions showed >90% nucleotide sequence similarity with those of Chi07 and 11167438_pmMC strains (Figure 2A). In the nonstructural coding region, the 2104 isolate displayed 83-87% nucleotide sequence similarity with the echovirus 1 Farouk strain (nucleotides 4200-5400), 81-83% similarity with the CV-B1 Chi07 strain (nucleotides 5500-5700), 80-93% similarity with the CV-B1 11167438_pmMC strain (nucleotides 6400-6800), and 80-85% similarity with the CV-B1 Conn-5 strain (nucleotides 7100-7200). Bootscan analysis showed multiple recombination crossover points mainly with the CV-B1 11167438_pmMC and echovirus 1 Farouk strains along the structural and nonstructural coding regions of 2104 genome (Figure 2B).

The 3936 isolate was most closely related to the CV-B5 1954/85/UK strain in the 5'UTR and P1 region, whereas the P2 and P3 regions displayed 75-95% nucleotide sequence similarity with EV-B86 and EV-B88 (Figure 2C). Bootscan analysis revealed recombination between the CV-B5 1954/85/UK strain, the EV-B86 BAN00-10354 strain and the EV-B88 BAN01-10398 strain (Figure 2D).

The 5'UTR of 3988 isolate had 80-95% nucleotide sequence similarity with EV-B86 (nucleotides 400-600), whereas P1 region had 80-91% nucleotide sequence similarity with the echovirus 7 UMMC strain. However, P2 region was most closely related to the EV-B86 BAN00-10354 strain, whereas P3 region had 80-90% nucleotide sequence similarity with the echovirus 7 DH22G strain (5400-7300), and 78-88% similarity with the EV-B86 BAN00-10354 strain (nucleotides 5057-7300) (Figure 2E). Evidence of recombination between the echovirus 7 UMMC strain, the EV-B86 BAN00-10354 strain, and the E7 DH22G strain was found by bootscan analysis (Figure 2F).

Sequence analysis

Table 3 shows the nucleotide and amino acid similarities of the 2104 isolate with the prototype and other CV-B1 strains. Within the complete genome, the similarities between the clinical isolate and the prototype strain were 80 % for nucleotides and 95% for amino acids. The highest nucleotide sequence similarity was found with the 1167438-pmMC strain (88.2%) whereas the highest amino acid sequence similarity was found with the Chi07 strain (96.25%) and the 1167438-pmMC strain (96.20%). The 5'UTR of the 2104 strain is 5 nucleotides longer than that of the Conn-5 strain. However, 5'UTRs of the 2104 and Conn-5 strains have comparable minimum free energy secondary structures (ΔG= -253.7 Kcal/mol and -251.7 Kcal/mol, respectively). The two virus genomes share identical nucleotide lengths within the ORF and 3'UTR.

Table 2. Typing results inferred from the phylogenetic analysis of enteroviral isolates

Genetic region	Isolate 2104		Isolate 3936		Isolate 3988	
	Type	Bootstrap	Type	Bootstrap	Type	Bootstrap
Full genome	CV-B1	100%	CV-B5	100%	E7	73%
5'UTR	CV-B1	99%	CV-B5	99%	EV-B80	<50%
P1	CV-B1	100%	CV-B5	100%	E7	100%
VP4	CV-B1	99%	CV-B5	79%	E7	<50%
VP2	CV-B1	99%	CV-B5	99%	E7	93%
VP3	CV-B1	99%	CV-B5	99%	E7	99%
VP1	CV-B1	100%	CV-B5	100%	E7	99%
P2	CV-B1	97%	CV-B5	<50%	EV-B86	70%
2A	CV-B1	99%	CV-B5	99%	EV-B98	<50%
2B	CV-B1	72%	EV-B87	<50%	EV-B86	<50%
2C	E11	69%	EV-B87	<50%	Unknown	-
P3	Unknown	-	Unknown	-	EV-B85	<50%
3A	EV-B84	<50%	EV-B88	<50%	EV-B86	57%
3B	Unknown	-	EV-B88	<50%	CV-B4	<50%
3C	Unknown	-	EV-B88	<50%	EV-B100	<50%
3D	Unknown	-	EV-B86	<50%	EV-B85	<50%
3'UTR	Unknown	-	EV-B85	<50%	Unknown	<50%

CV: coxsackievirus; E: echovirus; EV: enterovirus

Figure 2. Similarity plot (left panel) and bootscan analyses (right panel) of 2104 (A and B), 3936 (C and D) and 3988 (E and F) genomes relative to reference enteroviral sequences selected after analyzing the results of comparison with all reference sequences of species B enteroviruses. Plots were generated using SimPlot program and a sliding window of 200 bp with 20 bp step size. Positions containing gaps were excluded from the comparison.

