Evidence for enteroviral persistence in humans

Daniel N. Galbraith, Carron Nairn and Geoffrey B. Clements

Regional Virus Laboratory, Ruchill Hospital, Glasgow G20 9NB, UK

We have sought evidence of enteroviral persistence in humans. Eight individuals with chronic fatigue syndrome (CFS) were positive for enteroviral sequences, detected by PCR in two serum samples taken at least 5 months apart. The nucleotide sequence of the 5′ non-translated region (bases 174–423) was determined for each amplicon. Four individuals had virtually identical nucleotide sequences (> 97%) in both samples. The sequence pairs also each had a unique shared pattern indicating that the virus had persisted. In one individual (HO), it was clear that there had been infection with two different enteroviruses. In the remaining three individuals, the lack of unique shared features suggested that re-infection had occurred, rather than persistence. With the exception of HO, the sequences fell into a subgroup that is related to the Coxsackie B-like viruses.

It has long been recognized that in some cases of immune dysfunction such as agammaglobulinaemia enteroviruses can cause persistent infection in humans. Virus has been repeatedly isolated up to 23 months after initial culture (Wilfert *et al.*, 1977; O'Neill *et al.*, 1988). There is also some evidence for similar persistence in patients with heart muscle disease which has been reviewed recently (Muir & Archard, 1994). Enteroviral infection is a common feature of some groups of chronic fatigue syndrome (CFS) patients (Archard *et al.*, 1988; Clements *et al.*, 1995) and it has been suggested that enteroviral persistence may be occurring in some of these patients. To prove formally that persistence rather than re-infection is occurring it is necessary to identify a unique feature retained by serial viral isolates from one individual. We present here for the first time evidence for enteroviral persistence in humans based on sequence comparison of serial PCR products from the 5′ non-translated region (NTR). A group of CFS patients is being followed prospectively and those positive for enteroviral sequences in serum by PCR at two time points have been used. We show that in 4/8 cases closely related enteroviral sequences containing a unique shared pattern are detectable in sera of individual patients for up to 24 months, providing good evidence for viral persistence.

Each patient was assessed initially by a consultant in infectious disease. A thorough medical history was obtained, and physical examination and laboratory tests performed to exclude other causes of fatigue, thereby determining that the patients matched the Oxford criteria for CFS (Sharpe *et al.*, 1991). A blood sample was taken for detection of enteroviral sequences at first and subsequent attendance at the clinic. Each blood sample was tested for the presence of enteroviral sequences by PCR, using 'nested' primers specific for the 5′ NTR (Clements *et al.*, 1995). Appropriate systems were included to avoid contamination between samples during the PCR reactions. Positive and negative control samples gave expected results during each of the procedures (RNA from Coxsackievirus A9 was used as a positive control at all times to ensure successful nucleic acid amplification).

The PCR amplicons were sequenced as previously reported (Galbraith *et al.*, 1995). The pairs of sequences from each patient were compared using the computer program GAP

Author for correspondence: Geoff Clements.
Fax +44 141 946 2200.

The GenBank accession numbers of the sequences reported in this paper are X96897–X96912.

Table 1. GAP comparison statistics for pairs of enteroviral sequences obtained from individual CFS patients

Primers P and P9 (Galbraith *et al.*, 1995) were used to sequence the PCR products.

