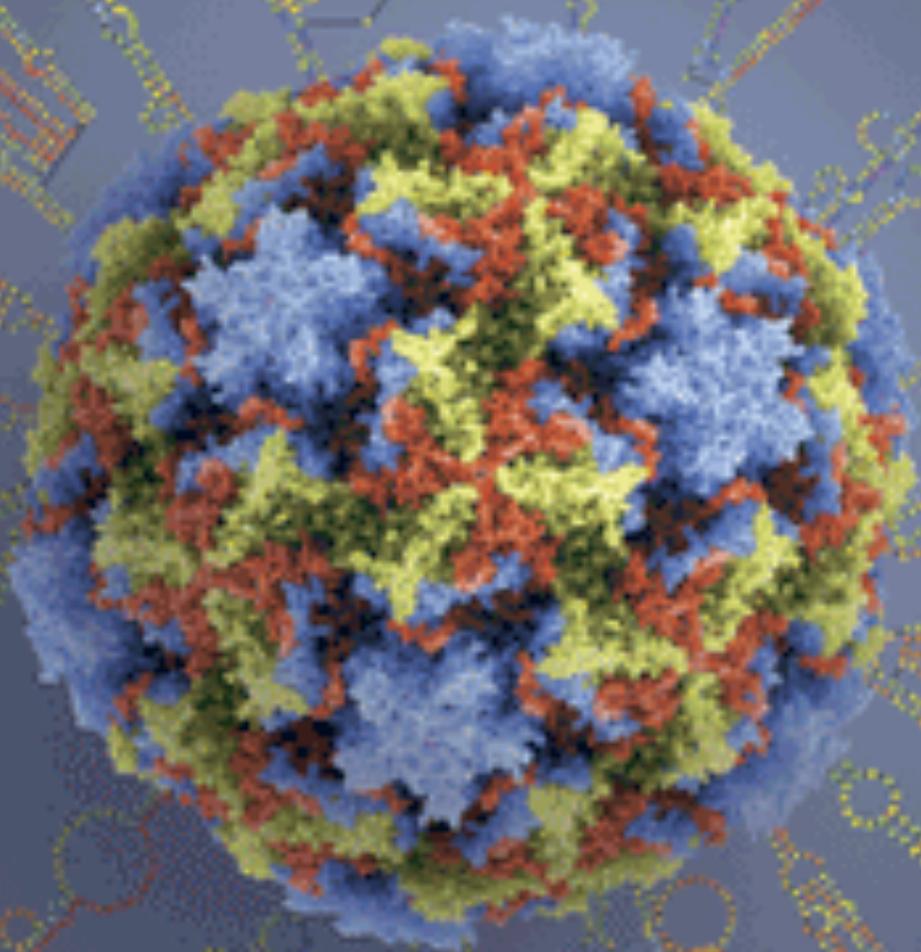


3 April 2009 | \$10

Science



Sequencing and Analyses of All Known Human Rhinovirus Genomes Reveal Structure and Evolution

Ann C. Palmenberg,^{1*} David Spiro,^{2*} Ryan Kuzmickas,² Shiliang Wang,² Appolinaire Djikeng,² Jennifer A. Rathe,³ Claire M. Fraser-Liggett,⁴ Stephen B. Liggett^{3†}

Infection by human rhinovirus (HRV) is a major cause of upper and lower respiratory tract disease worldwide and displays considerable phenotypic variation. We examined diversity by completing the genome sequences for all known serotypes ($n = 99$). Superimposition of capsid crystal structure and optimal-energy RNA configurations established alignments and phylogeny. These revealed conserved motifs; clade-specific diversity, including a potential newly identified species (HRV-D); mutations in field isolates; and recombination. In analogy with poliovirus, a hypervariable 5' untranslated region tract may affect virulence. A configuration consistent with nonscanning internal ribosome entry was found in all HRVs and may account for rapid translation. The data density from complete sequences of the reference HRVs provided high resolution for this degree of modeling and serves as a platform for full genome-based epidemiologic studies and antiviral or vaccine development.

Human rhinovirus (HRV), the disease agent for the common cold, is responsible for ~50% of asthma exacerbations and is one of the factors that can direct the infant immune system toward an asthmatic phenotype (1–4). Direct and indirect costs from the common cold and related complications in asthmatics amount to an estimated ~\$60 billion per year in the United States (5, 6). HRVs are single-stranded, positive-sense RNA enteroviruses in the *Picornaviridae* family and have been cataloged primarily by capsid serotyping relative to a historical repository of 99 strains, obtained from clinical specimens. HRVs are classified by their use of either intercellular adhesion molecule-1 (ICAM-1) (88 major viruses) or low-density lipoprotein receptor (LDLR) (11 minor viruses) as their receptor for cell entry (7). They have also been characterized by composite sensitivities across a panel of potential therapeutics (8) that have been used to parse the strains into two related drug-reactivity groups. The partial sequences of viral capsid-coding regions, noncoding regions, and a limited number of complete genomes have resulted in a division of the original 99 strains into two species: HRV-A (containing 74 serotypes) and HRV-B (containing 25 serotypes).

Recently, a number of previously unknown HRV-like sequences were detected in patients with influenza-like illnesses associated with severe respiratory compromise (9–11). The newly identified

viruses have not been cultured, but their sequences indicate that they likely represent a third (HRV-C) species. The lack of whole-genome sequence data for the full cohort of HRVs has made it difficult to understand basic molecular and evolutionary characteristics of the viruses and has hampered investigations of the epidemiology of upper respiratory tract infections and asthma epidemics. To define the extent and nature of HRV diversity and their evolution, we sequenced the genomes for every previously undetermined HRV in the reference repository, completing the full set of 99 serotypes, as well as 10 additional field samples.

Genome sequences and alignments. Modifications (12) were made to the sequence-independent, single-primer amplification (SISPA) method (13) to determine the complete genomes of 70 HRVs from the reference repository, as well as 10 nasal-wash samples from patients with HRV upper respiratory tract infections. We sequenced these viruses to an average of sixfold coverage for each of the ~7-kb genomes. To provide phylogenetic accuracy for these organisms with relatively small genomes and (often) high degrees of sequence similarity, a stringent approach was taken for aligning the sequences. The initial sequence fits for the polyproteins were performed on the basis of superimposition of the amino acid sequences within virion crystal structure maps (14) and supplemented with additional structure data from other viral proteins. In a stepwise manner, profile hidden Markov models (HMM) augmented the founder set with the remaining sequences (13). The published sequences (including redundant determinations) for the remaining serotypes were added so that the final collection consisted of 138 full-length HRV genomes, including at least one representative for each of the 99 original strains, 10 field samples, and 7 HRV-C strains (table S1). The genome-length RNA and polyprotein alignments for all considered sequences

are provided in tables S2 and S3. Regardless of species, all HRVs were found to have similar average base compositions. They are rich in A (31 to 34%) and U (25 to 30%) but low in G (19 to 22%) and C (18 to 22%). The third codon positions have the highest composition skew. An identity matrix for the polyproteins (fig. S1) shows that the average amino acid identity between pairs of HRV-C strains is slightly more diverse (78% identity, range 68% to 95%) than among the HRV-A (80%, range 64% to 99%) or HRV-B (83%, range 75% to 97%). The lack of broader diversity suggests that all HRVs are in a stable status for maintaining selection for certain traits, yet still have mutational flexibility for escape from immune responses.

Picornaviruses encode a single open reading frame (ORF) representing about 90% of the RNA length. Translation produces a polyprotein (~215 kD for HRV), subsequently cleaved in a viral protease-dependent cascade to form the 11 to 12 mature viral proteins required to initiate and sustain an infection (fig. S2A). Local and global RNA structures play established roles in HRV biology; however, the extent, character, and relatedness of serotype-specific 5'- and 3'UTR (untranslated region) variation are unknown. The role of this variability has thus not been related to function by modeling techniques or in vivo approaches. Our alignment methods included optimal energy RNA structure considerations (15) and were therefore sensitive to potential differences among the HRVs in the 5'- and 3'UTRs.

HRV RNA structures. All enteroviruses encode 5'-terminal cloverleaf-like motifs (CLs) that bind viral and cellular proteins for the initiation of RNA synthesis and also help convert infecting genomes from translation to replication templates. The HRV CLs [80 to 84 bases (b)] were predicted in every sequence with minimal structural variation among the species (representative structures are shown in Fig. 1A and additional structures in fig. S3). Immediately 3' to the CL, all HRVs were found to share an unusual pyrimidine-rich spacer segment with short oligo(C) and oligo(U) units interspersed with As (blue boxes, Fig. 1A and fig. S3A). The HRV-A have the shortest tracts (11 to 22 b) and HRV-B the longest (22 to 50 b). Nearly every HRV displayed a unique sequence in this region, and we identified unexpected variation even among isolates of the same serotype (Fig. 1A and fig. S3A). The equivalent genome location in poliovirus (10 b) interacts with poly(C)-binding protein 2 and is involved in the determination of the polio neurovirulent potential. The analogous regions in aphthoviruses or cardiociruses have homopolymeric poly(C) or poly(UC) tracts, the deletion of which markedly attenuates the virus through an RNA-activated protein kinase activation-dependent mechanism (16). If the HRV 5' spacer tracts are functional analogs to those of these other picornaviruses, then it is possible that the pathogenic potential of an individual HRV may also be encoded, in part, by this region.

Picornaviruses use internal ribosome entry sites (IRESs) to mediate translation initiation of their

¹Institute for Molecular Virology, University of Wisconsin, Madison, WI 53706, USA. ²J. Craig Venter Institute, Rockville, MD 20850, USA. ³Cardiopulmonary Genomics Program, University of Maryland School of Medicine, Baltimore, MD 21201, USA. ⁴Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD 21201, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: sligg001@umaryland.edu

polyprotein ORFs. The IRESs of all enteroviruses (termed type 1 IRESs) are thought to bind 40S ribosomal subunits internally within their 5'UTR and to then scan additional nucleotides to find the proper initiator AUG (17). Our modeling of the known serotypes confirms that all HRV IRESs start just 3' to the pyrimidine-rich spacer tract. We found that the internal IRES sequences are highly conserved, with an average nucleotide identity of 82%. Indeed, this region of the genome has the greatest degree of identity among all HRV (fig. S2B), exceeding 95% for regional motifs within the IRES. However, we also observed that the dominant IRES sequence conservation did not extend completely to the initiator AUG and that for the 18 to 40 bases 5'- to this codon, the region scanned by ribosomes, there was little species-specific conservation (<60% nucleotide identity). Despite this, our folding predictions configured every one of these regions into virtually the same RNA motif, which suggests that this structure is conserved even when the underlying nucleotides are not (Fig. 1B and fig. S3B).

Near the bottom of a long [15 to 20 base pair (bp)] minimum energy unbranched stem, the ORF AUG was invariably paired with a conserved upstream noncoding AUG (green boxes, Fig. 1B), marking the 3' boundary of the IRES, and the normal launch point for 40S scanning. Every HRV genome fold maintains this pairing. We predict that HRVs use the proximity of these AUGs to orient the 40S for direct transfer to the proper codon, without the need for scanning through the intervening nucleotides. The sequence and length variation between the AUGs was consistent with the idea that ribosomes would bypass this region entirely if they jump from one AUG to the other. This IRES folding is essentially a bait-and-switch mechanism, which we predict may enhance HRV translation competitiveness. This paired AUG motif is unique to HRV and is not found in other enteroviruses (e.g., poliovirus).

The HRV 3'UTRs (40 to 60 b) begin with the ORF termination codon and extend to the genetically encoded poly(A) tail. The ORF terminators themselves (solid red boxes in Fig. 1C and fig. S3D) included UAG, UAA, and UGA codons. The codon selection often differed among isolates from the same serotype. Multiple additional terminators (tan boxes), in and out of the ORF, were identified that punctuated each segment (3 to 9 per UTR). Despite large differences (>40%) in nucleotide identity (Fig. 1C), it was noted that all HRVs maintain a 13- to 16-bp unbranched stem, covering 67 to 88% of the 3'UTR, immediately abutting the poly(A) tail (see also fig. S3D). Some 3' stems have small interior loops, but nearly all, except for the unusual sequences of clade-D, present 5-base terminal loops, anchored with apical U-G or U-A pairs. Inevitably, the 3' sides of these terminal loops display UAG or UGA terminator codons, which may or may not synchronize with the local ORF. The function of these 3' stems is unknown, but such conservation is usually indicative of a putative protein recognition motif (such as translation termina-

tion factors) or, alternatively, an RNA:RNA tertiary interaction between the terminal-loop segment and a different (unknown) region of the genome.

Phylogenetic relationships of the HRV. Multiple methods were used to compute and compare phylogenetic trees for the aligned RNA and protein data. Figure 2 is a neighbor-joining consensus tree (18), which considered the 5'- and 3'UTRs and the first and second codon positions for the RNA genomes. All major nodes of the tree topology were stable over a range of calculation parameters, regardless of whether they invoked minimum evolution, parsimony, unweighted pair group method with arithmetic mean, or maximum likelihood (ML) methods. (See fig. S6A for ML tree with bootstrap, likelihood ratio tests, and comparison to the neighbor-joining tree.) Statistical comparisons were performed between the ML tree generated from the full genomes and the ML constrained-topology trees generated from capsid sequences that have been used as a surrogate for serotypes (fig. S6, A to C) (13), with the null hypothesis that topologies of these limited-sequence trees were the same as that of the full genome-based tree. The approximately unbiased (AU) test calculated from

the multiscale bootstrap *P* values for the capsid VP1 and VP0 trees were <10⁻⁶⁹, strongly rejecting the null hypothesis. In a similar manner, ML constrained-topology trees generated from IRES and 3D-polymerase sequences were not consistent with the full genome tree (fig. S6, A and B). Additional statistical tests also confirmed these results from the AU (fig. S6, B and C) for all four of the constrained-topology trees. Taken together, these results suggest that trees based on sequences from small, albeit biologically important, regions of the HRV genome have a limited capacity to reflect the evolutionary relationships and underlying diversity of HRVs. The tree shown in Fig. 2 also accurately represents parallel calculations on the aligned polyprotein dataset (*p*-distances differed by <2%). All aligned full-length sequences contributed to the tree calculations, but published sequences (noted with asterisks on Fig. 2) are shown only as needed, to complete the cohort of reference serotypes. (Were the redundant published sequences illustrated, without exception they would lie on the same-serotype branches, with *p*-distances <0.5%). The outer rings of this figure designate major (M: ICAM-1) or minor (m: LDLR) receptor preferences, and also each

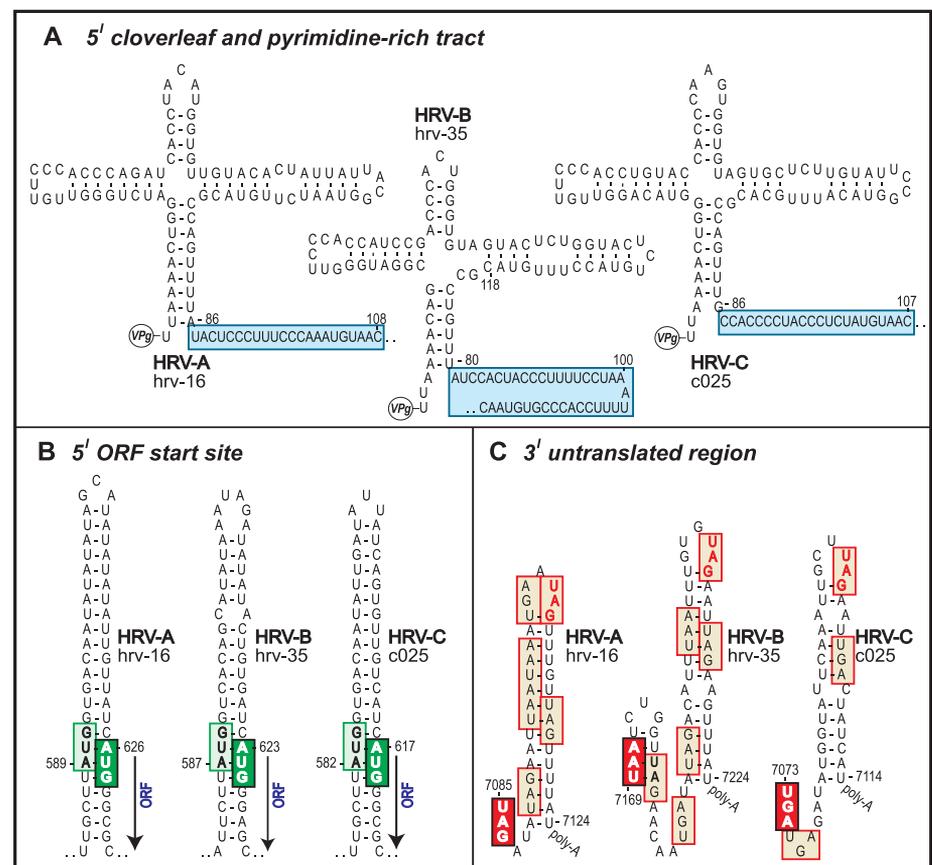


Fig. 1. Genome-wide optimal energy RNA configurations for select motifs representative of each HRV species. (A) All HRVs display characteristic 5'UTR CL elements with minimal predicted structural variation (blue boxes). (B) The alignments and predicted structures of the IRES of HRVs reveal a bait-and-switch arrangement for initiation of translation. A stem-pairing AUG (light green box) is found for each species, but a highly variable region of intervening sequence near the initiator AUG (dark green box) was noted. (C) HRV 3'UTRs have a unique unbranched stem motif before the poly-A tail. The left-most codon (red box, white text) is the ORF terminator. Other boxes highlight additional terminators (tan boxes), including a characteristic codon (tan box, red text) found in the apical loop.

strain's small-molecule drug reactivity group ("1" or "2") (8). The tree is shown rooted with three outgroup sequences from the *Human Enterovirus C* species (poliovirus 1m, coxsackieviruses a13 and a21), but the use of five additional outgroups had no effect on overall tree topology or HRV p-distances (fig. S6). The field samples (i.e., f01 to f10) were assigned tentative serotype names on the basis of capsid similarity relative to the known reference strains. Although some of the full genomes from these samples proved close to their repository cognates, a divergence in sequence for several strains was noted.