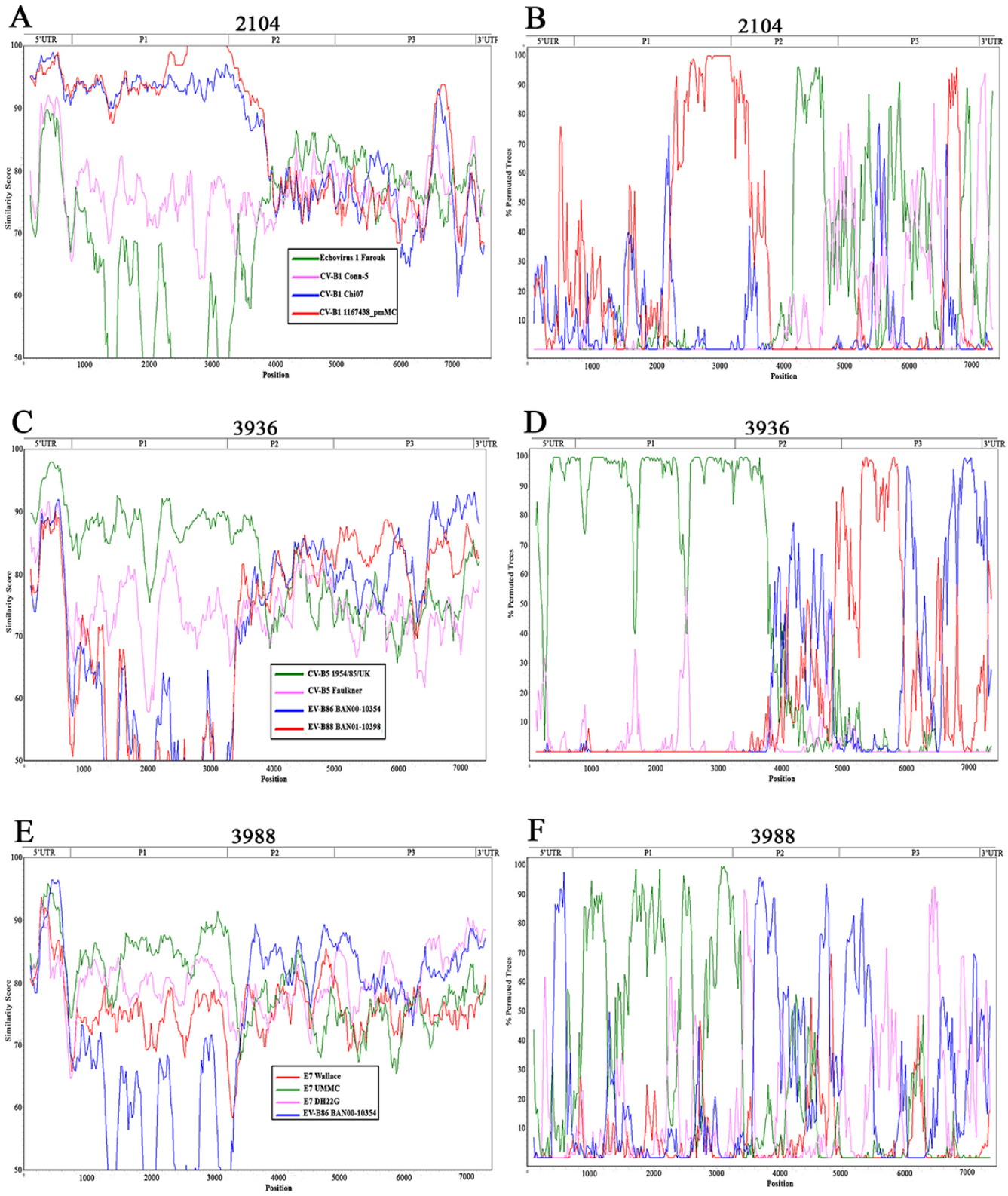


Table 3. Nucleotide and deduced amino acid sequence identity of 2104 isolate with CVB1 prototype strain Conn-5 and other CV-B1 strains