Patient (code)	Interval between samples (months)	Percentage similarity
1 (CR)	26	92·00
2 (HA)	10	98·20
3 (HO)	40	70·60
4 (MO)	5	97·50
5 (PA)	8	97·50
6 (TI)	24	99·20
7 (MC)	12	90·00
8 (HM)	41	89·50

```
     174                                                        214
Con  CTGTTACCCC  GGACTGAGTA  TC AATAAACT  GCTCACGCGG  TCGAAGGAGA
TI93 A.........  ..........  .. ........  ..........  ..........
TI95 A.........  ..........  .. ........  ..........  ..........
PA93 ....CT....  ..........  ..C.......  ..........  ..........
PA94 ..........  ..........  ..C.......  ..........  ..........
CR92 .....T....  ..........  .. ........  ..........  ..........
CR94 ..........  ..........  .. ....G...  ..........  CT........
MO93 ..........  ..........  .. ........  ..........  ..........
MO94 ..........  ..........  .. ........  .G........  ..........
HA93 ..........  ..........  .. ........  ..........  ..........
HA94 ..........  ..........  .. ........  ..........  ..........
MC93 ..........  ..........  .. ........  ..........  ..........
MC94 ..........  ..........  .. ....GG..  A......T..  CT........
HM90 ..........  ..........  .. ........  ..........  ..........
HM94 ..........  ....C.....  .. ....GG..  ...TG.....  CT........
HO91 ..........  ..........  .. ..A.G...  ..........  .T........
HO94 ..........  ........A.  CG ..G.G...  .....AAG..  C.........

     224                                                        264
Con  AAACG TCCGT  TACCCGGCTA  ACTACTTCGA  GAAACCCAGT  AACACCATGG
TI93 ..... .....  .......G..  ..........  ..........  .......C..
TI95 ..... .....  .......G..  ..........  ..........  .......C..
PA93 ..... .....  ......A..G  ..........  ..........  ..........
PA94 ..... .....  ......A..G  ..........  ..........  ..........
CR92 ..... .....  ......A..G  ..........  ..........  ..........
CR94 ..... .T...  ........C.  .T........  ..........  .C........
MO93 ..... .....  ..........  ..........  ..........  .......C..
MO94 ..... .....  ..........  ..........  ..........  .......C..
HA93 ..... .....  ......A..G  ..........  ..........  ..........
HA94 ..... .....  ......A..G  ..........  ..........  ..........
MC93 ..... .....  ......A..G  ..........  ..........  ..........
MC94 .....T..T..  ........C.  .T........  ......T...  .........A
HM90 ..... .....  ......A..G  ..........  ..........  ..........
HM94 ..... .T...  ........C.  .T........  A.........  .C..T.....
HO91 ..G.. .T...  ..T.....C.  ..........  ......T...  ......G...
HO94 GG... GGAAA  A....AC..G  G.........  ..........  ...T...C..

     274                                                        314
Con  AGATTGCGAA  GCGTTTCGCT  CAGCACACCC  CCAGTGTAGA  TCAGGCCGAT
TI93 .A........  .......A..  ......G...  .TG.......  ..........
TI95 .A........  ..........  ......G...  .TG.......  ..........
PA93 ..........  ..........  ..........  ..........  ..........
PA94 ..T.......  ..........  ..........  ..........  ..........
CR92 ..........  ..........  ..........  ..........  ..........
CR94 ..G.....C.  .T........  .C.....A..  .TG.......  .....T....
MO93 .A........  .......A..  ......G...  .TG.......  ..........
MO94 .A........  ..........  ......G...  .TG.......  ..........
HA93 ..........  ..........  ..........  ..........  ....T.....
HA94 ..........  ..........  ..........  ..........  ....T.....
MC93 ..........  ..........  ..........  ..........  ..........
MC94 .AG.....C.  .T........  .C.....A..  ..........  .....T....
HM90 ..........  ..........  ..........  ..........  ..........
HM94 ..G......G  .T........  .C.....A..  ..........  .....T....
HO91 .AG....AG.  .T........  ......TA..  ..........  .....T....
HO94 .N........  .T....AT.A  .TCA.AC...  .TG.......  .....A....
```