According to our tree(s), HRV-A and HRV-C share a common ancestor, which is a sister group to the HRV-B. Although the HRV-C clade currently has only seven full sequences, its genetic origin is clearly different from the reference set, and our phylogeny indicates that these represent a third HRV species, as has been recommended to the International Committee on Taxonomy of Viruses (19). The HRV-C have yet to be cultured or assessed for immunological cross-reactivity, but the sequence

space occupied by the available samples suggests that there may be many additional HRV-C strains awaiting discovery. Distance extrapolations relative to the new full reference cohort predict that HRV-C may have an even broader range of serotypes than the original 99, of which each confers only limited immunologic cross-protection to another.

A separate phylogenetic finding was the unexpected basal divergence within HRV-A of a small ($n = 3$) group of distinct strains, denoted clade D (20). Although the major basis for discriminating clade D from other HRV-A lies in their general, genome-length sequence divergence, these particular isolates have RNA elements—such as the cis-acting replication element, the 3'UTR terminal loop feature (see above and fig. S3), and local insertions/deletions and sequence motifs—that are somewhat atypical of other HRV-A strains. Some of the distinguishing characteristics are highlighted in fig. S4. Among all other major A clades, and major B clades, none have p-distances (>10%) that segregate them so distinctly. We are cautious in proposing clade D

as a fourth species (HRV-D), but the phylogenetic evidence (Fig. 2) and sequence characteristics (figs. S3 and S4) are highly suggestive. Other early topological divisions within HRV-A separate a major clade composed of 10 serotypes (counterclockwise, hrv-20 through hrv-12) from a second grouping composed of ~12 miniclades representing 61 serotypes (counterclockwise, hrv-89-f09 through hrv-100). These particular relationships were not readily apparent when only partial genome sequences were examined (21, 22). Fig. S6 shows a comparison of results from trees constructed with whole genomes, VP4/VP2, and VP1 sequence. Neither of the trees derived from these shorter sequences revealed the miniclades, clade D, or multiple other features. We thus contend that comparison of full-genome data, the context wherein evolutionary events occur, most likely provides the defining relationships among the HRVs, allows a more comprehensive assessment of strain diversity, and allows for more accurate historical extrapolations. The phylogenetic diversity we describe at the whole-genome level is consistent with the clinical heterogeneity of HRV infections in humans (1, 3, 4), although mapping specific clinical characteristics (i.e., incubation period, severity, respiratory compromise, and pro-asthmatic phenotypes) to responsible genomic regions will require additional field isolates from a large number of patients with multiple traits. Given the genome-wide diversity we have documented (e.g., 5' spacer elements, ORF start, protease, 3'UTRs), clinically relevant relationships may well depend on comparisons from multiple genome regions.

Recombination in HRV. Results from earlier sequencing of a subset of HRV reference genomes concluded that RNA recombination was not a major mechanism for HRV diversity (23, 24) and asserted that known isolates were independently segregating entities. We have reevaluated the potential for recombination by scanning the full reference set and the new field strains with a suite of recombination detection programs (25) relying on phylogenetic distance and sequence similarity. Stringent criteria ($P < 0.00001$ from two or more analyses modes) identified 23 genomes with probable origins resulting from at least 12 independent recombination events. Figure 3A shows representative data indicating that hrv-46 arose by recombination between hrv-53 (major parent) and hrv-80 (minor parent). Within the hrv-46 genome, nucleotides 32 to 3222 are most similar to hrv-80, whereas the rest of the genome (nucleotides 3223 to 7200) is common to hrv-53. The result is consistent with this trio's computed phylogenetic relationship (Fig. 2), placing the major parent (hrv-53) and the daughter (hrv-46) in the same clade and the minor parent (hrv-80) in a different, nearby clade. Results for all 23 identified recombination scenarios are summarized in Table 1. [See also (13) and table S4.] Of the recombination locales suggested by these events, the majority (10 of 12) involve the 5'UTR or the adjacent capsid genes, which seemingly have been collectively rearranged to produce at least 20 separate progeny strains. Among the

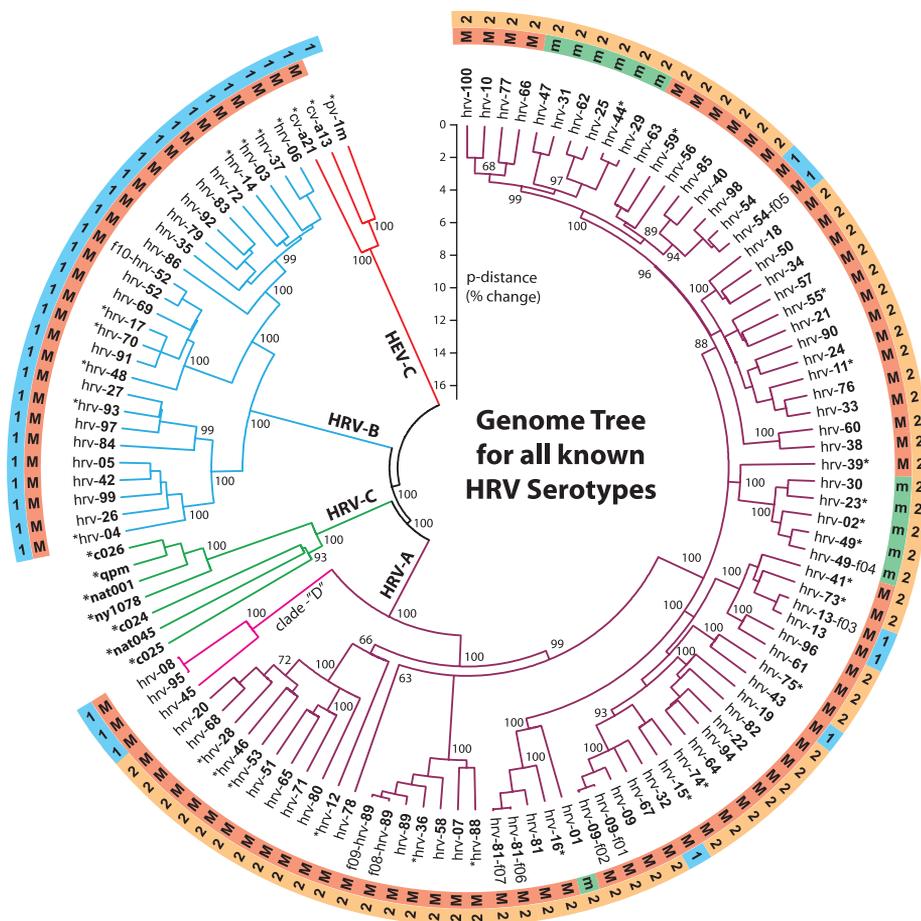


Fig. 2. A neighbor-joining (NJ) phylogenetic tree showing relationships between all known HRV serotypes created on the basis of full genome sequences. The HEV-C sequences (poliovirus 1M, coxsackievirus a13, and coxsackievirus a21) were used as outgroups. Branch lengths are proportional to similarity (p-distance). Key nodes on this tree are annotated with NJ bootstrap values (percentage of 2000 sampled trees). Asterisks in the strain names identify sequences obtained from published data (table S1). All other taxa are from this study. Letters in the outer rings designate whether that virus uses the major (M) ICAM-1 receptor or the minor (m) LDLR receptor (7) and whether its relative reactivity was more like group "1" or group "2" (if known) toward a panel of small-molecule antiviral compounds (8).

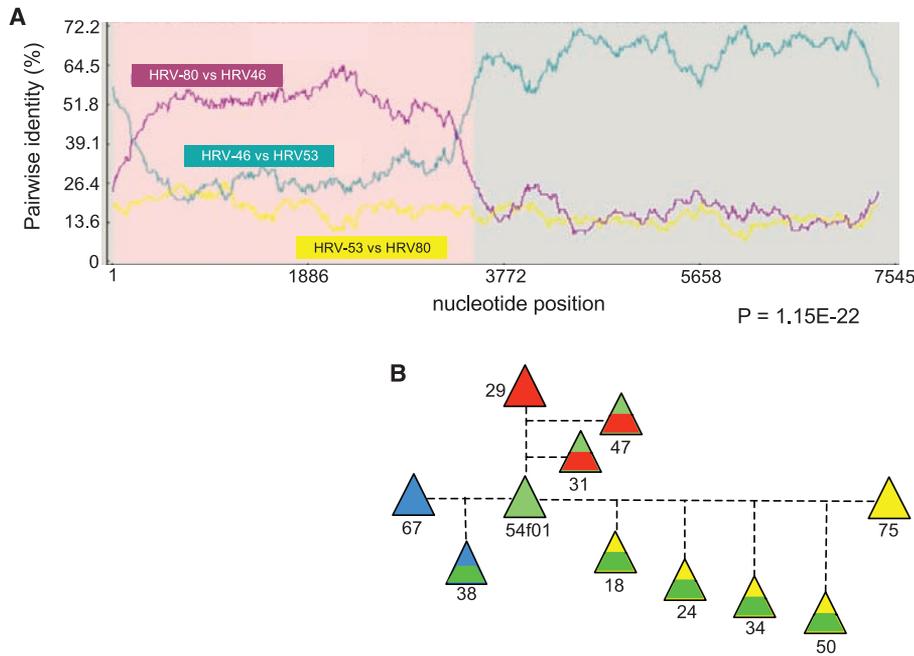


Fig. 3. Recombination of HRVs creates additional serotypes. **(A)** Representative results showing that hrv-46 arose from a recombination of hrv-53 (major parent) and hrv-80 (minor parent). Shown are normalized pairwise identities between each parent and the daughter hrv (purple and green) and the two parents (yellow). As indicated, hrv-46 nucleotides 32 to 3222 are from hrv-80, and nucleotides 3223 to 7200 are from hrv-53. **(B)** Recombination with an ancestor of hrv-54-f05 has resulted in seven serotype progeny. Each parental hrv is shown as a solid color. The contribution of each parent is proportional to the area of its color in the offspring. See Table 1 and table S4 for nucleotide boundaries and results from other recombination events.

Table 1. Recombination events in HRV serotypes. The *P* value listed is the lowest obtained. See table S4 for full data and *P* values.

Major parent	Minor parent	Recombinant	Genome region	<i>P</i>
hrv-45	hrv-21	hrv-8	5'UTR	1.094×10^{-21}
hrv-45	hrv-21	hrv-95	5'UTR	1.094×10^{-21}
hrv-65	hrv-21	hrv-80	5'UTR, VP4	6.129×10^{-23}
hrv-51	hrv-11	hrv-20	5'UTR, VP4	8.402×10^{-26}
hrv-51	hrv-11	hrv-68	5'UTR, VP4	8.402×10^{-26}
hrv-28	hrv-62	hrv-71	5'UTR	8.484×10^{-15}
hrv-54-f05	hrv-75	hrv-18	5'UTR, VP4, VP2, VP3, VP1	3.737×10^{-7}
hrv-54-f05	hrv-75	hrv-24	5'UTR, VP4, VP2, VP3, VP1	3.737×10^{-7}
hrv-54-f05	hrv-75	hrv-50	5'UTR, VP4, VP2, VP3, VP1, P2A	3.737×10^{-7}
hrv-54-f05	hrv-75	hrv-34	VP4, VP2, VP3, VP1	3.737×10^{-7}
hrv-54-f05	hrv-67	hrv-38	VP2, VP3, VP1, P2A, P2B, P2C, P3A	1.561×10^{-29}
hrv-54	hrv-67	hrv-60	VP2, VP3, VP1, P2A, P2B, P2C, P3A	1.561×10^{-29}
hrv-53	hrv-80	hrv-46	5'UTR, VP4, VP2, VP3, VP1	1.291×10^{-62}
hrv-13-f03	HRV41	hrv-73	5'UTR, VP4, VP2, VP3, VP1	9.419×10^{-12}
hrv-30	hrv-59	hrv-39	5'UTR, VP4, VP2, VP3, VP1	2.963×10^{-12}
hrv-68	hrv-20	hrv-28	VP4, VP2, VP3, VP1, P2A, P2B, P2C	6.839×10^{-12}
hrv-100	hrv-10	hrv-56	P2A, P2B, P2C, P3A, P3B, P3C, P3D	4.230×10^{-17}
hrv-29	hrv-54-f05	hrv-31	P3C, P3D	7.934×10^{-13}
hrv-29	hrv-54-f05	hrv-47	P3C, P3D	1.232×10^{-10}
hrv-42	hrv-97	hrv-4	5'UTR, VP4	8.233×10^{-35}
hrv-84	hrv-37	hrv-27	5'UTR, VP4	1.500×10^{-10}
hrv-84	hrv-37	hrv-93	5'UTR, VP4	1.500×10^{-10}
hrv-84	hrv-37	hrv-97	5'UTR, VP4	1.500×10^{-10}

138 full-length sequences, hrv-54 (or its ancestor) was apparently the most active in recombination. Field strain hrv-54-f05 links closely to three separate events (see Fig. 3B), contributing to at least seven different serotype progeny. In confirmation studies, we used a progressive alignment method for the HRV genomes (13), then repeated the suite of recombination detection programs. Of the 23 recombination events from the HMM alignment (Table 1), 19 were also found after the progressive alignment, although in some cases different major and/or minor parents were selected from within the same closely related clade (table S4).

Although the sequence fingerprints clearly trace these ancestral patterns, nevertheless all extant progeny also showed evidence of subsequent sequence divergence within the exchanged regions. Recombination was not identified between isolates from different species (i.e., HRV-A and HRV-B), but receptor binding preferences between ICAM and LDLR apparently presented no barriers to exchange. Major group hrv-54 and minor group hrv-29, for example, both contributed to the common ancestor of the minor group viruses, hrv-31 and hrv-47. These results, particularly for HRVs with different receptor preferences and those from distant clades, such as the parents hrv-21 and hrv-65, suggest that coinfection of the host with multiple parental strains is not uncommon and may lead to variant progeny with different biological properties. Our field isolates also deviated from the reference sequences in a manner that was not confined to any specific portion of the genome. In fact, field strain deviation relative to the reference isolates (for example, hrv-52-f01 versus hrv-52 differ by 838 nucleotides) was frequently greater than that observed between pairs of characterized serotypes (hrv-44 and hrv-29 differ by 385 nucleotides). Indeed, field samples of the same serotype collected from the same geographical region within 1 year showed marked variability (tables S2, S3, and S5). The propensity for such variation could underlie the marginal efficacy, or lack of efficacy, of anti-HRV therapeutics in clinical trials (26, 27). Given our reference set of full genomes, along with the means to rapidly sequence full genomes from field samples, new strategies may become apparent to engineer clade-specific agents by targeting their commonalities.

Conclusions. Our data complete and define the full set of genome-length sequences in the canonical reference repository of 99 HRV-A and HRV-B serotypes. Alignment and examination of these genomes confirmed species-specific sequence and RNA structural elements that differentiate the HRV-A and HRV-B from newly described HRV-C and further suggest that the HRV-A serotypes harbor a distinct, uncharacteristic clade, which may represent a fourth species (HRV-D). Local sequence variation, particularly in the 5'UTR, characterized each isolate within regions associated with the pathogenic potential for other picornaviruses. Parallel RNA structure comparisons defined several 5' and 3' elements as common to all isolates and unique to the HRV. Motifs like the AUG-presenting 5' ORF initiation stem, or the UAG-presenting 3' stem, may

contribute to HRV-specific IRES translation mechanisms, ORF termination, or polymerase recognition. We also found embedded within multiple sequences, including recent field isolates, clear evidence for repeated, historical genome recombination. Coinfection with multiple HRVs is known to occur (28), and we now know that this can lead to strains that may have distinct biologic properties and clinical characteristics. The required host environment for HRV recombination is not known, but with complete genome sequences from additional patient isolates such factors may become apparent. Our repository data set provides a baseline framework for the analysis of additional HRVs that may be in communities, including the HRV-Cs, and will enable larger-scale studies of basic molecular and evolutionary characteristics and assignment of disease phenotypes to specific genome regions. The clustering of small clades, the recombinations, and the mutations found in all regions of these genomes suggest that future HRV epidemiologic studies might benefit from full genome sequencing rather than the more limited serotyping. With such an approach, correlations may be more informative in inferring pathogenic potential and in designing antiviral agents and vaccines.