	Sequence identity with CV-B1 Conn-5 strain		Sequence identity with other CV-B1 strains	
Genetic region	**Nucleotide**	**Amino acid**	**Nucleotide**	**Amino acid**
Genome	80%	95%	80.5-88.2%	94.6-96.2%
5' UTR	88%	-	84.9-96.1%	-
P1				
VP4	82%	100%	81.2-94.7%	95.6-100%
VP2	80%	95%	79.1-93.2%	95.4-98.9%
VP3	79%	94.5%	79.5-94.8%	94.9-98.7%
VP1	78.5%	96%	78.6-99.6%	95.3-99.6%
P2				
2A	78%	91%	74.0-93.6%	87.3-96.0%
2B	76%	93%	76.1-83.5%	88.9-93.9%
2C	82%	98%	79.8-81.9%	96.3-98.2%
P3				
3A	80.5%	94%	6.4-81.3%	94.4-95.5%
3B	85%	100%	72.7-84.8%	90.9-100%
3C	79%	96%	77.6-82.5%	95.6-97.3%
3D	81%	92%	79.3-85.3%	91.3-98.3%
3' UTR	90%	-	80.4-84.3%	-

Table 4. Evolutionary divergence of coding regions in CV-B1 isolates*

Genetic region	Length (nt)	Average dN	Average dS	dN/dS	Evolutionary selection
P1					
VP4	207	0.014	0.575	0.024	Negative
VP2	789	0.019	0.638	0.029	Negative
VP3	714	0.015	0.610	0.024	Negative
VP1	834	0.020	0.589	0.034	Negative
P2					
2A	450	0.044	0.652	0.067	Negative
2B	297	0.036	0.694	0.052	Negative
2C	987	0.018	0.763	0.023	Negative
P3					
3A	267	0.026	0.756	0.034	Negative
3B	66	0.041	0.82	0.050	Negative
3C	549	0.029	0.755	0.038	Negative
3D	1386	0.031	0.711	0.044	Negative

* Complete coding sequence (cds) information was available for the following CV-B1 isolates: Conn-5 (M16560), Chi07 (KJ849619), MSH/KM9/2009 (JN596588), CVB1SD2011CHN (JX976769), CVB1Nm (EU147493), 1167438-pmMC (JN797615) and 2104/08 (KP260537)

Table 5. Nucleotide and deduced amino acid sequence identity of 3936 isolate with CVB5 prototype strain Faulkner and other CV-B5 strains

	Sequence identity with CV-B1 Conn-5 strain		Sequence identity with other CV-B1 strains	
Genetic region	**Nucleotide**	**Amino acid**	**Nucleotide**	**Amino acid**
Genome	79%	94%	79.1-84.6%	87.4-92.4%
5' UTR	87%	-	85.1-93.8%	-
P1				
VP4	79%	93%	77.3-86.5%	89.9-95.6%
VP2	79%	94%	79.1-88.9%	93.5-97.3%
VP3	78%	91%	75.6-87.8%	80.2-90.8%
VP1	78%	94%	76.2-89.1%	89.1-97.2%
P2				
2A	78%	91%	76.1-88.2%	86.0-94.0%
2B	76%	94%	76.4-79.8%	93.9-97.9%
2C	81%	99%	79.1-83.4%	96.7-98.5%
P3				
3A	67%	97%	77.5-83.5%	93.3-96.6%
3B	73%	96%	72.7-83.3%	90.9-95.5%
3C	78%	95%	77.9-81.1%	91.8-95.1%
3D	77%	94%	77.4-84.8%	93.3-95.5%
3' UTR	87%	-	81.0-88.0%	-