Fig. 1. For legend see opposite.

```
     324                                                        364
Con  GAGTCACCGC  ATTCCCCACA  GGCGACTGTG  GCGGTGGCTG  CGTTGGCGGC
TI93 ..........  G.........  ..T.......  ..........  ..........
TI95 ..........  G.........  ..T.......  ..........  ..........
PA93 ..........  ..........  ..........  ..........  ..........
PA94 ..........  ..........  ..........  ..........  ..........
CR92 ..........  ..........  ..........  ..........  ..........
CR94 ..........  .A........  ..........  ..........  ..........
MO93 ..........  G.........  ..T.......  ..........  ..........
MO94 ..........  G.........  ..T.......  ..........  ..........
HA93 ..........  ..........  ..........  ..........  ..........
HA94 ..........  ..........  ..........  ..........  ..........
MC93 ..........  ..........  ..........  ..........  ..........
MC94 ..........  .........G  ..G...C...  ..........  ..........
HM90 ..........  ..........  ..........  ..........  ..........
HM94 ..........  G........G  ......C...  ..........  ..........
HO91 ..........  .........G  ......C...  ..........  ..........
HO94 ..........  G.........  ..T.......  A.....A...  ..........

     374                                                        414
Con  CTG CCCATGG  GGCAACCCAT  GGGACGCTTC  AATATGGACA  TGGTGCGAAG
TI93 ... .......  ..........  C.........  ..........  ..........
TI95 ... .......  ..........  C.........  ..........  ..........
PA93 ... .......  ..........  ....T.....  ......CC..  ..........
PA94 ... .......  ..........  ....T.....  ......CG.T  A.........
CR92 ... .......  ..........  ..A.T.....  ..........  ..........
CR94 ... .......  ..........  .A........  ....CT....  ..........
MO93 ... .......  ..........  ..........  ..........  ..........
MO94 ... .......  .....A....  C.........  .......C..  ..........
HA93 ... .......  ..........  ....T.....  .......G..  ..........
HA94 ... .......  ..........  ....T.....  ......CN.T  AT........
MC93 ... .......  ..........  ....T.....  ..........  ..........
MC94 ... .......  .AT.......  ..........  ....CT....  ..........
HM90 ... .......  ..........  ....T.....  ......ACAT  ..........
HM94 ... .......  ..........  .......C..  ....CT....  ..........
HO91 ... .......  ..A.......  ........CT  ....CA....  ..........
HO94 ..AA.....T.  .....T....  ...C......  G...CC.CG.  ..........
```

Fig. 1. Nucleotide sequences of the partial 5′ NTR of enterovirus isolates from CFS patients as described in Table 1. Differences between the sequences and a consensus sequence (Con) are shown. Numbers refer to nucleotide positions of the complete genome of Coxsackievirus B3 (GenBank accession number M16572). N indicates either A or G at this position.

(GCG, Wisconsin package, version 8) and subsequent phylogenetic comparison was made using the programs PILEUP, DISTANCES and GROWTREE (GCG, Wisconsin package, version 8).

Table 1 gives the GAP comparison statistics for the maximum available sequence of the pairs obtained from individual patients. Four of the eight pairs of sequences [patients 2 (HA), 4 (MO), 5 (PA) and 6 (TI)] demonstrate a high level of similarity of 97·5 % or greater with samples taken up to 24 months apart. In the case of patient 6, the 0·8 % difference equates to 1 bp change within the region analysed (254 bases).

Fig. 1 presents the pairs of sequences from the eight patients compared to a consensus sequence of 250 bases. Four of the pairs of sequences (TI, PA, MO and HA) show their own unique pattern that is different from the consensus sequence. For example, pair TI at base numbers 174 and 241, pair HA at 318 and pair PA at 196 (where there is an additional cytosine). Pair MO shows features similar to TI but is missing that at 174. A comparison using the pairs TI, PA, MO and HA was made with 34 sequences derived from additional CFS patients from whom only a single sample was available (data not shown). Comparing these sequences with pair TI, one contained guanine at position 241, but did not have the feature at 174 and furthermore was dissimilar at seven other bases. None of the 34 sequences in the comparison contained the inserted cytosine present in pair PA. In the case of pair MO, the features were present in one of the comparison sequences, and in pair HA the thymine at position 318 was present in two of the 34 sequences. However, the rest of these three individual sequences showed at least ten differences from the consensus