References and Notes

- J. M. Gwaltney Jr., J. O. Hendley, G. Simon, W. S. Jordan Jr., *N. Engl. J. Med.* **275**, 1261 (1966).
- S. L. Johnston *et al.*, *Am. J. Respir. Crit. Care Med.* **154**, 654 (1996).
- K. G. Nicholson, J. Kent, D. C. Ireland, *Br. Med. J.* **307**, 982 (1993).
- D. J. Jackson *et al.*, *Am. J. Respir. Crit. Care Med.* **178**, 667 (2008).
- A. M. Fendrick, A. S. Monto, B. Nightengale, M. Sarnes, *Arch. Intern. Med.* **163**, 487 (2003).
- K. B. Weiss, S. D. Sullivan, *J. Allergy Clin. Immunol.* **107**, 3 (2001).
- M. Vlasak *et al.*, *J. Virol.* **79**, 7389 (2005).
- K. Andries *et al.*, *J. Virol.* **64**, 1117 (1990).
- S. R. Dominguez *et al.*, *J. Clin. Virol.* **43**, 219 (2008).
- P. McErlean *et al.*, *PLoS ONE* **3**, e1847 (2008).
- S. K. Lau *et al.*, *J. Clin. Microbiol.* **45**, 3655 (2007).
- A. Djikeng *et al.*, *BMC Genomics* **9**, 5 (2008).
- Materials and methods are available as supporting material on Science Online.
- A. C. Palmenberg, J.-Y. Sgro, in *Molecular Biology of Picornaviruses*, B. L. Semler, E. Wimmer, Eds. (ASM Press, New York, 2001).
- A. C. Palmenberg, J.-Y. Sgro, *Semin. Virol.* **8**, 231 (1997).
- L. R. Martin, G. M. Duke, J. E. Osorio, D. J. Hall, A. C. Palmenberg, *J. Virol.* **70**, 2027 (1996).
- R. J. Jackson, in *Translational Control*, W. B. Hershey, M. B. Mathews, N. Sonenberg, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1996).
- K. Tamura, J. Dudley, M. Nei, S. Kumar, *Mol. Biol. Evol.* **24**, 1596 (2007).
- On the basis of these data, the International Committee on Taxonomy of Viruses Picornavirus Study Group recently proposed to recognize them as a new species (29).
- hrv-08 and hrv-95 are designated as two different serotypes by ATCC but differ by 67 nucleotides, perhaps indicating a misidentification in the repository.
- R. M. Ledford *et al.*, *J. Virol.* **78**, 3663 (2004).
- C. Savolainen, S. Blomqvist, M. N. Mulders, T. Hovi, *J. Gen. Virol.* **83**, 333 (2002).
- A. L. Kistler *et al.*, *Viral. J.* **4**, 40 (2007).
- P. Simmonds, *J. Virol.* **80**, 11124 (2006).
- D. Martin, E. Rybicki, *Bioinformatics* **16**, 562 (2000).
- R. Fleischer, K. Laessig, *Clin. Infect. Dis.* **37**, 1722 (2003).
- D. L. Barnard, *Curr. Pharm. Des.* **12**, 1379 (2006).
- C. Savolainen, M. N. Mulders, T. Hovi, *Virus Res.* **85**, 41 (2002).
- N. Knowles, personal communication.
- Funded by the University of Maryland School of Medicine internal funds and by NIH grant U19-AI070503. We thank A. Wolf and J.-Y. Sgro for implementation of the tree topology tests. The accession numbers for the hrv strains are FJ445111 to FJ445190 (see table S1).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1165557/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S5
References

5 September 2008; accepted 4 February 2009
Published online 12 February 2009;
10.1126/science.1165557
Include this information when citing this paper.

REPORTS

Photodegradable Hydrogels for Dynamic Tuning of Physical and Chemical Properties

April M. Kloxin,¹ Andrea M. Kasko,^{1,2*} Chelsea N. Salinas,¹ Kristi S. Anseth^{1,2,†}

We report a strategy to create photodegradable poly(ethylene glycol)-based hydrogels through rapid polymerization of cytocompatible macromers for remote manipulation of gel properties *in situ*. Postgelation control of the gel properties was demonstrated to introduce temporal changes, creation of arbitrarily shaped features, and on-demand pendant functionality release. Channels photodegraded within a hydrogel containing encapsulated cells allow cell migration. Temporal variation of the biochemical gel composition was used to influence chondrogenic differentiation of encapsulated stem cells. Photodegradable gels that allow real-time manipulation of material properties or chemistry provide dynamic environments with the scope to answer fundamental questions about material regulation of live cell function and may affect an array of applications from design of drug delivery vehicles to tissue engineering systems.

Hydrogels are hydrophilic polymers swollen by water that are insoluble owing to physical or chemical cross-links. These water-swollen gels are used extensively as biomaterials for complex device fabrication (1), cell culture for

tissue regeneration (2), and targeted drug release (3). Often, sophisticated control of the gel structure in space and time is required to elucidate the dynamic relationship between biomaterial properties and their influence on biological function (4, 5). For example, progenitor cells are often expanded and differentiated in hydrogel microenvironments, and researchers have demonstrated how the initial gel properties, including mechanics (6, 7) and chemical functionality (8), influence cellular fate. In regenerative medicine, the structure and composition of gels are also regulated temporally, through hydrolytic (9) and enzymatic (10–12) degradation mech-

anisms, to promote cell secretory properties and encourage the development of tissue-like structures *in vitro* and *in vivo*. A major challenge is determining which biochemical and biophysical features must be presented in a gel culture environment.

Hydrogel structure and functionality have evolved from the direct encapsulation of cells in simple homogeneous materials to those with highly regulated structures spanning multiple size scales [e.g., through self-assembly (13) or micro-engineering (14)]. These hydrogel structures are further modified locally by cells with the synthetic incorporation of bioresponsive functionalities (15) or externally by advanced patterning to create spatially varying functionalities. For example, the chemical patterning of a gel by the addition of a second, interpenetrating network or peptide tether has been demonstrated by diffusing chemical moieties into a gel and covalently linking these functionalities to the network by photocoupling (16) or reaction with a photolytically uncaged reactive group (17). Although these are important advances, such processes do not allow modulation of the gel chemistry in real time or photodegradation of the gel structure. Few synthetic materials provide a cellular microenvironment in which physical or chemical cues are initially present and subsequently regulated on demand. We have synthesized monomers capable of polymerizing in the presence of cells to produce photolytically degradable hydrogels whose physical or chemical properties are tunable temporally and spatially with light. The desired gel property for altering cell function or fabricating a

¹Department of Chemical and Biological Engineering, University of Colorado, 424 UCB ECCH 111, Boulder, CO 80309, USA. ²Howard Hughes Medical Institute, University of Colorado, 424 UCB ECCH 111, Boulder, CO 80309, USA.

*Present address: Department of Bioengineering, University of California, Los Angeles, CA, 90095–1600, USA.

†To whom correspondence should be addressed. E-mail: Kristi.Anseth@colorado.edu

Supporting Online Material

Table of Contents

Materials and Methods

Figure S1: Amino Acid Identities

Figure S2: Genome Map

Figure S3A, B, C, D: RNA Features

Figure S4: Unique Features of Clade-D

Figure S5: Distribution of Field Sample Variation

Figure S6A, B, C: Alternative Phylogenetic Tree, comparisons, and statistical analysis with constrained-topology trees

Table S1: Guide to Sequences and Strains

Table S2: RNA Alignments

Table S3: Amino Acid Alignments

Table S4: Recombination Analysis

Table S5: Specific Amino Acid Variations in Field Samples

Supporting Online Material References

Materials and Methods

HRV samples and preparation of viral nucleic acids

The HRV reference repository was considered to include all unique HRV serotypes available from American Type Culture Collection (Manassas, VA, (<http://www.atcc.org>)). These amounted to a total of 99 HRVs, as the previously designated “HRV-Hanks” is now considered hrv-21, hrv-87 is now classified as a strain of EV-86, a serotype within the human enterovirus D species, and hrv-1b and hrv-1a are now considered hrv-1. One described HRV serotype (hrv-57) was not available from ATCC and we utilized a field sample whose sequence was consistent with regions of the hrv-57 genome that had been previously reported. Additional field samples were obtained from the Wisconsin State Laboratory of Public Hygiene (Madison, WI) collected from 2005-2006. These isolates were amplified in HeLa cells, the virus was concentrated, then snap frozen. Full genome sequencing was performed as described below with 10^5 - 10^6 virions from the contents of one vial of frozen virus provided by the aforementioned sources. Briefly, viral RNA and DNA was prepared in a manner previously described in detail (*SI-S3*) with minor modifications. Each biological sample was first spun to remove cellular debris and processed through a 0.22 μ M filter to enrich viral particles in the flow-through while retaining bacteria and other cells in the filter. To eliminate residual nucleic acid contaminants in the filtrate, 100 units of DNase I and 3 μ g RNase A were added to the viral resuspension, followed by incubation at 37°C for 1 hour. RNA was extracted with Trizol-LS reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. The RNA pellet was resuspended in 20 μ l of nuclease-free water.

Construction of a library of random PCR fragments and sequencing

The extracted RNA was processed as previously described (S1-S3). Briefly, 800 ng of purified RNA was reverse-transcribed with SSII Reverse Transcriptase (Invitrogen) using the FR26RV-N primer (5' GCC GGA GCT CTG CAG ATA TCN NNN NN 3') at a concentration of 1 μ M. In addition, primer FR40RV-T (5' GCC GGA GCT CTG CAG ATA TC (T)₂₀ 3') was added at a concentration of 5 nM to specifically amplify the polyadenylated 3' end. The second strand was synthesized by addition of Klenow exo-polymerase (New England Biolabs, Ipswich, MA) in the presence of the FR26RV-N random primer. To capture the 5' end, the Klenow reaction was supplemented with 10-30 nM of primers FR30RVA (5' GCC GGA GCT CTG CAG ATA TC TTA AAA CTG G 3') and FR30RVB (5' GCC GGA GCT CTG CAG ATA TC TTA AAA CAG C 3') where the final 10bp match the 5' ends of A-type and B-type rhinoviruses, respectively. PCR amplification used high fidelity Taq Gold DNA polymerase (Applied Biosystems, Foster City, CA) with the FR20RV primer (5' GCC GGA GCT CTG CAG ATA TC 3'). PCR amplicons were A-tailed with dATP and 5 units of DNA polymerase (Invitrogen) at 72°C for 30 minutes. A-tailed PCR amplicons were fractionated on a 1% agarose gel and fragments between 500 and 1000 nt were extracted. Amplicons were ligated *en masse* into the Topo TA cloning vector (Invitrogen) and transformed into competent one-shot Topo top 10 bacteria (Invitrogen). Cells were plated on LB/Amp/XGal agar, and individual colonies were picked for sequencing. The inserted fragments were sequenced bidirectionally with the M13 primers from the Topo TA vector. We routinely sequenced a total of 192 fragments or more per library. Sequencing reactions were performed at the Joint Technology Center (Rockville, MD) on an Applied Biosystems 3730 xl sequencing system with Big Dye Terminator chemistry (Applied Biosystems).

Assembly of Viral Genomes

Sequence reads were downloaded, trimmed to remove primer sequence as well as low quality sequence, and assembled with the program ELVIRA, the Executive for Large-scale Viral Assembly (<http://sourceforge.net/projects/elvira>). Additional manual inspections identified ambiguities or potential single nucleotide variants and were interrogated by further RT-PCRs, cloning, and sequencing. To close gaps between assembled contigs, strain-specific primers were utilized. Additional primer design, cDNA synthesis and sequencing were performed to ensure at least 4X sequence coverage along all genomes.

Polyprotein and RNA genome alignments

Founder data for the HRV polyprotein alignment are from superimposition of virion crystal structure hydrogen-bonding maps as described (*S4*). Profile hidden Markov models, derived from the founder data, were progressively augmented with published sequences and those derived from this study. The HMMER program suite (Accelrys, San Diego, CA) reported both high and low road fits, building the alignment possibilities with additions from highest to lowest similarity. Each insertion and deletion (indel) in the output iterations was examined and eased within the confines of its high road and low road fits, to maximize conservation of viral cleavage sites, catalytic sites, determined structure landmarks (P1, 2A, 3B-3D) and sequence similarity. The HMM profile of composite HRVs was used to fit 3 additional HEV-C out-group sequences into the final alignment. Reverse translation of the polyprotein alignment relative to the original RNA sequences formed a core ORF alignment with analogous indels. Preliminary tree-building exercises identified 20 sequences representative of the dominant HRV clades,

including 3 examples from the putative HRV-C species. The complete genome of each select sequence was analyzed separately by mFOLD (S5). The consensus topography (optimal plus 100 suboptimal structures) for each the 5' and 3' regions of each fold was superimposed into common alignments which maximized analogous base-pair superimposition (5' cloverleaf, IRES, 3' stems, etc.) and minimized indels. These foundations were converted into HMM profiles to which the remaining UTR sequences of all HRV and out-groups were fit. In the interest of alignment length, the HEV-C out-group sequences were truncated to remove a 5' fragment without analog in the HRV (100 b ribosome read-through, 5' to ORF AUG). The 5', ORF and 3' aligned segments for all included sequences were joined contiguously into full-length genome alignments. Again, each indel in the file was re-examined for plausibility, consistency and biological conservation, before the composite alignment was finalized.

Phylogenetic Analysis

Phylogenetic analyses on the polyprotein and genome alignments (msf file format) were conducted with MEGA version 4 (S6) and PhyML version 2.4.5 (S7). HRV sequences from this study (including field strains) were augmented as needed with published data to include at least one representative of the 99 described HRV-A and HRV-B serotypes, and all available (full-length) data from the HRV-C. Multiple tree iterations were evaluated for both the protein (single gene and polyprotein) and RNA genome data, with UPGMA (MEGA), maximum parsimony (MEGA), neighbor joining methods (MEGA) with bootstrap tests (2000x), and maximum likelihood (PhyML) with approximate likelihood ratio tests (minimum of SH-like and Chi²-based aLRT). None of these methods showed significant topological differences for major branch points with p-values >2 (i.e., 2% change), especially if the 3rd positions of the ORF

codons were omitted from consideration in the RNA trees. In the absence of full-genome data for all reference strains, HRV relationships have been approximated according to more limited sequence sets derived from the VP1, VP0, IRES and 3D regions. To determine whether “serotype” (i.e., VP0 or VP1 sequences) and/or other more conserved regions of the genomes (i.e., IRES or 3D gene) were useful indicators of the full strain relationships, defined maximum likelihood (ML) topologies optimal for these regions were compared statistically with the ML full-genome tree. Optimal ML topologies for the VP1-only (966 b), 3D-only (1389 b), IRES-only (547 b) and VP0-only (86 b VP4 + 352 b VP2) fragments within the RNA alignment were computed within PhyML, then compared individually against the optimal topology of the full-genome ML tree using PAML and the CONSEL (V0.1i) suite of programs (S37). The tested (null) hypothesis was the expectation that the HRV relationships established using only these limited (albeit commonly used) regions of sequence would be similar to those derived from full genome data. See also the legends to Fig. S6A, B for specific definitions of the implemented topology tests.

5'- and 3'-structural predictions

We considered the accepted notion that thermodynamically derived models for phylogenetically related viruses should exhibit common RNA structural motifs, if such motifs are required for biological activity. Our approach has been previously described in detail (S5). Briefly, rather than form structure predictions on the basis of sequence similarities, folding was undertaken first, and then we searched among the most probable configurations (energy minimization) for regions with consistent structures. As described above for the RNA alignments, full genome sequences for 20-40 hrv, representing different phylogenetic clades,

were evaluated in their entirety by mFOLD (S5) asking for the optimal, and up to 100 closely related (+12 Kcal) suboptimal configurations. Without exception, the consensus fold for each sequence (required >80% of queried connect files) agreed that the 5' and 3' ends of each RNA generally configured independently. That is, few if any segments within these regions made preferred (low energy) long-range contacts with interior portions of the genome. The UTR topologies folded regionally as a series of local, connected motifs. This tendency was confirmed by assessing the P-num values, computed for the whole-genome folds. The pairing number (P-num) is a quantitative measure of the propensity of any given base to become involved with the same or alternative pairing partners in a collection of suboptimal folds. We have shown for other viral sequences that low P-num bases and their correlate partners usually dominate the most important helices and stems supporting biologically significant motifs, especially within the lowest energy configurations. These bases and their partners were therefore used to identify and align true homologues (functional analogues) at the primary sequence level, even in regions with less-than obvious conservation. Superimposition of the 5' and 3' low P-num motifs from the genome folds formed the core consensus profiles for the RNA alignments in these regions (as described above) and also identified multiple specific conserved motifs throughout the genomes. Once these commonalities were identified for the 5' cloverleaf, IRES, 5' ORF initiation stem, cre element, and 3'UTR, all other sequences in the alignment were re-folded in these regions (mFold, with 50 suboptimals), to confirm that they too had similar, conserved, low energy, low P-num motifs.

Recombination analysis

The recombination predictions of the genomic sequences, aligned as described above, were conducted with a suite of programs within the RDP3 package (*S8*). The individual programs RDP (*S9*), Bootscan (*S10*), Maximum X2 (*S11*), Chimaera (*S8*), SiScan (*S12*) and 3Seq (*S13*), were implemented for the analysis. Since no single program provides optimal performance under all conditions, any event supported by evidence from two or more analyses with P-values <0.00001 was considered a result consistent with recombination. Potential recombination events were also assessed by phylogenetic analysis, breakpoint polishing and alignment consistency checks. For each individual program, default settings were used except as specified: RDP, internal reference only, window size for recombination, 100 bp with step size 10 bp; GENECONV, G-scale was set as 3; BootScan, number of bootstrap replicates, 200, window size, 100 bp, step size, 10 bp, model options: Jukes and Cantor, 1969; Maxichi, variable window size was used, strip gap was selected; SiScan, window size, 100 bp with step size 10 bp, P-value permutation number was set as 1000. In additional analyses, to confirm the recombinations that were found, genomic sequences were aligned using progressive alignment methods (*S14*). Briefly, the progressive alignments were performed with non-coding (*S15*) and coding regions (*S14*, *S16*) using the referenced programs and the alignments were concatenated by a custom script. The subsequent final alignment was utilized for recombination predictions using the RDP3 package as described above.