Table 6. Evolutionary divergence of coding regions in CV-B5 isolates*

Genetic region	Length (nt)	Average dN	Average dS	dN/dS	Evolutionary selection
P1					
VP4	207	0.014	0.484	0.029	Negative
VP2	789	0.010	0.447	0.022	Negative
VP3	714	0.025	0.457	0.054	Negative
VP1	849	0.016	0.427	0.038	Negative
P2					
2A	450	0.032	0.442	0.072	Negative
2B	297	0.015	0.413	0.036	Negative
2C	987	0.013	0.615	0.021	Negative
P3					
3A	267	0.019	0.683	0.028	Negative
3B	66	0.021	0.688	0.030	Negative
3C	549	0.025	0.665	0.038	Negative
3D	1386	0.024	0.686	0.035	Negative

*Complete coding sequence (cds) information was retrieved from the following isolates: Faulkner (AF114383), 1954/85/UK (X67706), 2000/CSF/KOR (AY875692), 19CSF (JX017381), 20CSF(JX017380), SH1 (GU376747), COXB5/Henan/2010 (HQ998851), CVB5/CC10/10 (JN580070), A210/KM/09 (JX843811), CVB5/SD/09 (JX276378), CV-B5/P727/2013/China (KP289438), 2009-148-4 (KP266578), 2009-154-2 (KP266577), 2001-01-101 (KP266576), 2001-01-11 (KP266575),03001N (JX017383), 17Y (JX017382) and 3936/09 (KP233830).

Table 7. Nucleotide and deduced amino acid sequence identity between E7-Wallace and 3988 isolates

	Length (nt)		Sequence identity	
Genetic region	**Wallace**	**3988**	**Nucleotide**	**Amino acid**
Genome	7427	7429	80%	96%
5' UTR	742	742	86%	-
P1				
VP4	207	207	79%	97%
VP2	783	783	79%	98%
VP3	714	714	79%	94.5%
VP1	894	894	78%	94%
P2				
2A	432	432	79%	92%
2B	297	297	78%	97%
2C	987	987	82%	98%
P3				
3A	267	267	77%	94%
3B	66	66	71%	95%
3C	549	549	80%	97%
3D	1389*	1389*	80%	96%
3' UTR	100	102	87%	-

* including stop codon

Table 8. Evolutionary divergence of coding regions in echovirus 7 isolates*

Genetic region	Length (nt)	Average dN	Average dS	dN/dS	Evolutionary selection
P1					
VP4	207	0.016	0.764	0.021	Negative
VP2	783	0.017	0.703	0.024	Negative
VP3	714	0.017	0.680	0.025	Negative
VP1	849	0.024	0.663	0.036	Negative
P2					
2A	432	0.053	0.666	0.079	Negative
2B	297	0.025	0.725	0.034	Negative
2C	987	0.015	0.755	0.020	Negative
P3					
3A	267	0.028	0.676	0.041	Negative
3B	66	0.031	0.613	0.050	Negative
3C	549	0.026	0.742	0.035	Negative
3D	1386	0.024	0.688	0.034	Negative