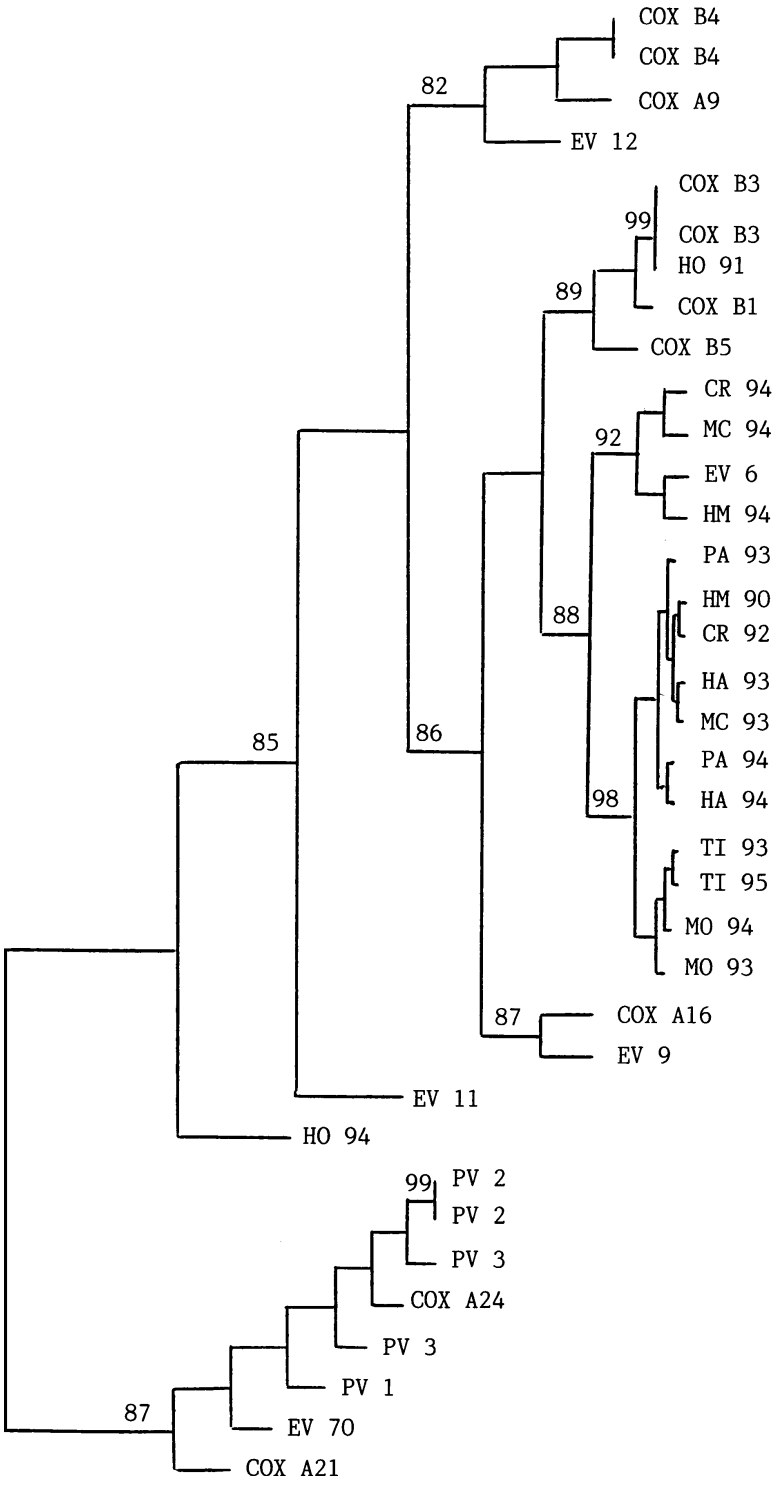

Fig. 2. Dendrogram of the genetic relationship between the 5′ NTRs of pairs of enteroviral sequences derived from PCR amplicons from patients with CFS. Included in the dendrogram are known enteroviral sequences obtained mainly from GenBank/EMBL sequence libraries, with accession numbers in parentheses: PV 1, poliovirus type 1 (V01150); PV 2, poliovirus type 2 (M12197, D00625); PV 3, poliovirus type 3 (K01392, X04468); Cox A24, Coxsackievirus A24 (D90457); EV 70, enterovirus 70 (D00820); Cox A21, Coxsackievirus A21 (D00538); EV 12, ECHOvirus 12 (X77708); EV 11, ECHOvirus 11 (X80059); EV 6, ECHOvirus 6 (U16283); Cox B3, Coxsackievirus B3 (M16572, M33854); Cox B1, Coxsackievirus B1 (M16560); EV 9, ECHOvirus 9 (X84981); Cox A16, Coxsackievirus A16 (U05876); Cox A9, Coxsackievirus A9 (D00627); Cox B4, Coxsackievirus B4 (X05690, D00149); Cox B5, Coxsackievirus B5 (X67706). The figures on the dendrogram indicate the percentage similarity of the sequences within each branch.