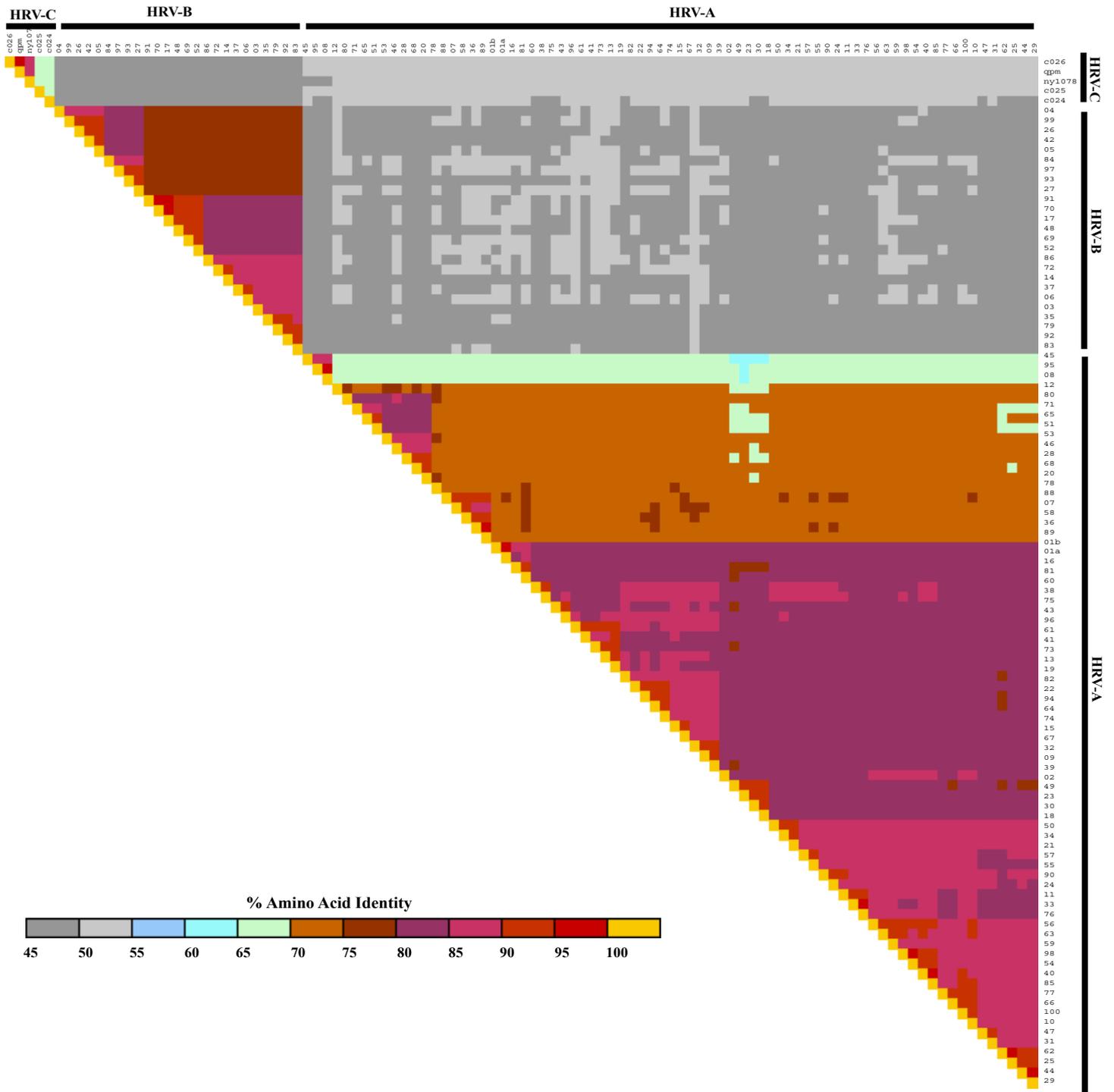


Fig. S1. Whole genome amino acid sequence identity comparison. Amino acid sequences were deduced from the coding region of the sequenced Rhinovirus genome sequences. The aligned amino acid sequences were compared in a pairwise fashion to calculate the identities. The identity matrices of the compared genomes visualized in square arrays of sequence identity values were clustered by the phylogenetic ordering.

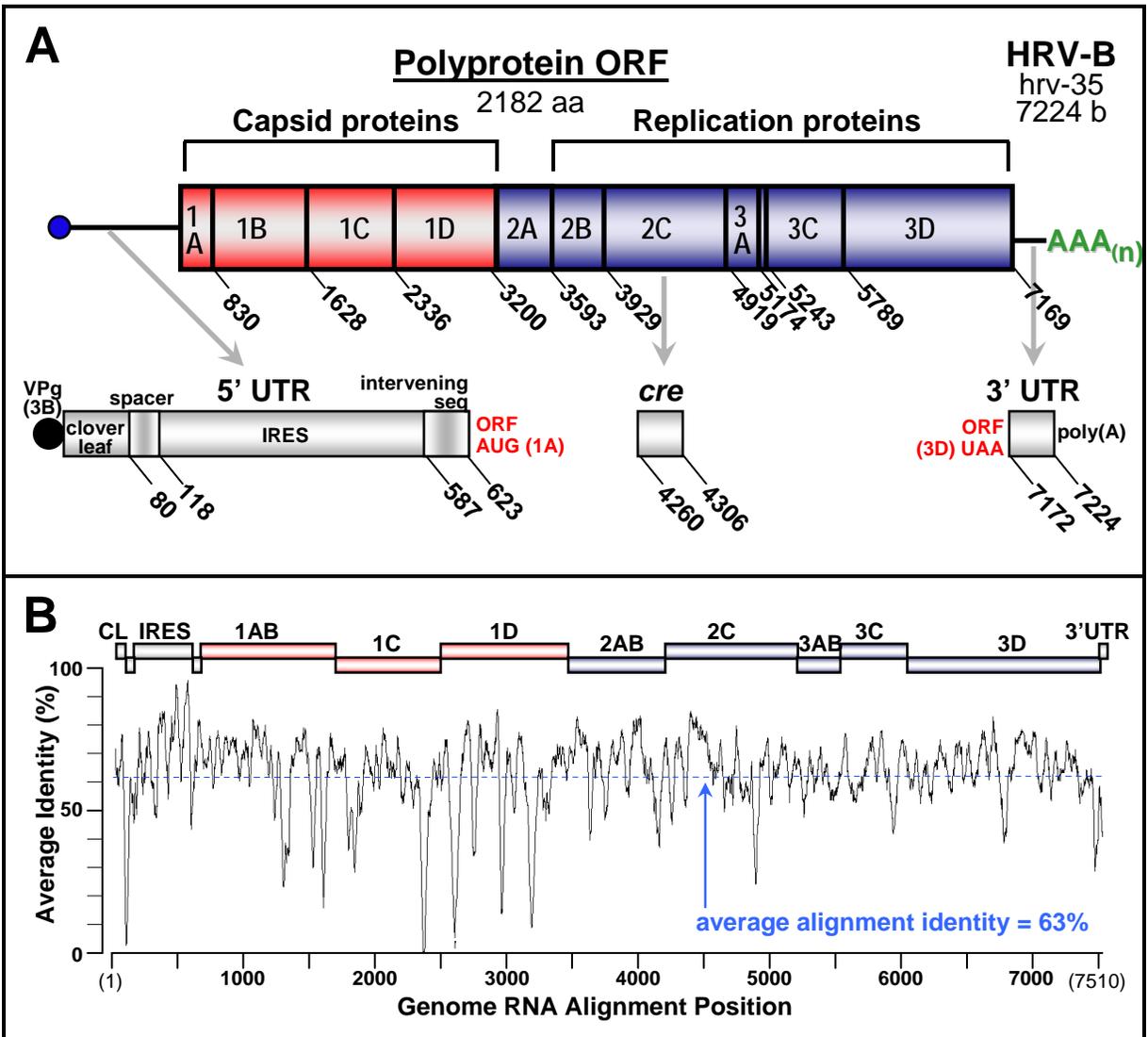


Figure S2: A: An HRV genome map. The 5' untranslated region (5' UTR) is linked to VPg(3B), and encodes several important RNA structures which function during RNA synthesis and genome translation. The single open-reading frame (ORF) encodes a polyprotein, which is cleaved in a series of co-translational and post-translational reactions to provide all mature viral proteins required to establish and perpetuate an infection. The capsid proteins (red) and proteins involved in replication functions (blue) are common to all HRV, and all sequences share analogous cleavage sites delineating these locations in the polyprotein. The illustrated genome is hrv-35; the base numbering system in this panel is for that sequence. **B:** The HRV sequences in the RNA genome alignment (Table S2) were queried pairwise at each position in the alignment. All sequences were given equivalent weight. The arithmetic average of the scores (0 or 1 for each pair) was reported for each position, re-averaged over a sliding window of 30 adjacent residues (± 15), then plotted relative to the alignment as a whole. The strongest sequence conservation is in the 5' IRES region. The lowest conservation is in the 5' spacer region, upstream of the IRES, and also in the capsid coding regions, where the 1B, 1C and 1D troughs correspond to the respective, mapped immunogenic surface loops. The averaged identity across all HRVs, considering all alignment positions, is indicated with the blue, horizontal line.

Figure S3 B: Region Between IRES and ORF Start Codon (AUG)

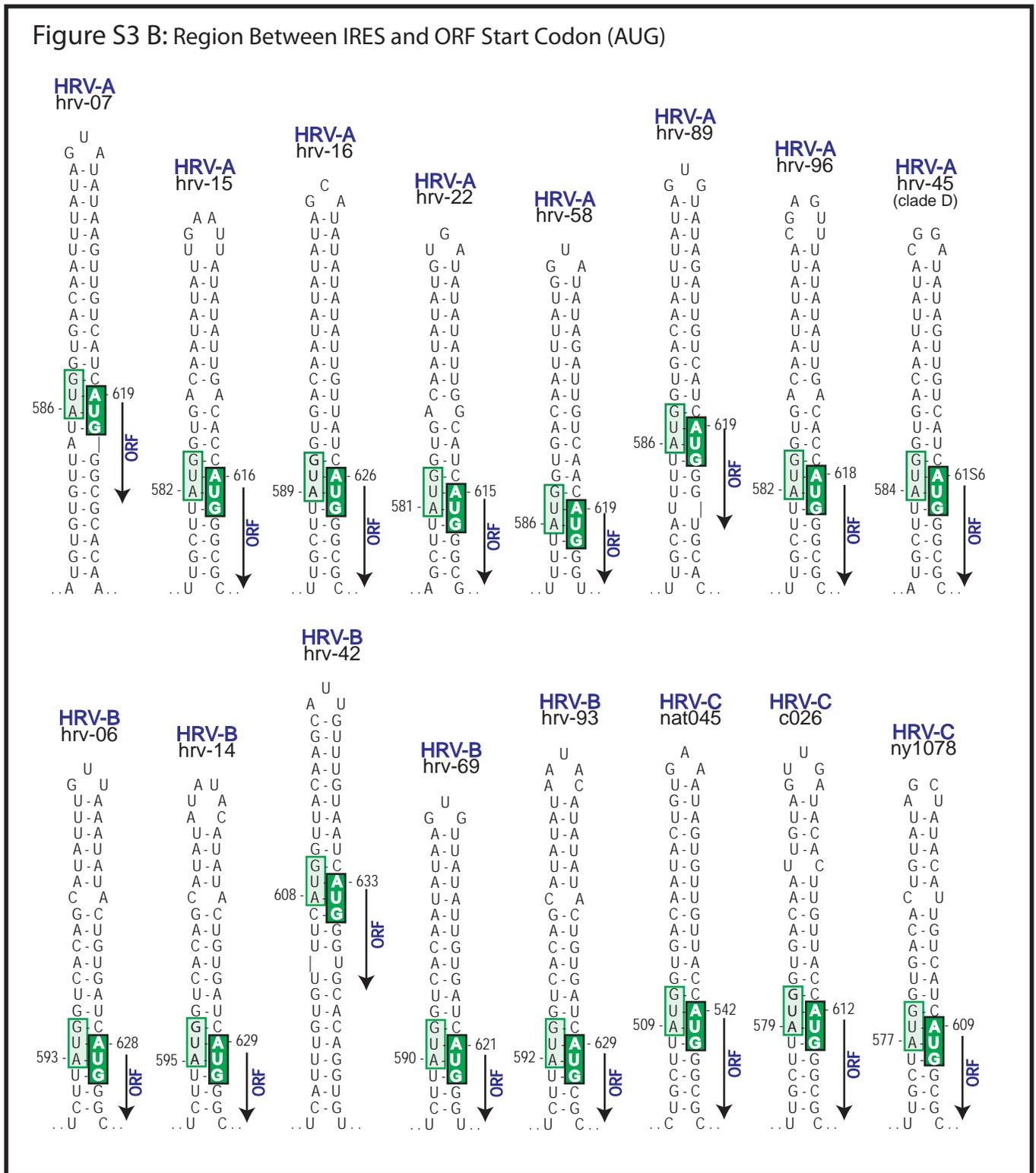


Figure S3 B: The RNA structures near the AUG which begins the polyprotein open-reading frame (ORF) are depicted for representatives of each HRV species. Every sequenced HRV folds this region into a similar configuration when the genome is probed by minimum-free energy calculations. The ORF AUG (white letters, green box) is always paired with a conserved upstream AUG (light green box), which marks the 3' boundary of the IRES. Other species in the *Enterovirus* genus do not make similar pairings. Rather, other enteroviruses are believed to launch their ribosomes for scanning, from the upstream AUG. This HRV configuration makes it likely that an IRES-bound ribosome could readily “switch” to the ORF AUG, without scanning the intervening spacer. Between these AUGs, neither the specific sequence or it's length is well conserved among HRVs. Rather, all form these related structures, preserving the specific pairings only in the region of the AUGs.

Figure S3 C: Cis-acting Replication Elements (cre)

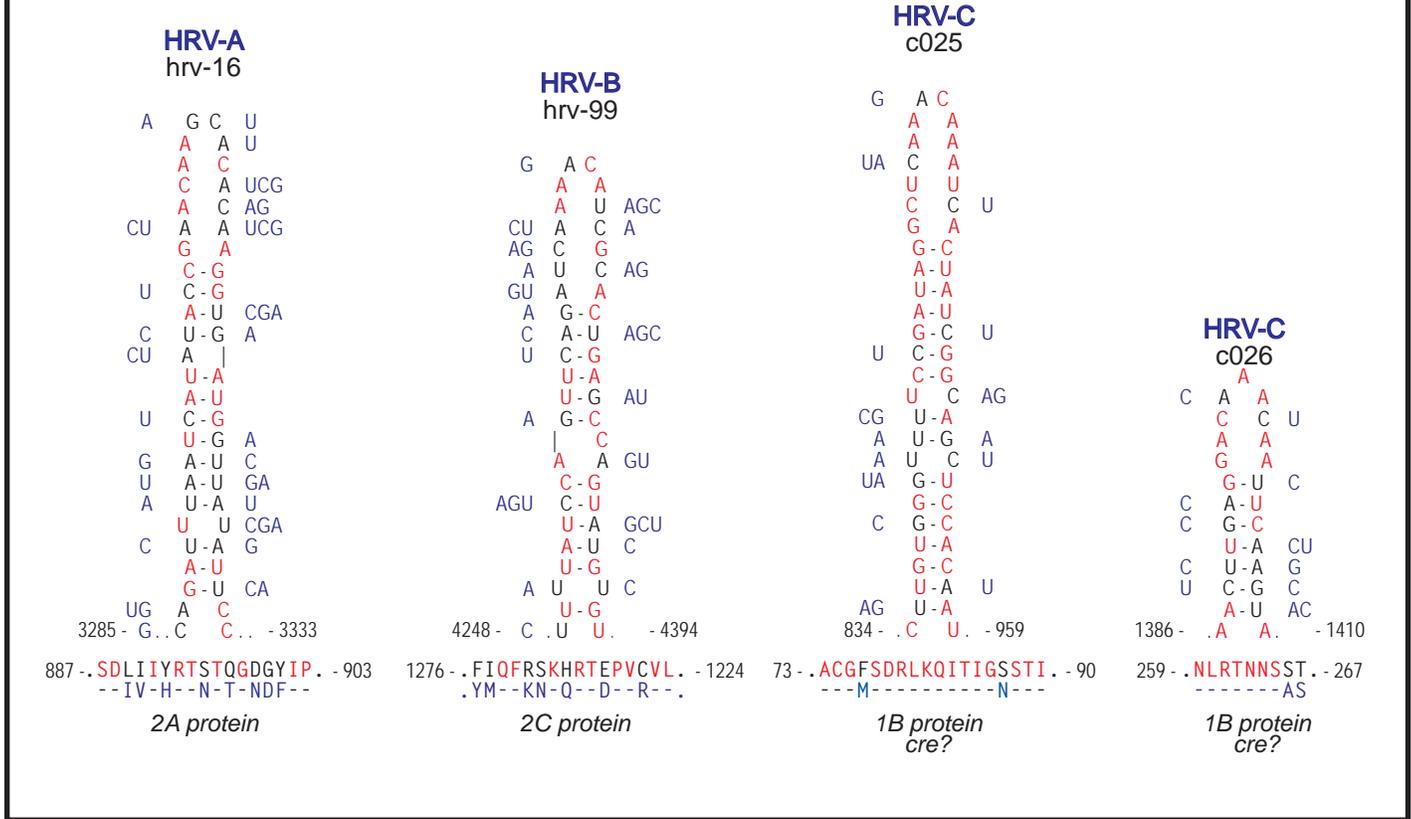


Figure S3 C: Picornavirus RNA synthesis is protein primed. The viral RNA polymerase 3Dpol, uridylylates protein 3B (VPg) to form VPg-pUpU, then uses this protein:RNA as primer for the initiation of positive and negative strand RNA synthesis. The uridylylation reaction is templated by a special RNA structure called the *cre* (cis-acting replication element) whose location varies for every known species of picornavirus. For the HRV-A, the *cre* has been mapped within the genome region encoding the 2A protein. For the HRV-B, the *cre* is within the 2C region. The HRV-C *cre* has not been mapped genetically, but common to all *cre*s, is a distinct stem motif, displaying CAAACAA or a closely related sequence. Within the sequenced genomes of the HRV-C there are 2 locations in the 1B gene which fit this description. These, and the HRV-A and HRV-B *cre*s are depicted. The structure is for the indicated sequence. Those bases colored red are conserved among all other sequences in that species. All observed (aligned) base changes are indicated in blue. Below each element is the protein sequence encoded by that region. Red residues are conserved among all sequences in that species. Blue residues are observed in at least 1 other sequence. Note: the HRV-A structure omits sequence variation contributed by the clade-D viruses. It is not yet clear whether the *cre* for these viruses occupies the same genome location.