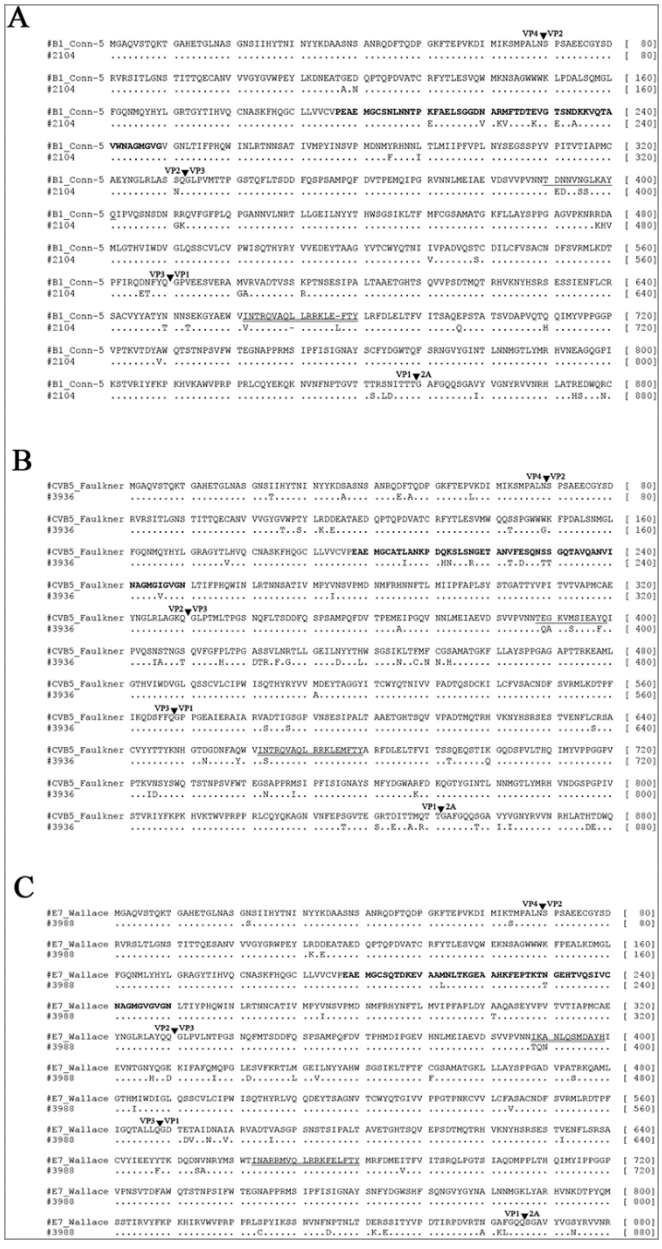
*Complete coding sequence information was available for the following isolates: Wallace (AY302559), UMMC (AY036578), DH22G (KJ765699), 2001-31 (KP266570) and 3988/09 (KP202389)

The 3'UTR of the Conn-5 strain has a thermodynamically more favourable free energy level (ΔG= -28 Kcal/mol) compared with that of the 2104 isolate (ΔG= -23.8 Kcal/mol). The 2A, 2B, 3A and 3D proteins had the highest level of amino acid sequence variability (Table 3). Examination of relative frequencies of synonymous and nonsynonymous substitutions in different genomic regions of the CV-B1 isolates showed no evidence for selective pressure (Table 4). Based on alignment with the three-dimensional structure of coxsackievirus B3 [15], a number of amino acid polymorphisms could be mapped to specific capsid regions. In comparison with the Conn-5 strain, seven amino acid substitutions were noted in the Puff region of the VP2 protein (Figure 3A, in bold), four in the Knob region of the VP3 protein (Figure 3A, underlined), and two within the hydrophobic pocket region of the VP1 protein (Figure 3A, double underlined).

The similarities between the 3936 isolate and the CV-B5 prototype strain (Faulkner) were 79% for nucleotides and 94 % for amino acids (Table 5). The highest nucleotide and amino acid sequence similarity was found with the 1654/85/UK strain (84.6% for nucleotides and 92.4 % for amino acids). The 5'UTR of the 3936 strain was 1 nucleotide longer than that of the Faulkner strain. Both the 5' and 3'UTR RNA secondary structures of the 3936 isolate (ΔG= -262.6 Kcal/mol and -45.4 Kcal/mol) have a thermodynamically more favourable free energy level compared with those of the Faulkner strain (ΔG= -255.4 Kcal/mol and -35.3 Kcal/mol). When the extra nucleotide in the 5'UTR of the 3936 isolate was removed, the RNA folded again at a thermodynamically more stable free energy level (ΔG= -262.8 Kcal/mol) than that of the Faulkner strain. No deletions or insertions were observed in the ORF region. The highest level of amino acid sequence variability were detected at the VP3 and 2A proteins (Table 5). Nine amino acid substitutions were noted in the Puff region of the VP2 protein (Figure 3B, in bold), four in the Knob region of the VP3 protein (Figure 3B, underlined), and one within the hydrophobic pocket region of the VP1 protein (Fig 3B, double underlined). Calculation of the dN/dS ratios indicated no evidence of selective pressure exerted in the structural and nonstructural proteins (Table 6).