sequence and thus could not be related closely to the genomes of patients MO and HA. The pairs of sequences from CR, MC and HM do not share unique identifying features but overall there is a close similarity to the other CFS sequences. In the case of MC and HM, the second sequences share 15 nucleotide changes, the most likely explanation of which is re-infection of both by a related circulating enterovirus strain. These changes are present in a number of the 34 sequences sampled in 1994. The second specimens from these patients were taken in March and May in 1994. The interval between the two samples in the case of MC was 12 months and in HM 41 months.

As has been reported previously (Galbraith *et al.,* 1995) the

sequences from CFS patients form a group demonstrating a close genetic relationship with each other, and fall into a subgroup that is related to the Coxsackie B viruses. In this study, phylogenetic analysis (Fig. 2) demonstrated that 7/8 of sequences from the CFS patients grouped together. Two pairs of sequences from patients 4 (MO93 and MO94) and 6 (TI93 and TI95) group alongside each other showing a high degree of similarity. The five others [patients 1 (CR92 and CR94), 2 (HA93 and HA94), 5 (PA93 and PA94), 7 (MC 93 and MC94) and 8 (HM90 and HM94)] also group closely together but cannot be identified as belonging to one of the known enterovirus groupings on the basis of sequence comparison in this region.

The two sequences from patient 3 (HO91 and HO94) mapped to separate sites on the phylogenetic tree as expected from the similarity figures (70·6%). This strongly suggests that these sequences are derived from two enteroviruses that caused separate infections. The enteroviral sequence derived from the serum of patient 3 (HO91) was 99% identical with a Coxsackie B3 virus sequence (GenBank accession number M33854) and the other was most similar to ECHOvirus 6 (GenBank accession number U16283).

We have previously reported an association between enterovirus and CFS in slightly less than 50% of patients (Clements *et al.*, 1995). In virtually all cases where we have sequenced enterovirus amplicons from CFS patients, they have proved to be atypical. Furthermore these atypical sequences have only been found in one comparison non-CFS patient (Galbraith *et al.*, 1995). By taking sequential samples we have now directly sought evidence for viral persistence. However, determining if a particular virus has persisted presents considerable difficulties when dealing with RNA viruses. Enteroviruses cannot replicate their RNA genome without mistakes occurring due to the high error rate of the RNA-dependent RNA polymerase ($1–5 \times 10^{-3}$ per base for each replication cycle) and the absence of a proof reading function. Consequently, as the number of replication cycles increases, the divergence from the original sequence also increases. There are constraints on this process, however, as some of these changes could make the virus non-viable.

Our results show that in 4/8 cases paired enteroviral sequences have at least 97·5% similarity and also a unique shared pattern in each individual. The simplest explanation is that these sequences have persisted in these patients and there has been some divergence in the sequence due to errors in replication. Some of the nucleotide positions undergo changes more frequently than others, suggesting that there are some constraints on the variation. There is evidence from poliovirus that this region has a stem–loop secondary structure which may explain the constraints on variation. In our analysis we are likely to sample only the most common sequence type present in the serum samples, therefore the true extent of the heterogeneity of these enteroviral sequences is yet to be determined. Patient 3 (HO) in this study shows indications of two distantly related viral sequences, which provides evidence for there being separate infections with different enteroviruses. In the case of MC and HM, the second samples were taken within 3 months of each other and the sequences shared 15 nucleotide changes different from the consensus sequence. The most likely explanation is a re-infection of both with a related strain. An alternative explanation, which is less likely, is convergent evolution. In the case of patient CR, the differences between the two sequences also suggest the probability of a re-infection having occurred.

Co-existence of populations of different enteroviral sequences has been shown in poliovirus where reversion of attenuated vaccine strains to a neurotropic type can occur in an individual (Kinnunen *et al.*, 1990).

There have been no molecular studies carried out on non-polio enteroviruses which persistently infect human subjects. However studies using foot-and-mouth disease virus (a picornavirus closely related to enteroviruses) did show that 0·2–2·4% of nucleotides changed over 5 months in infected cattle (Malirat *et al.*, 1994). This investigation also found evidence for mixtures of populations of virus in infected animals evolving independently over time. This is analogous to our findings reported here.