Figure S3 D: 3' Untranslated Regions (3' UTR)

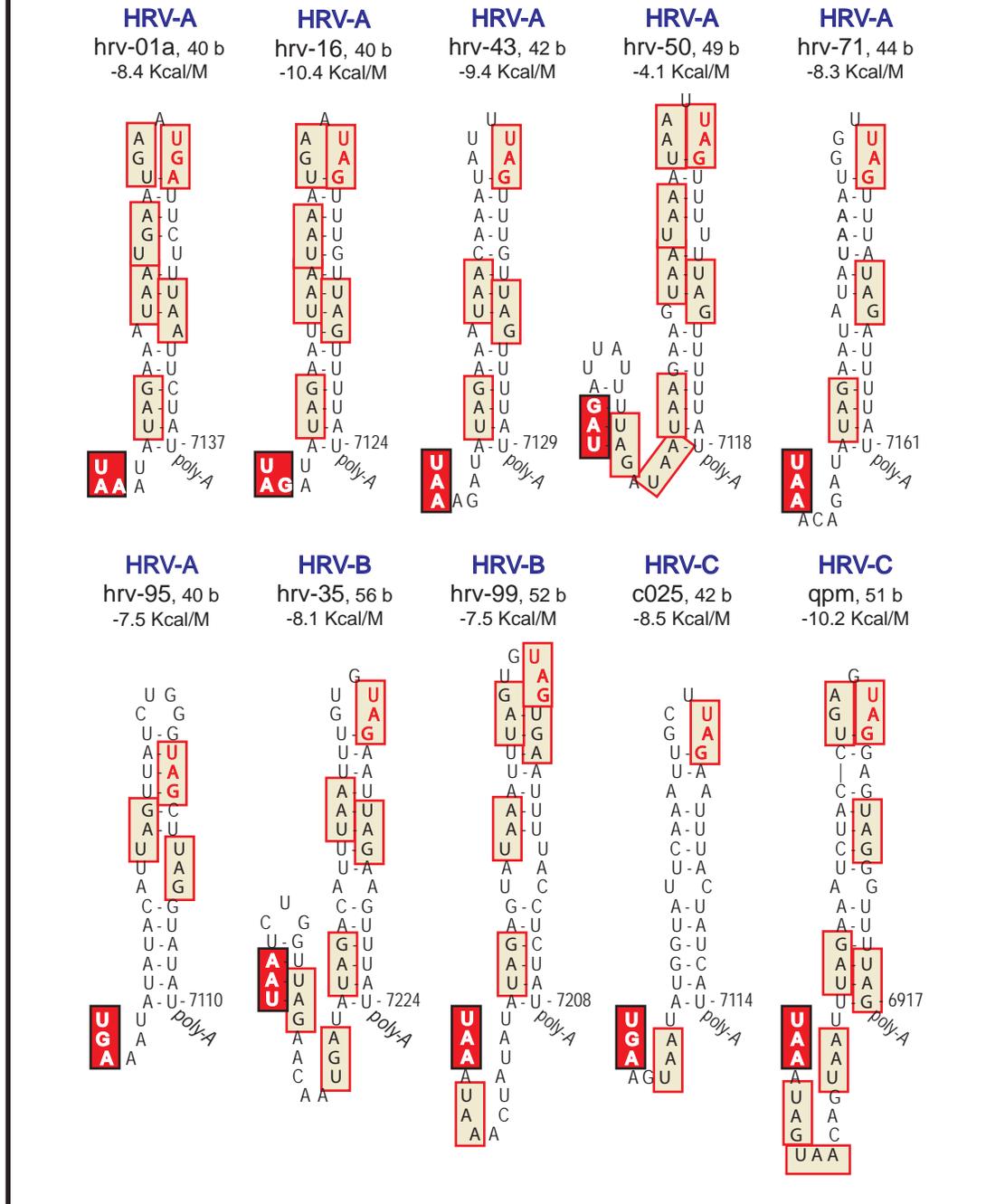


Figure S3 D: Representative HRV genomes were analyzed for their minimum free-energy configurations. Within those datasets, the regions representing the 3' untranslated regions (3'UTR) fold into motifs independent of the rest of the genome. Examples of these regions, typical of each HRV species are illustrated (ORF terminator, red box, white letters). In every case a single stem motif dominates the region, ending exactly at the 3' poly(A) tail, which may contribute 1-2 additional base-pairs to the bottom of the stem. Sequence conservation is poor throughout this region, but every virus, with the exception of the clade-D isolates, displays a termination codon within the terminal loop of the 3' stem (tan box, red letters). Although multiple other termination codons are scattered frequently throughout the region (tan boxes), this particular triplet is the only one which is conserved among most of the isolates, regardless of species. The clade-D viruses (e.g. hrv-95) have shorter 3' stems and display an UAG triplet near, but not within the terminal loop (Fig S4).

Figure S4 : Some Unique Features of Clade D Viruses (hrv-08, hrv-45, hrv-95)

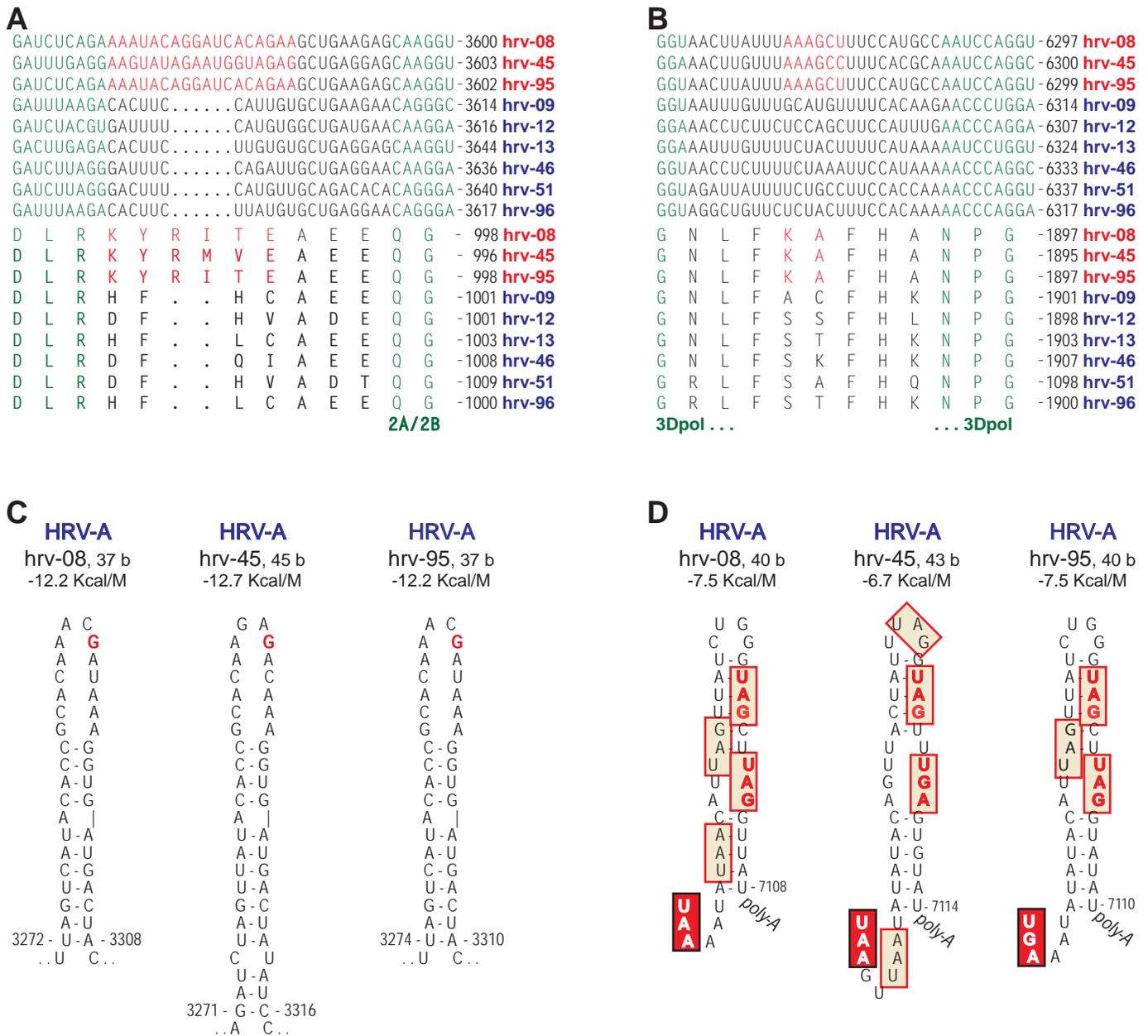


Figure S4: Among the HRV-A, 3 serotypes (hrv-08, hrv-45 and hrv-95) form a distinct clade, called “clade-D”. Of these, hrv-08 and hrv-95 share >98% nucleotide identity, perhaps indicating a misidentification within the original reference collection. Despite the paucity of unique isolates in this clade, the identified members share characteristics in certain key genome locations which clearly distinguish them from all other HRV-A sequences (or HRV-B and HRV-C). A few examples are shown in this figure. **A:** Near the C-terminus of the 2A protein, in a region known to contribute to substrate specificity, all clade-D viruses have an indel unique among all other HRV. **B:** Throughout the alignments, there are numerous locations, like this one in the 3Dpol gene, where clade-D sequences vary independently with highly variant, non-synonymous substitutions. **C:** The *cre* elements of all HRV-A are located 2A gene (see Fig S3C). The loop sequences, especially in the essential terminal loop invariably display ACAACAA motifs providing the template for 3D-dependent VPg uridylylation. The “G” residue (red) displayed within the clade-D loop has not been observed in any other functional *cre* element. This may indicate that for this clade, the (authentic) *cre* lies elsewhere in the genome, or that these particular *cre*s have a different templated activity. **D:** The 3' UTR sequences of all HRVs, with the exception of the clade-D viruses, display a conserved termination codon in the terminal 5-6 b stem loop (Fig S3D), and moreover, there are very few examples of extended interior loops within the terminal stem. The clade-D viruses pair the conserved termination codon, near the top of the stem, and display a second, unpaired codon, within an interior loop (tan boxes, red letters). They are also the only sequences to have just 4 b in the terminal loop. This entire 3' configuration is unique to this clade, and not shared by any other HRV-A, HRV-B or HRV-C.

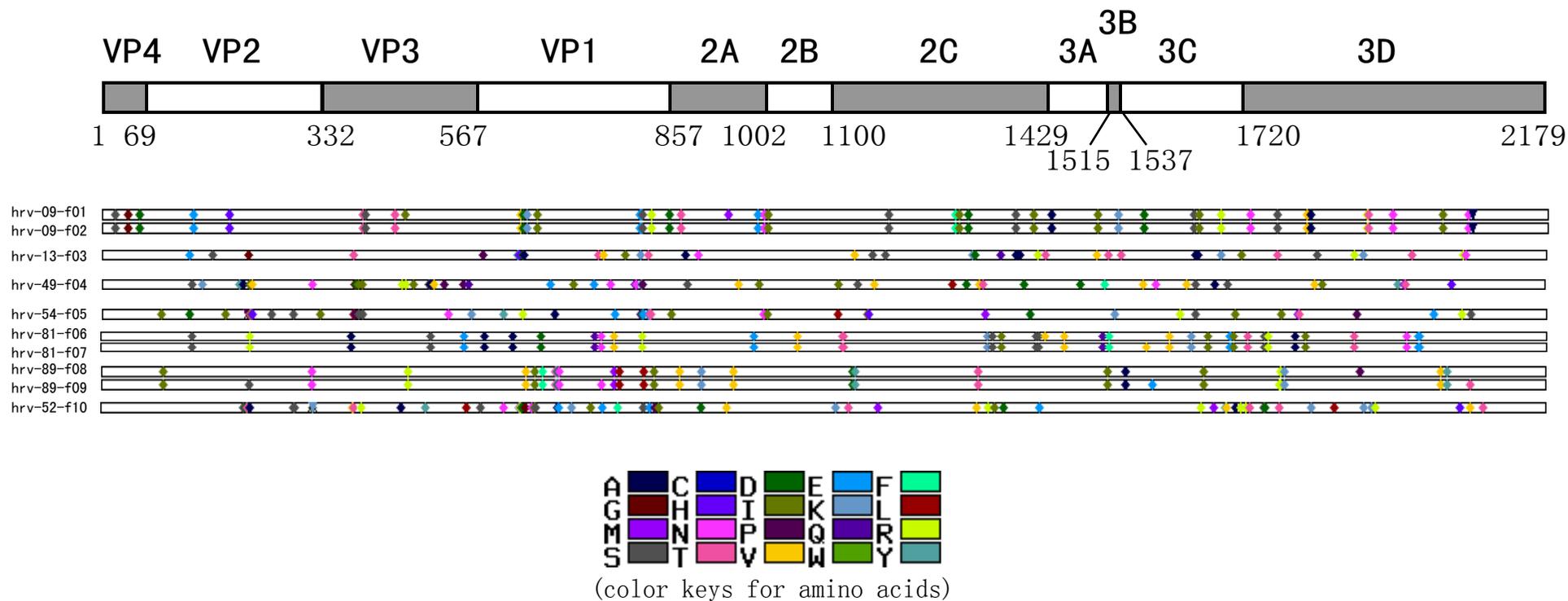


Figure S5. The distribution of amino acid variations in field samples. Each field sample was aligned with the reference sample of the same serotype. Amino acid differences in the field sample are designated by color changes as identified in the key at the bottom of the figure. The upper panel portrays a diagram of hrv-14 genome structure for orientation.

Figure S6b. The serotype of an HRV is determined by sequences within the immunogenic surface loops of capsid proteins VP0 (VP4+VP2) and VP1. Frequently, these segments and/or the IRES and/or the 3Dpol regions are sequenced separately and used to place new isolates onto HRV species trees. Established statistical methods within the program set CONSEL (V0.1i) tested the null hypothesis that the topology (Fig S6c) of such “serotype only” trees, or an IRES-only tree, or a 3D-only tree were consistent with the Fig S6a full genome ML tree. The program output includes: the observed log-likelihood difference (obs); the p-value of the approximately unbiased test calculated from the multiscale bootstrap (au); the bootstrap probability calculated from multiscale bootstrap (np); the bootstrap probability calculated in the usual manner (bp); the Kishino-Hasegawa test (kh); the Shimodaira-Hasegawa test (sh); the weighted Kishino-Hasegawa test (wkh); and the weighted Shimodaira-Hasegawa test (wsh). The last 6 tests had identical output values for all considered trees, and are summarized together. The data, individually and collectively strongly reject the null hypotheses for all alternative tested topologies.