The 3988 isolate has 80% nucleotide and 96% amino acid sequence similarity with the echovirus 7 prototype strain Wallace (Table 7). The highest nucleotide similarity was found with the DH22G strain (82.88%) and the UMMC strain (82.85%), whereas the amino acid sequence similarity was higher with the DH22G strain (97.30%) than with the UMMC strain (96.89%). The 3'UTR of 3988 strain is 2 nucleotides longer than that of the Wallace strain. The two virus genomes share identical nucleotide lengths within the ORF and 5'UTR. The 5'UTRs of the Wallace and 3988 strains have comparable minimum free energy secondary structures (ΔG= -261.0 Kcal/mol and -264.7

Figure 3. Alignment of the deduced amino acid sequences of the P1 region of prototype and clinical CV-B1 (A), CV-B5 (B) and echovirus 7 (C) strains. The cleavage sites and the adjacent functional proteins are indicated by downward-pointing arrows. The amino acids forming the VP2 puff region are in bold. The amino acids forming the VP3 knob region are underlined. The sequence corresponding to the VP1 hydrophobic pocket is double-underlined.



Kcal/mol, respectively). However, the 3'UTR of the 3988 strain has a thermodynamically more favourable free energy level (ΔG= -35.5 Kcal/mol) compared with that of the Wallace strain (ΔG= -31.1 Kcal/mol). The 2A, 3A and VP1 proteins of the 3988 isolate had the highest level of amino acid sequence variability compared to other viral proteins. Two amino acid substitutions were noted in the Puff region of the VP2 protein (Figure 3C, in bold), and three in the Knob region of the VP3 protein (Figure 3C, underlined). Nonsynonymous mutations were not detected in the hydrophobic pocket region of the VP1 protein (Fig 3C, double underlined). Furthermore, no evidence for selective pressure was found in the structural and nonstructural coding regions (Table 8).

Discussion

The CV-B1 2104 strain was isolated in 2008 from cerebrospinal fluid of 1-year-old female with respiratory distress and neurological signs. It had shown decreased susceptibility to the antiviral activity of MxA protein [16], and increased potential to induce inflammatory cytokines [17], compared to the prototype strain. The molecular basis for the phenotypic differences observed between the clinical and prototype CVB1 has not been yet investigated. The 2104 isolate shared high nucleotide and amino acid sequence similarity with CV-B1 Chi07 strain which caused 6 neonatal deaths in USA in 2007 due to heart failure [18,19]. However, results of similarity plot and bootscanning analyses suggest that a recombination had occurred in the structural and nonstructural coding regions between CV-B1 11167438_pmMC and echovirus 1 Farouk.

The CVB5 3936 strain was isolated in 2009 from stool of 4-year-old male with aseptic meningitis. It is most closely related to CV-B5 1954/85/UK strain that was isolated from the vesicular fluid within skin lesions of a patient with hand food and mouth disease (HFMD) in UK in 1985; its structural coding region was reported to be most closely related to that of swine vesicular disease virus [20]. However, according to the bootscan analysis, the 3936 strain is a mosaic virus resulting from recombination between CV-B5 1954/85/UK, EV-B86 BAN00-10354 and EV-B88 BAN01-10398 strain. The last two strains were isolated from stool specimens of patients presenting with acute flaccid paralysis (AFP) in Bangladesh during the years 2000 and 2001 [21].

The E7 3988 strain was isolated in 2009 from the blood of a neonate with sepsis-like illness. It shared high nucleotide and amino acid sequence similarity with E7 UMMC at P1 genetic region. UMMC strain was isolated from the cerebrospinal fluid of a Malaysian child with HFMD and fatal encephalomyelitis in 2000 [22,23]. P2 region was most closely related to EV-B86 BAN00-10354 strain, whereas P3 region was most closely related to E7 DH22G strain that was isolated from the stool of a Chinese patient with HFMD in 2012 [24]. Recombination between E7 UMMC, E7 DH22G, and EV-B86 BAN00-10354 strains was detected with strong bootstrap support.