A clinical evaluation of these patients has been carried out and will be presented separately. In all eight cases the symptoms persisted essentially unchanged over the time between the two samples. There was no evidence for there being a clinical difference between patient 3, in whom we have evidence of a re-infection, and the other seven. The site of persistence also remains to be determined, but detecting the presence of viral sequences in serum reflects a viraemia which indicates a replication site somewhere in the body. It is clear therefore that the complete pathology of enteroviruses is yet to be determined and requires further investigation. In this study we have reported strong evidence for persistence of enteroviruses in some individuals with CFS.

This research was supported by grants from the Linbury Trust. We would like to thank clinicians from the Department of Infectious Disease and Tropical Medicine, Ruchill Hospital for their help. We also acknowledge the staff of the Regional Virus Laboratory for their assistance. We also thank Dr Keith Vass and Mr Robert MacFarlane of CRC Beatson Laboratories for their assistance with the computer analysis and sequencing, respectively. We thank two of the referees for drawing our attention to similarities in the sequence of the second specimens from MC and HM.

References

Archard, L. C., Bowles, N. E., Behan, P. O., Bell, E. J. & Doyle, D. (1988). Post viral fatigue syndrome: persistence of enterovirus RNA in muscle and elevated creatine kinase. *Journal of the Royal Society of Medicine* **81**, 326–329.

Clements, G. B., McGarry, F., Nairn, C. & Galbraith, D. N. (1995). Detection of enterovirus-specific RNA in serum: the relationship to chronic fatigue. *Journal of Medical Virology* **45**, 156–161.

Galbraith, D. N., Nairn, C. & Clements, G. B. (1995). Phylogenetic

analysis of short enteroviral sequences from patients with chronic fatigue syndrome. *Journal of General Virology* **76**, 1701–1707.

Kinnunen, L., Huovilainen, A., Pöyry, T. & Hovi, T. (1990). Rapid molecular evolution of wild type 3 poliovirus during infection in individual hosts. *Journal of General Virology* **71**, 317–324.

Malirat, V., Mello, P. A. D., Tiraboschi, B., Beck, E., Gomes, I. & Bergmann, I. E. (1994). Genetic variation of foot and mouth disease virus during persistent infection in cattle. *Virus Research* **34**, 31–48.

Muir, P. & Archard, L. C. (1994). There is evidence for persistent enterovirus infections in chronic medical conditions in humans. *Reviews in Medical Virology* **4**, 245–250.

O'Neill, K. M., Pallansch, M. A., Winkelstein, J. A., Lock, T. M. & Modlin, J. F. (1988). Chronic group A coxsackievirus infection in agammaglobulinemia: demonstration of genomic variation of serotypically identical isolates persistently excreted by the same patients. *Journal of Infectious Disease* **157**, 183–186.

Sharpe, M. C., Archard, L. C., Banatvala, J. E., Borysiewicz, L. K., Clare, A. W., David, A., Edwards, R. H. T., Hawton, K. E. H., Lambert, H. P., Lane, R. J. M., McDonald, E. M., Mowbray, J. F., Pearson, D. J., Petto, T. E. A., Preedy, V. R., Smith, A. P., Smith, D. G., Taylor, D. J., Tyrrell, D. A. J., Wessely, S., White, P. D., Behan, P. O., Rose, F. C., Peters, T. J., Wallace, P. G., Warrell, D. A. & Wright D. J. M. (1991). A report on chronic fatigue syndrome: guidelines for research. *Journal of the Royal Society of Medicine* **84**, 118–121.

Wilfert, C. M., Buckley, R. H., Mohanakumar, T., Griffith, J. F., Katz, S. L., Whisnant, J. K., Eggleston, P. A., Moore, M., Treadwell, E., Oxman, M. N. & Rosen, F. S. (1977). Persistent and fatal central nervous system ECHOvirus infections in patients with agammaglobulinemia. *The New England Journal of Medicine* **296**, 1485–1489.

Received 25 March 1996; Accepted 24 October 1996