Topology tests of ML complete (full) genome tree					
Rank	item	obs	au	np	bp, pp, kh, sh, whk, wsh
Full	1	-1900.6	1.000	1.000	1.000
VP1	4	1900.6	8e-76	8e-23	0
3D	2	3320.5	2e-06	7e-06	0
VP0	3	6528.1	7e-70	2e-19	0
IRES	5	15947.7	1e-100	7e-26	0

Maximum Likelihood Tree for complete genomes of all known HRV serotypes

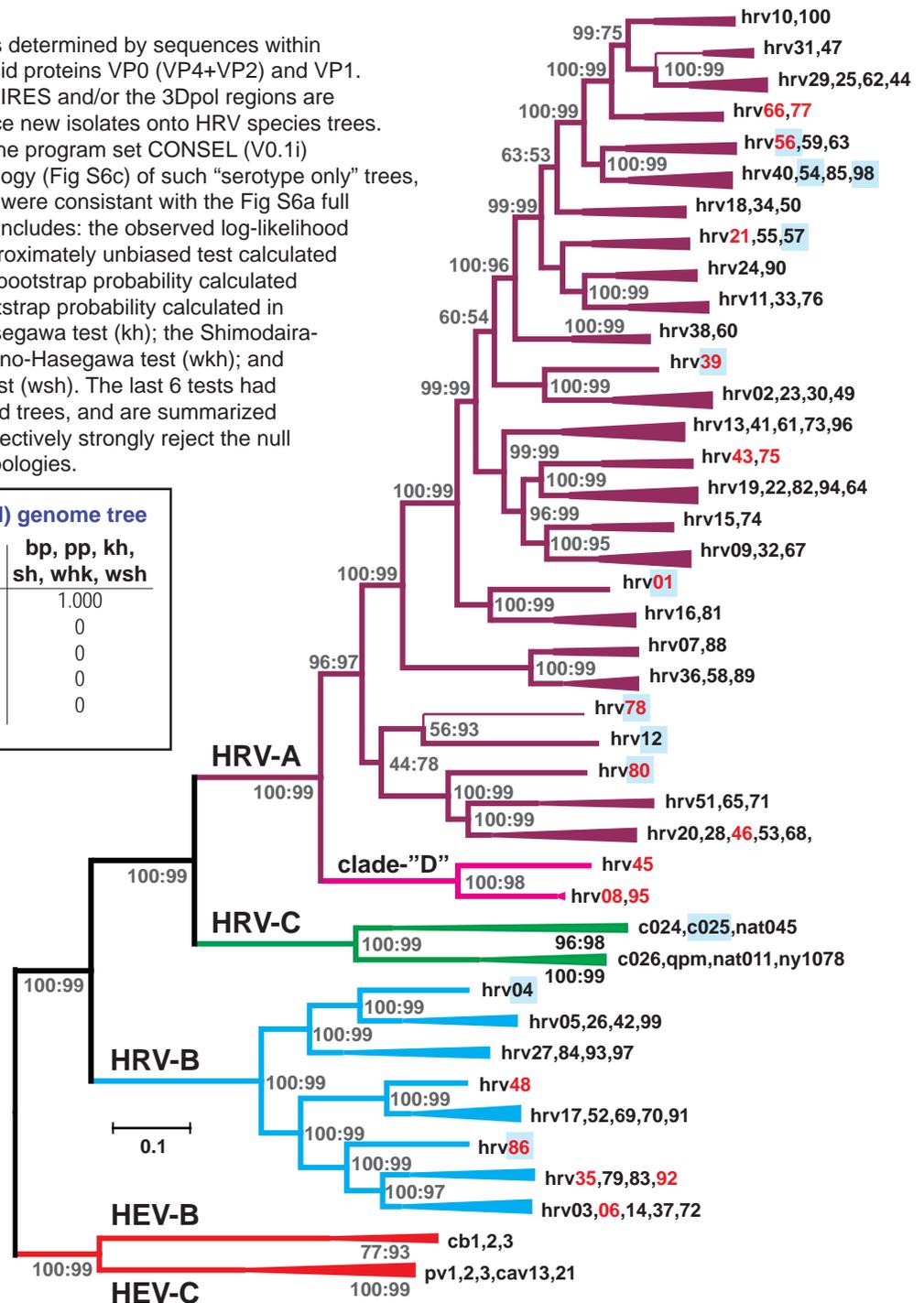


Figure S6a. The RNA genome alignment (TableS2) was augmented with additional outgroup sequences from the HEV-C species (pv2l:M12197, pv3l:K01392) and the HEV-B species (cvb1:M16560, cvb2:AF081485, cb3:M33854), with profile:profile fits in Clustal. A maximum likelihood (ML) tree was calculated using PHYML version 2.4.8 for the full genome dataset (“full”). TN93 nucleotide substitution model gave the best likelihood (loglk -462223) over HKY, GTR, JP69, K2P and F81 models. The Ts/tv ratio and invariable sites were estimated. Approximate likelihood ratio tests (aLRT, minimum of SH-like and Chi2-based values) and bootstrap tests (BS, 200x) were applied. The test values label the tree nodes (gray text, %BS:aLRT). Tree length (sum of branch lengths) is 19.39. When compared to the Fig 2 neighbor-joining (NJ) tree, terminal clades were collapsed if they were identical to the NJ tree. Of all considered linkages, only hrv78 and the hrv(31+47) clade showed minor rearrangements by ML analyses relative to NJ (thin line branches). Parallel ML trees were also calculated using only the VP1, VP0, IRES or 3D-encoding RNA sequences (see Methods). Taxa names with different clade affiliations (>0.2 p-value) on the VP0-only tree are shown in red. Taxa names with different clade affiliations on VP1-only tree are highlighted in blue.

Figure S6C. The five specific maximum likelihood (ML) tree topologies tested in panel S6b are indicated.

Unconstrained, full genome, ML topology

((((((((((((((((hrv74*, hrv15*) (hrv67 (hrv32 ((hrv09-f02, hrv09-f01) hrv09)))) (((((hrv94, hrv64) hrv22) hrv82) hrv19) (hrv75*, hrv43))) ((hrv96, hrv61) ((hrv73* (hrv13-f03, hrv13) hrv41*)) ((((((hrv90 (hrv24*, hrv24) ((hrv76, hrv33) hrv11*)) ((hrv21, hrv55*) hrv57) (((((hrv98 (hrv54-f05, hrv54) (hrv85, hrv40) ((hrv63, hrv59*) hrv56) (((hrv100, hrv10) (((hrv62, hrv25) (hrv44*, hrv29)) (hrv47, hrv31))) (hrv77, hrv66))) ((hrv50, hrv34) hrv18))) (hrv60, hrv38) (((hrv49-f04, hrv49*) hrv02*) (hrv30, hrv23*)) hrv39*)) (((hrv81-f07, hrv81-f06) hrv81) hrv16*) hrv01) (((((hrv89-f09, hrv89-f08) hrv89) hrv36*) hrv58) (hrv88*, hrv07))) ((hrv80 ((hrv71 (hrv65, hrv51) (((hrv68, hrv20) hrv28*) (hrv53*, hrv46*))) (hrv78, hrv12*))) ((hrv95, hrv08) hrv45) (((c025* (nat045*, c024*) (((c026*, qpm*) nat001*) ny1078))) (((((hrv99 (hrv42, hrv05) hrv26) hrv04*) ((hrv97 (hrv93*, hrv27) hrv84) ((((((hrv92, hrv83) hrv79) hrv35) (hrv03* ((hrv37*, hrv06*) (hrv72, hrv14*)))) hrv86) (((hrv52-f10, hrv52) (hrv69 (hrv91 (hrv70*, hrv17*))) hrv48*))) (cva21* (cva13* ((pv2l*, pv3l*) pv1m*))) cvb2*) cvb3*, cvb1*)

VP0-only constrained ML tree topology

((((((((((((((((((((hrv39* ((hrv98 (hrv54-f05, hrv54) hrv01)) (hrv85, hrv40)) (hrv63, hrv59*)) (((((((((hrv32 ((hrv09-f02, hrv09-f01) hrv09) hrv67) hrv96) ((hrv82, hrv61) ((hrv73* (hrv41* (hrv13-f03, hrv13)) hrv55*)) hrv57) (((hrv49-f04, hrv49*) hrv02*) (hrv30, hrv23*)) hrv21)) (hrv60, hrv38) ((hrv73* (hrv76, hrv11*) (hrv90 (hrv24*, hrv24) (((hrv22 (hrv64, hrv94) (hrv74*, hrv15*) hrv19))) ((hrv34, hrv18) hrv50))) (((hrv81-f07, hrv81-f06) hrv81) hrv16*) hrv56) (hrv75*, hrv43) (((hrv100, hrv10) (hrv77, hrv66) (((hrv62, hrv25) (hrv44*, hrv29)) (hrv47, hrv31)))) (((((hrv89-f09, hrv89-f08) hrv89) hrv58) hrv36*) (hrv88*, hrv07))) (((hrv95, hrv08) hrv45) hrv78) hrv12*) (hrv53*, hrv28*) (hrv68, hrv20) (hrv65 (hrv71, hrv51))) (hrv80, hrv46*)) (((c025*, c024*) nat045*) (((c026*, qpm*) nat001*) ny1078))) (((((((hrv91, hrv70*) hrv17*) hrv69) hrv48*) (hrv52-f10, hrv52) ((hrv14* ((hrv37*, hrv06*) hrv03*) (hrv86, hrv72))) ((hrv79, hrv83) hrv92) hrv35) (((hrv42, hrv05) hrv99) hrv04*) hrv26) ((hrv97 (hrv93*, hrv27) hrv84)))) cva21* (cva13* ((pv3l*, pv2l*) pv1m*)) cvb2*) cvb3*, cvb1*)

VP1-only constrained ML tree topology

((((((((((((((((((((hrv90 (hrv24*, hrv24) ((hrv76, hrv11*) hrv33)) (hrv55*, hrv21)) (hrv57 ((hrv50, hrv34) hrv18))) (((((((((hrv62, hrv25) (hrv44*, hrv29)) (hrv47, hrv31)) (hrv77, hrv66) (hrv100, hrv10) (((hrv98 (hrv54-f05, hrv54) ((hrv85, hrv40) hrv56) ((hrv63, hrv59*) hrv39*)) hrv01) (((hrv81-f07, hrv81-f06) hrv81) hrv16*)) (((hrv96, hrv61) (hrv73* (hrv41* (hrv13-f03, hrv13))) (((((hrv94, hrv64) hrv22) hrv82) hrv19) (((hrv74*, hrv15*) (hrv67 (hrv32 ((hrv09-f02, hrv09-f01) hrv09)))) (hrv60, hrv38))) (hrv75*, hrv43) (((hrv49-f04, hrv49*) hrv02*) (hrv30, hrv23*)) (((((hrv89-f09, hrv89-f08) hrv89) hrv36*) hrv58) (hrv88*, hrv07) (((hrv80, hrv46*) ((hrv71 (hrv65, hrv51) (((hrv68, hrv20) (hrv53*, hrv28*))) (hrv78, hrv12*))) ((hrv95, hrv08) hrv45) ((nat045*, c024*) (((c026*, qpm*) nat001*) ny1078) c025*)) (((((((hrv91 (hrv70*, hrv17*) hrv69) hrv48*) (hrv52-f10, hrv52) hrv48*) (((hrv37*, hrv06*) (hrv72, hrv14*) hrv03*) (hrv35 ((hrv92, hrv83) hrv79)) hrv86) (((hrv99, hrv04*) (hrv42, hrv05) hrv26) ((hrv97 (hrv93*, hrv27) hrv84)))) (cvb3*, cvb1*) cvb2*) cva21* (cva13* pv1m*) pv2l*, pv3l*)

IRES-only constrained ML tree topology

((((((((((((((((((((hrv47, hrv31) (hrv77, hrv66) ((hrv100, hrv10) hrv56) (((hrv80, hrv55*) hrv21) (hrv57 ((hrv85, hrv40) ((hrv62, hrv25) (hrv44*, hrv29) (((hrv98, hrv54) hrv54-f05) hrv38) (((hrv63, hrv59*) hrv50) (((hrv68, hrv20) (hrv90 (hrv24*, hrv24) hrv11*)) (hrv33, hrv76)))) ((hrv95, hrv08) ((hrv46*, hrv34) hrv18))) (((hrv23*, hrv02*) (hrv49-f04, hrv49*) hrv30) (((hrv67 (hrv32 ((hrv09-f02, hrv09-f01) hrv09))) hrv60) (hrv74*, hrv15*))) (((hrv81-f07, hrv81-f06) hrv81) hrv16*) (hrv53*, hrv28*) ((hrv19, hrv82) ((hrv94, hrv64) hrv22) (((hrv96, hrv61) (hrv73* (hrv41* (hrv13-f03, hrv13))) ((hrv75*, hrv01) hrv43)))) hrv39*) (((c026*, qpm*) nat001*) ny1078) hrv45) (hrv78, hrv12*))) (((((hrv89-f09, hrv89-f08) hrv89) hrv36*) hrv58) (hrv88*, hrv07) (((c025*, nat045*) c024*) (hrv71 (hrv65, hrv51))) (((((((hrv99 (hrv42, hrv05) hrv26) hrv84) hrv86) (hrv72 ((hrv03* (hrv06*, hrv37*) hrv14*) ((hrv92 (hrv83, hrv35) hrv79) ((hrv97, hrv04*) (hrv93*, hrv27)))) (((hrv52-f10, hrv52) hrv48*) hrv69) (hrv70* (hrv91, hrv17*))) (cva21* ((pv2l*, pv3l*) pv1m*) cva13*)) cvb2*) cvb3*, cvb1*)

3D-only constrained ML tree topology

((((((((((((((((((((hrv62, hrv25) (hrv44*, hrv29), hrv100) (hrv10 (hrv63, hrv56))) ((hrv57, hrv21), hrv55*)) ((hrv50, hrv34) (((hrv98 (hrv54-f05, hrv54) (hrv47, hrv31)) (hrv60, hrv38) (((hrv85, hrv18), hrv40), hrv77)) (hrv90 (hrv66 (hrv24*, hrv24) (((hrv76, hrv33), hrv11*), hrv59*) (((hrv49-f04, hrv49*), hrv02*) (hrv30, hrv23*), hrv39*)) ((hrv75* (((hrv94, hrv64), hrv22), hrv82) (hrv43, hrv19)) (hrv67 (hrv32 ((hrv09-f02, hrv09-f01), hrv09))) (hrv74*, hrv15*))) ((hrv96, hrv61) ((hrv73* (hrv13-f03, hrv13), hrv41*)) (((hrv81-f07, hrv81-f06), hrv81) hrv16*) hrv01) ((hrv88*, hrv07) ((hrv36*, hrv58) (hrv89-f09, hrv89-f08), hrv89))) (hrv78, hrv12*)) (hrv80 (((hrv53*, hrv46*) ((hrv68, hrv28*), hrv20) (hrv71 (hrv65, hrv51)))) (((((hrv97 (hrv93*, hrv27), hrv84) (((hrv42, hrv05), hrv26), hrv99), hrv04*)) (((((((hrv92, hrv83), hrv79), hrv35) (hrv03* (hrv37*, hrv06*) (hrv72, hrv14*))), hrv86) (hrv48* ((hrv69 (hrv52-f10, hrv52) (hrv17* (hrv70*, hrv91)))))) (((cva21*, pv2l*) (cva13*, pv3l*), pv1m*)) (cvb3* (cvb2*, cvb1*))) (((c025* ((c026*, qpm*), nat001*), ny1078), c024*, nat045*)), hrv45), hrv95, hrv08)

Table S1:

Guide to Sequences and Strains

msf file order (1)	species (2)	sero-type (3)	strain (4)	name in alignment (5)	name on tree (6)	accession # or seq origin (7)	receptor (8)	drug group (9)	ORF start (10)	ORF end (11)	protein length (12)	seq length (13)	3'NTR (14)	seq desc (15)	published sequence reference
-	-	-	-	hrv-ann (18)	-	-	-	-	-	-	-	-	-	-	(S15-S19)
1	HEV-C	a13	Flores	af499637-cva13	cva-13*	AF499637	ICAM1	1	746	7387	2214	7458	71	cg	(S20)
2	HEV-C	a21	Kuykendall	af546702-cva21	cva-21*	AF546702	ICAM1	1	714	7334	2207	7406	72	cg	(S20)
3	HEV-C	pv1	Mahoney	v01149-pv1m	pv-1m*	V01149	PVR	1	743	7369	2209	7440	71	cg	(S21)
4	HRV-A	1	b, B632	d00239-01	not inc	D00239	LDLR	2	623	7093	2157	7133	40	cg	(S22)
5	HRV-A	1	a, 20601-Ohio	pico-01	not inc	AF343633, AY458604, M12166, M12169, AY436674	LDLR	2	627	7097	2157	7137	40	cg	Pico-DB (7)
6	HRV-A	1	A	hrv-01	hrv-01	FJ445111	LDLR	2	627	7097	2157	7137	40	cg	this study
7	HRV-A	2	HGP	x02316-02	hrv-02*	X02316	LDLR	2	611	7060	2150	7102	42	cg	(S23)
8	HRV-A	7	ATCC	hrv-07	hrv-07	FJ445176	ICAM1	2	619	7104	2162	7146	42	cg	this study
9	HRV-A	7	68-CV11	dq473503-07	not inc	DQ473503	ICAM1	2	619	7104	2162	7146	42	cg	(S24)
10	HRV-A	8	ATCC	hrv-08	hrv-08	FJ445113	ICAM1	2	607	7068	2154	7108	40	cg	this study
11	HRV-A	9	211-CV13	pico-09	not inc	AF343605, AY450525	ICAM1	2	611	7081	2157	7128	47	cg	Pico-DB (7)
12	HRV-A	9	ATCC	hrv-09	hrv-09	FJ445177	ICAM1	2	612	7085	2158	7132	47	cg	this study
13	HRV-A	9	fs ship#2, isolate#A	hrv-09-f01	hrv-09-f01	FJ445114	ICAM1	2	614	7087	2158	7134	47	cg	this study
14	HRV-A	9	fs ship#2, isolate#B	hrv-09-f02	hrv-09-f02	FJ445115	ICAM1	2	613	7086	2158	7133	47	cg	this study
15	HRV-A	10	204-CV14	dq473498-10	not inc	DQ473498	ICAM1	2	609	7088	2160	7137	49	cg	(S24)
16	HRV-A	10	ATCC	hrv-10	hrv-10	FJ445178	ICAM1	2	609	7088	2160	7137	49	cg	this study
17	HRV-A	11	1-CV15	ef173414-11	hrv-11*	EF173414	ICAM1	2	619	7077	2153	7125	48	cg	(S25)
18	HRV-A	12	181-CV16	ef173415-12	hrv-12*	EF173415	ICAM1	2	614	7078	2155	7124	46	cg	(S25)
19	HRV-A	13	ATCC	hrv-13	hrv-13	FJ445116	ICAM1	1	616	7095	2160	7140	45	cg	this study
20	HRV-A	13	fs ship#1	hrv-13-f03	hrv-13-f03	FJ445117	ICAM1	1	617	7096	2160	7143	47	cg	this study
21	HRV-A	15	1734- South Carolina/60	dq473493-15	hrv-15*	DQ473493	ICAM1	2	616	7092	2159	7134	42	cg	(S24)
22	HRV-A	16	11757- Washington DC/60	l24917-16	hrv-16*	L24917	ICAM1	2	626	7084	2153	7124	40	cg	(S26)