There was no evidence of positive selective pressure exerted on the three EV strains tested in this study. However, a sequence divergence of 12-14% was detected in the 5'UTR between the clinical isolates and their corresponding prototype strains. Mutations in the EV 5'UTR which plays an important role in the control of replication and translation [25], have been proposed as major determinants of cardiovirulence [26], neurovirulence [23], and myopathogenicity [27]. Similar to the 5′ UTR, the EV 3′ UTR contains secondary structural elements and has been shown to play a role in both translation and replication [28]. It is noteworthy that the secondary structures of both untranslated regions of the 3936 isolate fold at a thermodynamically more favourable free energy level than those of the prototype strain. More investigations are required to determine whether the lower free energy level of EV RNA secondary structures observed in this study is associated with viral virulence and more severe clinical outcome, as has been described earlier for EV71 [29].

Interestingly, compared to the prototype strain, several amino acid substitutions were noted in the Puff, Knob, and the hydrophobic pocket regions of the viral capsid protein. The VP2 Puff region is known to be the largest and most variable surface loop of VP2 among group B coxsackieviruses [15]. It is a major neutralization site in both the polioviruses and rhinoviruses, suggesting its role in receptor binding or recognition [30]. The three clinical isolates in this study had amino acid substitutions in the Puff region of the VP2 protein. Amino acid substitutions in the Puff region of CVB3 VP2 protein were previously linked to cardiovirulence [31]. Like the Puff region of VP2, the Knob region of VP3 is known to contain a major neutralization site for both poliovirus and rhinovirus [30,32]. Furthermore, mutations in the VP3 Knob region have been suggested to have a role in the CVB3-induced cardiovirulence [31]. Whether or not the nonsynonymous mutations detected in the capsid region of the three clinical isolates play a role in virulence and pathogenesis of enterovirus infection, needs further investigation.

Conclusion

Phylogenetic analysis of the enterovirus strains suggested that the cases resulted from circulation of different genetic lineages of species B enteroviruses. Results from the similarity plots and bootscan analyses indicated accumulation of multiple inter- and intratypic recombination events along the EV genome. Whether the observed recombination results or not into better adaptation to new environments, or more potential to persist, replicate and cause disease, needs to be investigated.

Acknowledgements

The authors would like to thank Aljawhara Al-Sabah and Susan Silpi Kurien for their valuable technical support.

References

1. Sawyer MH (1999) Enterovirus infections: diagnosis and treatment. Pediatr Infect Dis J 18: 1033-1039.
2. Racaniello VR (2007) Picornaviridae: the viruses and their replication. In Knipe DM, Howley PM, editors. Fields Virology, 5th edition. Philadelphia: Lippincott Williams & Wilkins p. 795-838.
3. International Committee on Taxonomy of Viruses (ICTV) (2015) Order: Picornavirales. Available: http://www.ictvonline.org/virusTaxonomy.asp. Accessed 06 April 2015.
4. Dalwai A, Ahmad S, Pacsa A, Al-Nakib W (2009) Echovirus type 9 is an important cause of viral encephalitis among infants and young children in Kuwait. J Clin Virol 44: 48-51.
5. Dalwai A, Ahmad S, Al-Nakib W (2010) Echoviruses are a major cause of aseptic meningitis in infants and young children in Kuwait. Virol J 7: 236.
6. Ahmad S, Dalwai A, Al-Nakib W (2013) Frequency of enterovirus detection in blood samples of neonates admitted to hospital with sepsis-like illness in Kuwait. J Med Virol 85: 1280-1285.
7. Lukashev AN, Lashkevich VA, Ivanova OE, Koroleva GA, Hinkkanen AE, Ilonen J (2005) Recombination in circulating Human enterovirus B: independent evolution of structural and non-structural genome regions. J Gen Virol 86: 3281-3390.
8. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596-1599.
9. Reuter JS, Mathews DH (2010) RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics. 11: 129.
10. Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3: 418-426.
11. Yang Z, Bielawski JP (2000) Statistical methods for detecting molecular adaptation. Trends Ecol Evol 15: 496-503.
12. Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol and Evol 4: 406-425.
13. Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111-120.
14. Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783-791.
15. Muckelbauer JK, Kremer M, Minor I, Diana G, Dutko FJ, Groarke J, Pevear DC, Rossmann MG (1995) The structure of coxsackievirus B3 at 3.5 A° resolution. Structure 3: 653–667.
16. Chehadeh W, Abdulkareem HA (2010) Difference in susceptibility to MxA protein between a coxsackievirus B1 isolate and prototype, impact of serial cell culture passage. J Med Virol 82: 424-432.
17. Chehadeh W, Alkhabbaz M (2013) Differential TLR7-mediated expression of proinflammatory and antiviral cytokines in response to laboratory and clinical enterovirus strains. Virus Res 174: 88-94.
18. Verma NA, Zheng XT, Harris MU, Cadichon SB, Melin-Aldana H, Khetsuriani N, Oberste MS, Shulman ST (2009) Outbreak of life-threatening coxsackievirus B1 myocarditis in neonates. Clin Infect Dis 49: 759–763.
19. Quinn KK, Wollersheim SK, Krogstad P (2014) Complete genome sequence of coxsackievirus B1 isolated during case outbreaks in 2007 in the United States. Genome Announc 2: e00574-14.
20. Zhang G, Wilsden G, Knowles NJ, McCauley JW (1993) Complete nucleotide sequence of a coxsackie B5 virus and its relationship to swine vesicular disease virus. J Gen Virol 74: 845-853.
21. Oberste MS, Maher K, Nix WA, Michele SM, Uddin M, Schnurr D, al-Busaidy S, Akoua-Koffi C, Pallansch MA (2007) Molecular identification of 13 new enterovirus types, EV79-88, EV97, and EV100-101, members of the species Human Enterovirus B. Virus Res 128: 34-42.
22. Chua BH, McMinn PC, Lam SK, Chua KB (2001) Comparison of the complete nucleotide sequences of echovirus 7 strain UMMC and the prototype (Wallace) strain demonstrates significant genetic drift over time. J Gen Virol 82: 2629-2639.
23. Lum LC, Chua KB, McMinn PC, Goh AY, Muridan R, Sarji SA, Hooi PS, Chua BH, Lam SK (2002) Echovirus 7 associated encephalomyelitis. J Clin Virol 153: 153-160.
24. Yao X, Bian LL, Mao QY, Zhu FC, Ye Q, Liang ZL (2015) Echovirus 7 associated with hand, foot, and mouth disease in mainland China has undergone a recombination event. Arch Virol 160: 1291-1295.
25. Hunziker IP, Cornell CT, Whitton JL (2007) Deletions within the 5′UTR of coxsackievirus B3: consequences for virus translation and replication. Virology 360: 120-128.
26. Tu Z, Chapman NM, Hufnagel G, Tracy S, Romero JR, Barry WH, Zhao L, Currey K, Shapiro B (1995) The cardiovirulent phenotype of coxsackievirus B3 is determined at a single site in the genomic 5' nontranslated region. J Virol 69: 4607-4618.
27. Tam PE, Weber-Sanders ML, Messner RP (2003) Multiple viral determinants mediate myopathogenicity in coxsackievirus B1-induced chronic inflammatory myopathy. J Virol 77: 11849-11854.
28. Merkle I, van Ooij MJ, van Kuppeveld FJ, Glaudemans DH, Galama JM, Henke A, Zell R, Melchers WJ (2002) Biological significance of a human enterovirus B-specific RNA element in the 3′ nontranslated region. J Virol 76: 9900–9909.
29. Li R, Zou Q, Chen L, Zhang H, Wang Y (2011) Molecular analysis of virulent determinants of enterovirus 71. PLoS One 6: e26237.
30. Page GS, Mosser AG, Hogle JM, Filman DJ, Rueckert RR, Chow M (1988) Three-dimensional structure of poliovirus serotype 1 neutralizing determinants. J Virol 62: 1781–1794.

31. Stadnick E, Dan M, Sadeghi A, Chantler JK (2004) Attenuating mutations in coxsackievirus B3 map to a conformational epitope that comprises the puff region of VP2 and the knob of VP3. J Virol 78: 13987-14002.
32. Sherry B, Rueckert R (1985) Evidence for at least two dominant neutralization antigens on human rhinovirus 14. J Virol 53: 137-143.

Corresponding author
Dr Wassim Chehadeh
Department of Microbiology
Faculty of Medicine
Kuwait University
PO Box 24923 Safat
13310 Kuwait
Phone.: 965 4 63 65 19
Fax: 965 5 33 27 19
E-mail: wchehadeh@hsc.edu.kw

Conflict of interests: No conflict of interests is declared.