23	HRV-A	18	ATCC	hrv-18	hrv-18	FJ445118	ICAM1	2	612	7070	2153	7119	49	cg	this study
24	HRV-A	19	ATCC	hrv-19	hrv-19	FJ445119	ICAM1	2	615	7088	2158	7135	47	cg	this study
25	HRV-A	20	ATCC	hrv-20	hrv-20	FJ445120	ICAM1	2	622	7119	2166	7163	44	cg	this study
26	HRV-A	21	ATCC	hrv-21	hrv-21	FJ445121	ICAM1	2	615	7085	2157	7134	49	cg	this study
27	HRV-A	22	ATCC	hrv-22	hrv-22	FJ445122	ICAM1	2	615	7082	2156	7129	47	cg	this study
28	HRV-A	23	5124-CV24	dq473497-23	hrv-23*	DQ473497	LDLR	2	535	6981	2149	7025	44	cg-5	(S24)
29	HRV-A	24	ATCC	hrv-24	hrv-24	FJ445190	ICAM1	2	622	7083	2154	7132	49	cg	this study
30	HRV-A	24	5146-CV25	ef173416-24	not inc	EF173416	ICAM1	2	622	7083	2154	7132	49	cg	(S25)
31	HRV-A	25	ATCC	hrv-25	hrv-25	FJ445123	LDLR	2	613	7077	2155	7126	49	cg	this study
32	HRV-A	28	6101-CV29	dq473508-28	hrv-28*	DQ473508	ICAM1	2	622	7104	2161	7148	44	cg	(S24)
33	HRV-A	29	ATCC	hrv-29	hrv-29	FJ445125	LDLR	2	613	7074	2154	7123	49	cg	this study
34	HRV-A	30	106F	dq473512-30	not inc	DQ473512	LDLR	2	609	7055	2149	7099	44	cg	(S24)
35	HRV-A	30	ATCC	hrv-30	hrv-30	FJ445179	LDLR	2	609	7055	2149	7099	44	cg	this study
36	HRV-A	31	ATCC	hrv-31	hrv-31	FJ445126	LDLR	2	612	7082	2157	7131	49	cg	this study
37	HRV-A	32	ATCC	hrv-32	hrv-32	FJ445127	ICAM1	1	611	7084	2158	7133	49	cg	this study
38	HRV-A	33	ATCC	hrv-33	hrv-33	FJ445128	ICAM1	2	620	7084	2155	7133	49	cg	this study
39	HRV-A	34	ATCC	hrv-34	hrv-34	FJ445189	ICAM1	2	612	7070	2153	7119	49	cg	this study
40	HRV-A	34	137-3	dq473501-34	not inc	DQ473501	ICAM1	2	612	7070	2153	7119	49	cg	(S24)
41	HRV-A	36	342H	dq473505-36	hrv-36*	DQ473505	ICAM1	2	617	7099	2161	7141	42	cg	(S24)
42	HRV-A	38	ATCC	hrv-38	hrv-38	FJ445180	ICAM1	2	611	7087	2159	7136	49	cg	this study
43	HRV-A	38	CH79	dq473495-38	not inc	DQ473495	ICAM1	2	611	7087	2159	7136	49	cg	(S24)
44	HRV-A	39	209-Maryland/62	ay751783-39	hrv-39*	AY751783	ICAM1	2	615	7085	2157	7136	51	cg	(S27)
45	HRV-A	40	ATCC	hrv-40	hrv-40	FJ445129	ICAM1	2	613	7089	2159	7138	49	cg	this study
46	HRV-A	41	56110- North Carolina/61	dq473491-41	hrv-41*	DQ473491	ICAM1	2	616	7098	2161	7145	47	cg	(S24)
47	HRV-A	43	ATCC	hrv-43	hrv-43	FJ445131	ICAM1	1	614	7087	2158	7129	42	cg	this study
48	HRV-A	44	71560- North Carolina/61	dq473499-44	hrv-44*	DQ473499	LDLR	2	613	7074	2154	7123	49	cg	(S24)
49	HRV-A	45	ATCC	hrv-45	hrv-45	FJ445132	ICAM1	2	616	7071	2152	7114	43	cg	this study
50	HRV-A	46	Crell - Baylor 2- Texas/64	dq473506-46	hrv-46*	DQ473506	ICAM1	2	613	7104	2164	7149	45	cg	(S24)
51	HRV-A	47	ATCC	hrv-47	hrv-47	FJ445133	LDLR	2	613	7083	2157	7132	49	cg	this study
52	HRV-A	49	8213	dq473496-49	hrv-49*	DQ473496	LDLR	2	613	7062	2150	7106	44	cg	(S24)
53	HRV-A	49	fs ship#2	hrv-49-f04	hrv-49-f04	FJ445134	LDLR	2	616	7065	2150	7109	44	cg	this study
54	HRV-A	50	ATCC	hrv-50	hrv-50	FJ445135	ICAM1	2	611	7069	2153	7118	49	cg	this study

55	HRV-A	51	ATCC	hrv-51	hrv-51	FJ445136	ICAM1	2	614	7108	2165	7152	44	cg	this study
56	HRV-A	53	F01-3928	dq473507-53	hrv-53*	DQ473507	ICAM1	2	613	7098	2162	7143	45	cg	(S24)
57	HRV-A	54	ATCC	hrv-54	hrv-54	FJ445138	ICAM1	1	612	7085	2158	7134	49	cg	this study
58	HRV-A	54	fs ship#1	hrv-54-f05	hrv-54-f05	FJ445139	ICAM1	1	611	7084	2158	7133	49	cg	this study
59	HRV-A	55	Wis315E-Wisconsin-64	dq473511-55	hrv-55*	DQ473511	ICAM1	2	523	6987	2155	7036	49	cg-5	(S24)
60	HRV-A	56	ATCC	hrv-56	hrv-56	FJ445140	ICAM1	2	614	7087	2158	7136	49	cg	this study
61	HRV-A	57	fs ship#1	hrv-57	hrv-57	FJ445141	ICAM1	2	615	7085	2157	7134	49	cg	this study
62	HRV-A	58	ATCC	hrv-58	hrv-58	FJ445142	ICAM1	2	619	7098	2160	7140	42	cg	this study
63	HRV-A	59	611-CV35	dq473500-59	hrv-59*	DQ473500	ICAM1	2	612	7085	2158	7135	50	cg	(S24)
64	HRV-A	60	ATCC	hrv-60	hrv-60	FJ445143	ICAM1	2	614	7090	2159	7139	49	cg	this study
65	HRV-A	61	ATCC	hrv-61	hrv-61	FJ445144	ICAM1	2	619	7098	2160	7139	41	cg	this study
66	HRV-A	62	ATCC	hrv-62	hrv-62	FJ445145	LDLR	2	615	7082	2156	7131	49	cg	this study
67	HRV-A	63	ATCC	hrv-63	hrv-63	FJ445146	ICAM1	2	619	7092	2158	7141	49	cg	this study
68	HRV-A	64	ATCC	hrv-64	hrv-64	FJ445181	ICAM1	2	615	7082	2156	7129	47	cg	this study
69	HRV-A	64	6258-CV44	ef173417-64	not inc	EF173417	ICAM1	2	615	7082	2156	7129	47	cg	(S25)
70	HRV-A	65	ATCC	hrv-65	hrv-65	FJ445147	ICAM1	2	624	7118	2165	7162	44	cg	this study
71	HRV-A	66	ATCC	hrv-66	hrv-66	FJ445148	ICAM1	2	617	7090	2158	7139	49	cg	this study
72	HRV-A	67	ATCC	hrv-67	hrv-67	FJ445149	ICAM1	2	612	7085	2158	7135	50	cg	this study
73	HRV-A	68	ATCC	hrv-68	hrv-68	FJ445150	ICAM1	2	623	7120	2166	7164	44	cg	this study
74	HRV-A	71	ATCC	hrv-71	hrv-71	FJ445152	ICAM1	2	626	7117	2164	7161	44	cg	this study
75	HRV-A	73	107E	dq473492-73	hrv-73*	DQ473492	ICAM1	2	616	7095	2160	7140	45	cg	(S24)
76	HRV-A	74	328A	dq473494-74	hrv-74*	DQ473494	ICAM1	2	611	7078	2156	7120	42	cg	(S24)
77	HRV-A	75	328F	dq473510-75	hrv-75*	DQ473510	ICAM1	2	618	7091	2158	7137	46	cg	(S24)
78	HRV-A	76	H00062	dq473502-76	not inc	DQ473502	ICAM1	2	620	7084	2155	7129	45	cg	(S24)
79	HRV-A	76	ATCC 1x 3185 - 3310	hrv-76	hrv-76	FJ445182	ICAM1	2	619	7083	2155	7128	45	cg	this study
80	HRV-A	77	ATCC	hrv-77	hrv-77	FJ445154	ICAM1	2	614	7087	2158	7136	49	cg	this study
81	HRV-A	78	ATCC	hrv-78	hrv-78	FJ445183	ICAM1	2	623	7099	2159	7145	46	cg	this study
82	HRV-A	78	2030-65	ef173418-78	not inc	EF173418	ICAM1	2	623	7099	2159	7145	46	cg	(S25)
83	HRV-A	80	ATCC	hrv-80	hrv-80	FJ445156	ICAM1	2	616	7095	2160	7138	43	cg	this study
84	HRV-A	81	ATCC	hrv-81	hrv-81	FJ445157	ICAM1	1	618	7076	2153	7116	40	cg	this study
85	HRV-A	81	fs ship#1	hrv-81-f06	hrv-81-f06	FJ445158	ICAM1	1	618	7076	2153	7116	40	cg	this study
86	HRV-A	81	fs ship#2	hrv-81-f07	hrv-81-f07	FJ445159	ICAM1	1	618	7076	2153	7116	40	cg	this study
87	HRV-A	82	ATCC	hrv-82	hrv-82	FJ445160	ICAM1	2	615	7079	2155	7123	44	cg	this study
88	HRV-A	82	Santa Cruz,	dq473509-82	not inc	DQ473509	ICAM1	2	621	7085	2155	7129	44	cg	(S24)

			CA												
89	HRV-A	85	ATCC	hrv-85	hrv-85	FJ445163	ICAM1	2	615	7091	2159	7140	49	cg	this study
90	HRV-A	85	50-525-CV54	pico-85	not inc	AF343642, AY450517	ICAM1	2	615	7091	2159	7140	49	cg	Pico-DB (7)
91	HRV-A	88	CVD 01-0165- Dam-brauskas	dq473504-88	hrv-88*	DQ473504	ICAM1	2	622	7101	2160	7143	42	cg	(S24)
92	HRV-A	89	na	a10937-89	not inc	A10937	ICAM1	2	619	7110	2164	7152	42	cg	US patent DE3628658-A1 03/03/88
93	HRV-A	89	ATCC	hrv-89	hrv-89	FJ445184	ICAM1	2	619	7110	2164	7152	42	cg	this study
94	HRV-A	89	fs ship#2 isolate#A	hrv-89-f08	hrv-89-f08	FJ445166	ICAM1	2	617	7108	2164	7150	42	cg	this study
95	HRV-A	89	fs ship#2 isolate#B	hrv-89-f09	hrv-89-f09	FJ445165	ICAM1	2	619	7110	2164	7152	42	cg	this study
96	HRV-A	89	41467-Gallo	m16248-89	not inc	M16248	ICAM1	2	619	7110	2164	7152	42	cg	(S28)
97	HRV-A	90	ATCC	hrv-90	hrv-90	FJ445167	ICAM1	2	617	7075	2153	7124	49	cg	this study
98	HRV-A	94	ATCC	hrv-94	hrv-94	FJ445185	ICAM1	2	616	7083	2156	7132	49	cg	this study
99	HRV-A	94	SF-1803	ef173419-94	not inc	EF173419	ICAM1	2	616	7083	2156	7132	49	cg	(S25)
100	HRV-A	95	ATCC	hrv-95	hrv-95	FJ445170	ICAM1	2	609	7070	2154	7110	40	cg	this study
101	HRV-A	96	ATCC	hrv-96	hrv-96	FJ445171	ICAM1	2	618	7088	2157	7134	46	cg	this study
102	HRV-A	98	ATCC	hrv-98	hrv-98	FJ445173	ICAM1	2	611	7084	2158	7133	49	cg	this study
103	HRV-A	100	ATCC	hrv-100	hrv-100	FJ445175	ICAM1	2	612	7091	2160	7140	49	cg	this study
104	HRV-C	na	QPM	ef186077-qpm	qpm*	EF186077	na	na	438	6866	2143	6917	51	cg-5	(S29)
105	HRV-C	na	Nat001	ef077279-nat001	nat001*	EF077279 (16)	na	na	614	7039	2142	7079	40	cg	(S24) and this study
106	HRV-C	na	ny1078	ny1078	ny1078	unpub (17)	na	na	609	7025	2139	7072	47	cg	(S30, S31)
107	HRV-C	na	c024	ef582385-c024	c024*	EF582385	na	na	616	7047	2144	7099	52	cg	(S32)
108	HRV-C	na	Nat045	ef077280-nat045	nat045*	EF077280	na	na	542	6973	2144	7015	42	cg-5	(S24)
109	HRV-C	na	c026	ef582387-c026	c026*	EF582387	na	na	612	7037	2142	7086	49	cg	(S32)
110	HRV-C	na	c025	ef582386-c025	c025*	EF582386	na	na	617	7072	2152	7114	42	cg	(S32)
111	HRV-B	3	FEB	dq473485-03	hrv-03*	DQ473485	ICAM1	1	627	7160	2178	7208	48	cg	(S24)
112	HRV-B	3	FEB	ef173422-03	not inc	EF173422	ICAM1	1	627	7160	2178	7211	51	cg	(S25)
113	HRV-B	4	16/60	dq473490-04	hrv-04*	DQ473490	ICAM1	1	626	7156	2177	7212	56	cg	(S24)
114	HRV-B	5	ATCC	hrv-05	hrv-05	FJ445112	ICAM1	1	632	7162	2177	7212	50	cg	this study
115	HRV-B	6	Thompson	dq473486-06	hrv-06*	DQ473486	ICAM1	1	628	7164	2179	7216	52	cg	(S24)
116	HRV-B	14	1059- South Carolina/59	l05355-14	hrv-14*	L05355	ICAM1	1	629	7165	2179	7212	47	cg	(S33)

117	HRV-B	14	1060- South Carolina/59	k02121-14	not inc	K02121	ICAM1	1	629	7165	2179	7212	47	cg	(S34)
118	HRV-B	17	33342- North Carolina/59	ef173420-17	hrv-17*	EF173420	ICAM1	1	619	7176	2186	7219	43	cg	(S25)
119	HRV-B	26	ATCC	hrv-26	hrv-26	FJ445124	ICAM1	1	633	7160	2176	7211	51	cg	this study
120	HRV-B	27	ATCC	hrv-27	hrv-27	FJ445186	ICAM1	1	628	7161	2178	7217	56	cg	this study
121	HRV-B	27	5870-CV28	ef173421-27	not inc	EF173421	ICAM1	1	628	7161	2178	7217	56	cg	(S25)
122	HRV-B	35	164A	dq473487-35	not inc	DQ473487	ICAM1	1	623	7168	2182	7224	56	cg	(S24)
123	HRV-B	35	ATCC	hrv-35	hrv-35	FJ445187	ICAM1	1	623	7168	2182	7224	56	cg	this study
124	HRV-B	37	151-1	ef173423-37	hrv-37*	EF173423	ICAM1	1	631	7164	2178	7216	52	cg	(S25)
125	HRV-B	42	ATCC	hrv-42	hrv-42	FJ445130	ICAM1	1	633	7163	2177	7223	60	cg	this study
126	HRV-B	48	1505	dq473488-48	hrv-48*	DQ473488	ICAM1	1	623	7174	2184	7214	40	cg	(S24)
127	HRV-B	52	ATCC	hrv-52	hrv-52	FJ445188	ICAM1	1	625	7173	2183	7216	43	cg	this study
128	HRV-B	52	fs ship#1	hrv-52-f10	hrv-52-f10	FJ445137	ICAM1	1	625	7173	2183	7216	43	cg	this study
129	HRV-B	52	F01-3772	ef173424-52	not inc	EF173424	ICAM1	1	625	7173	2183	7216	43	cg	(S25)
130	HRV-B	69	ATCC	hrv-69	hrv-69	FJ445151	ICAM1	1	621	7169	2183	7211	42	cg	this study
131	HRV-B	70	F02-2547-Treganza	dq473489-70	hrv-70*	DQ473489	ICAM1	1	623	7180	2186	7223	43	cg	(S24)
132	HRV-B	72	ATCC	hrv-72	hrv-72	FJ445153	ICAM1	1	631	7167	2179	7216	49	cg	this study
133	HRV-B	79	ATCC	hrv-79	hrv-79	FJ445155	ICAM1	1	625	7170	2182	7224	54	cg	this study
134	HRV-B	83	ATCC	hrv-83	hrv-83	FJ445161	ICAM1	1	627	7175	2183	7230	55	cg	this study
135	HRV-B	84	ATCC	hrv-84	hrv-84	FJ445162	ICAM1	1	622	7152	2177	7201	49	cg	this study
136	HRV-B	86	ATCC	hrv-86	hrv-86	FJ445164	ICAM1	1	632	7165	2178	7213	48	cg	this study
137	HRV-B	91	ATCC	hrv-91	hrv-91	FJ445168	ICAM1	1	621	7178	2186	7221	43	cg	this study
138	HRV-B	92	ATCC	hrv-92	hrv-92	FJ445169	ICAM1	1	628	7176	2183	7233	57	cg	this study
139	HRV-B	93	SF-1492	ef173425-93	hrv-93*	EF173425	ICAM1	1	629	7162	2178	7215	53	cg	(S25)
140	HRV-B	97	ATCC	hrv-97	hrv-97	FJ445172	ICAM1	1	626	7156	2177	7207	51	cg	this study
141	HRV-B	99	ATCC	hrv-99	hrv-99	FJ445174	ICAM1	1	626	7156	2177	7208	52	cg	this study

Table S1 Legend

- (1) Order in which these sequences are listed in accompanying protein (TableS3) and RNA (TableS2) alignment files.
- (2) Designation of Human Rhinovirus-C (HRV-C) as a species has been accepted by the ICTV Study Group on Picornaviruses.
- (3) Field strains were assigned serotypes on the basis of sequence similarity to reference (ATCC) strains.
- (4) Strain designation from GenBank sequence description if available; ATCC origin; or field-strain isolate from this study.
- (5) Protein and RNA alignments use these names. Accession number, if available, precedes serotype; f01-f10: field strains; entries beginning with "hrv-" are from this study.
- (6) (abbreviated) Strain names used on phylograms. Not all alignment entries are included (not inc). Asterisk (*) designates GenBank origin, sequence is not original to this study.
- (7) GenBank accession numbers for published genomes (black) or from this study (green). Pico-01, pico-09 and pico-85 genomes were pieced from the indicated GenBank fragments (see accession numbers) according to linkage guides on The Picornavirus Database (Pico-DB) at: <http://www.picornaviridae.com/>
- (8) The Major (M) group HRV use ICAM-1 receptors. The Minor (m) group HRVs use LDLR-like receptors. See reference: (S35).
- (9) Drug group reactivity according to reference: (S36). na: not part of that study.
- (10) 1st base of polyprotein open reading frame within individual sequence file.
- (11) Last base of polyprotein open reading frame within individual sequence file. Count excludes the ORF termination codon.
- (12) Polyprotein length, in amino acids, encoded by this genome.
- (13) Genome sequence length (as available) as it is included within TableS2 alignment file.
- (14) 3' non-coding fragment length, as included in TableS2 alignment file. Count does not include poly(A).
- (15) cg: complete genome. cg-5: almost complete genome, except for 5' end.
- (16) Chimeric sequence: bases 1-135 from this study, remainder from EF077279.
- (17) Thomas Briese and Ian Lipkin, personal communication to ACP.
- (18) First entry of TableS2 and TableS3 alignments include an annotation line with protein structure features, polyprotein cleavage sites, ORF start/stop, etc.

Major Parent	Minor Parent	Recombinant	start	end	Genome Region	P-RDP	P-BooScan	P-MaxChi	P-Chimera	P-Siscan	P-3Seq
hvv-45	hvv-21	hvv-8 (a)	13	631	5'-UTR	3.351x10 ⁻²⁰	1.094x10 ⁻²¹	5.518x10 ⁻⁰⁸	4.7x10 ⁻⁹	6.061x10 ⁻¹²	
hvv-45	hvv-21	hvv-95 (a)	34	599	5'UTR	3.351x10 ⁻²⁰	1.094x10 ⁻²¹	5.518x10 ⁻⁰⁸	4.7x10 ⁻⁹	6.061x10 ⁻¹²	
hvv-65	hvv-21	hvv-80 (b)	10	756	5'-UTR, VP4	1.077x10 ⁻²¹	6.129x10 ⁻²³	9.417x10 ⁻⁸	1.584x10 ⁻³	5.099x10 ⁻¹²	2.235x10 ⁻⁵
hvv-51	hvv-11	hvv-20 (b)	6	690	5'-UTR, VP4	1.023x10 ⁻¹⁹	8.402x10 ⁻²⁶	1.606x10 ⁻⁶	7.286x10 ⁻⁷	2.187x10 ⁻²⁰	2.101x10 ⁻¹²
hvv-51	hvv-11	hvv-68 (b)	11	691	5'-UTR, VP4	1.023x10 ⁻¹⁹	8.402x10 ⁻²⁶	1.606x10 ⁻⁶	7.286x10 ⁻⁷	2.187x10 ⁻²⁰	2.101x10 ⁻¹²
hvv-28	hvv-62	hvv-71	1	661	5'-UTR	2.582x10 ⁻⁹	8.484x10 ⁻¹⁵	1.194x10 ⁻³	4.255x10 ⁻⁶		1.899x10 ⁻⁸
hvv-54-f05	hvv-75	hvv-18	511	3058	5'-UTR, VP4, VP2, VP3, VP1	1.371x10 ⁻³		1.555x10 ⁻⁶	3.737x10 ⁻⁷	8.391x10 ⁻³	2.173x10 ⁻³
hvv-54-f05	hvv-75	hvv-24 (b)	540	3071	5-UTR, VP4, VP2, VP3, VP1	1.371x10 ⁻³		1.555x10 ⁻⁶	3.737x10 ⁻⁷	8.391x10 ⁻³	2.173x10 ⁻³
hvv-54-f05	hvv-75	hvv-50 (b)	554	3294	5-UTR, VP4, VP2, VP3, VP1, P2A	1.371x10 ⁻³		1.555x10 ⁻⁶	3.737x10 ⁻⁷	8.391x10 ⁻³	2.173x10 ⁻³
hvv-54-f05	hvv-75	hvv-34	651	3088	VP4, VP2, VP3, VP1	1.371x10 ⁻³		1.555x10 ⁻⁶	3.737x10 ⁻⁷	8.391x10 ⁻³	2.173x10 ⁻³
hvv-54-f05	hvv-67	hvv-38 (b)	1006	4935	VP2, VP3, VP1, P2A, P2B, P2C, P3A		1.561x10 ⁻²⁹	9.616x10 ⁻¹⁴	1.322x10 ⁻¹⁴	8.465x10 ⁻⁵	3.552x1 ⁻¹⁸
hvv-54	hvv-67	hvv-60 (b)	995	4979	VP2, VP3, VP1, P2A, P2B, P2C, P3A	8.952x10 ⁻¹²	1.561x10 ⁻²⁹	9.616x10 ⁻¹⁴	1.322x10 ⁻¹⁴	8.465x10 ⁻⁵	3.552x1 ⁻¹⁸
hvv-53	hvv-80	hvv-46 (a)	32	3222	5'-UTR, VP4, VP2, VP3, VP1	1.153x10 ⁻²²	2.641x10 ⁻¹⁶	3.295x10 ⁻²³	1.261x10 ⁻²⁵	3.511x10 ⁻¹⁰	1.291x10 ⁻⁶²
hvv-13-f03	HRV41	hvv-73	64	3333	5'-UTR, VP4, VP2, VP3, VP1	1.606x10 ⁻²		3.503x10 ⁻⁹	9.419x10 ⁻¹²		2.363x10 ⁻⁸
hvv-30	hvv-59	hvv-39 (b)	450	3148	5'-UTR, VP4, VP2, VP3, VP1	1.958x10 ⁻⁷	6.012x10 ⁻³	9.356x10 ⁻¹¹	2.963x10 ⁻¹²	8.613x10 ⁻⁷	1.023x10 ⁻¹¹
hvv-68	hvv-20	hvv-28 (a)	691	4143	VP4, VP2, VP3, VP1, P2A, P2B, P2C			4.245x10 ⁻¹¹	9.232x10 ⁻⁶	1.932x10 ⁻⁶	6.839x10 ⁻¹²
hvv-100	hvv-10	hvv-56 (b)	3487	6889	P2A, P2B, P2C, P3A, P3B, P3C, P3D	4.315x10 ⁻³	4.230x10 ⁻¹⁷	1.23x10 ⁻⁷	1.921x10 ⁻¹⁰		5.416x10 ⁻⁷
hvv-29	hvv-54-f05	hvv-31 (a)	5181	7007	P3C, P3D	3.3x10 ⁻¹⁰	7.934x10 ⁻¹³	1.038x10 ⁻⁶	5.174x10 ⁻⁸	1.232x10 ⁻¹⁰	
hvv-29	hvv-54-f05	hvv-47 (a)	5321	7008	P3C, P3D	3.3x10 ⁻¹⁰		1.038x10 ⁻⁶	5.174x10 ⁻⁸	1.232x10 ⁻¹⁰	
hvv-42	hvv-97	hvv-4 (a)	1	640	5'-UTR, VP4	1.322x10 ⁻¹²	8.233x10 ⁻³⁵	1.315x10 ⁻¹⁰	1.126x10 ⁻⁸	9.599x10 ⁻²²	4.975x10 ⁻⁴
hvv-84	hvv-37	hvv-27 (b)	48	718	5'-UTR, VP4	7.842x10 ⁻⁴	1.500x10 ⁻¹⁰	9.948x10 ⁻⁴	1.061x10 ⁻³	2.364x10 ⁻¹⁰	1.792x10 ⁻⁴
hvv-84	hvv-37	hvv-93 (b)	100	772	5'-UTR, VP4	7.842x10 ⁻⁴	1.500x10 ⁻¹⁰	9.948x10 ⁻⁴	1.061x10 ⁻³	2.364x10 ⁻¹⁰	1.792x10 ⁻⁴
hvv-84	hvv-37	hvv-97 (b)	101	748	5'-UTR, VP4	7.842x10 ⁻⁴	1.500x10 ⁻¹⁰	9.948x10 ⁻⁴	1.061x10 ⁻³	2.364x10 ⁻¹⁰	1.792x10 ⁻⁴

Table S4. Recombination analysis based on the HHM-based alignment using the complete genomes from the reference set and the field samples. Results are shown for all recombination events with $P < 0.00001$ by at least 2 analysis modes (see Methods). P-values represent the results using the indicated analysis modes. The nucleotide is referenced to the recombinant (daughter) hrv. See Fig. 3 for representative recombinations in graphical forms. When a progressive alignment was performed and recombination tested, 19 of the 23 recombinants were confirmed (a, major and minor parents were the same as with HHM; b, one or both parents differed from the HHM-based analysis).

Table S5. Amino acid differences between the HRV reference genomes and the analogous field isolates. The position is the amino acid number of the reference genome for each comparison.

	hrv-52-f10	hrv-89-f08	hrv-89-f09	hrv-81-f06	hrv-81-f07	hrv-13-f03	hrv-09-f01	hrv-09-f02	hrv-49-f04	hrv-54-f05
40							R40G	R40G		
59							S59D	S59D		
136						K136E				E136D
142				N142S	N142S		G142N	G142N		
143							D143E	D143E		
156									R156K	
173						C173S				
193										V193I
199							D199H	D199H		
215									N215K	
216									-f216Y	
222	Q222S								T222A	
227	N227D									
229	I229T					I229T				-229P
230						N230G				A230-
232	E232A		T232S						S232N	V232T
233									M233I	
234										T234I
237				K237R	K237R				L237V	N237H
271										T271S
305										P305S
308	N308S									
310	T310S									
335		S335N	S335N							
339	V339I									
340	P340-									
344	-344K								S344N	
351										V351I
404										-f404L
405	T405V			T405A	T405A					A405P
406						N406T				
410	S410T									
411									E411G	
412									N412D	N412S
417									V417I	
419										-f419S
421							Q421T	Q421T		
422	K422R								V422I	
428							T428S	T428S		
474							S474T	S474T		
487	S487A									
488									A488R	
491		K491R	K491R				T491I			
494									K494R	
506									V506I	
526	-f526Y									
529									P529S	
531									T531A	
534				X534S	X534S					
535									S535A	
538									I538V	
554									A554P	
555										S555N
584									V584P	
586				T586E	T586E					
589	-f589L									
592									H592Q	R592K
625						S625K				

626	P626S					S626P				
631				S631A	S631A					
656										-f656Y
662	T662N									
675				S675A	S675A					
679						H679Y				
681						S681H				
685							I685V	I685V		
686	E686K									Q686R
687	I687V									
688						T688D	S688D	S688D		
689						N689A				
690	N690D	T690V	T690V							
696	T696G									
706							E706K	E706K		
709	S709N									
713	V713I	V713I	V713I							
714	N714S									
721							L721I	L721I		
725		Y725-f	Y725-f							
734				N734D	N734D					
745		A745T	A745T							
749	T749A	Q749K	Q749K							
751	Q751E	N751D	N751D						K751E	V751A
753		S753N	S753N							
772	E772K									
786									V786I	
802	V802I									
827			S827N						Q827E	
828	A828P					S828T				
829	N829E			H829Q	H829Q					
836						I836V				
838				X838N	X838N					
847		I847M	I847M							
853	L853-f									K853E
855									S855N	
856		S856L	S856L							
858				I858V	I858V					
869						V869I				
892									V892I	
893	N893S	S893L	S893L			G893K			E893Q	R893Q
894							T894E	T894E	D894N	E894D
896							D896E	D896E		D896E
898							A898S	A898S		
911				K911R	K911R	A911T			T911K	
914	D914E								A914P	T914E
915										N915T
916		V916I	V916I							
917							K917R	K917R		V917T
922	N922S									
923	S923P									
928	V928I									
948							N948D	N948D		
952										V952I
959		I959V	I959V							
966							I966T	I966T		
973						S973A				
987									N987S	
993		R993K	R993K			E993N				

995	E995D									
1035	M1035V									
1041							I1041M			
1042		I1042V	I1042V							
1071									I1071V	
1092				D1092E	D1092E		D1092E	D1092E		
1102							S1102N	S1102N		S1102N
1105									V1105I	
1109										V1109I
1110							V1110I	V1110I		
1164				I1164V	I1164V					
1209	R1209K									
1223										-f1223L
1229	A1229T									
1233									V1233I	
1236				A1236T	A1236T					
1238		V1238D	V1238D	S1238T	S1238T					
1246		-f1246Y	-f1246Y							
1249							I1249V			
1265									G1265S	
1272										V1272I
1274										Y1274H
1277	V1277M									
1278							N1278S			
1294									I1294V	
1299							L1299S			
1305							P1305S	P1305S		
1412							Y1412-f	Y1412-f		
1418							V1418I	V1418I	I1418L	
1431							N1431D	N1431D		
1432	A1432V									
1437							R1437K			
1440									N1440D	
1441							V1441I			
1442							K1442D			
1443		I1443T	I1443T							
1450	K1450R									
1460	V1460I								-f1460V	
1461										V1461M
1467									S1467T	
1469				R1469K	R1469K					
1474	X1474D									
1478				N1478S	N1478S					
1482							E1482Q			
1491				V1491I	V1491I					
1514							V1514A	L1514S	L1514S	
1521							T1521A			
1530	K1530E									
1538									V1538D	Y1538D
1542							T1542I	T1542I		
1550							K1550R			
1556				N1556S	N1556S					
1560				N1560S	N1560S					
1563							A1563T			
1569				I1569V						
1572							V1572A	V1572A		
1600				I1600V	I1600V					
1630									E1630D	
1645							I1645V			

Supplementary References

- S1. A. Djikeng *et al.*, *BMC Genomics* **9**, 5 (2008).
- S2. T. Allander, S. U. Emerson, R. E. Engle, R. H. Purcell, J. Bukh, *Proc. Natl. Acad. Sci. U. S. A.* **98**, 11609 (2001).
- S3. T. Allander *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12891 (2005).
- S4. A. C. Palmenberg, J.-Y. Sgro, in *Molecular Biology of Picornaviruses*, B. L. Semler, E. Wimmer, Eds. (ASM Press, New York, 2001).
- S5. A. C. Palmenberg, J.-Y. Sgro, *Seminars in Virology* **8**, 231 (1997).
- S6. K. Tamura, J. Dudley, M. Nei, S. Kumar, *Mol. Biol. Evol.* **24**, 1596 (2007).
- S7. S. Guindon, O. Gascuel, *Syst. Biol.* **52**, 696 (2003).
- S8. D. P. Martin, C. Williamson, D. Posada, *Bioinformatics* **21**, 260 (2005).
- S9. D. Martin, E. Rybicki, *Bioinformatics* **16**, 562 (2000).
- S10. M. O. Salminen, J. K. Carr, D. S. Burke, F. E. McCutchan, *AIDS Res. Hum. Retroviruses* **11**, 1423 (1995).
- S11. J. M. Smith, *J. Mol. Evol.* **34**, 126 (1992).
- S12. M. J. Gibbs, J. S. Armstrong, A. J. Gibbs, *Bioinformatics* **16**, 573 (2000).
- S13. M. F. Boni, D. Posada, M. W. Feldman, *Genetics* **176**, 1035 (2007).
- S14. C. Notredame, D. G. Higgins, J. Heringa, *J. Mol. Biol.* **302**, 205 (2000).
- S15. R. C. Edgar, *Nucleic Acids Res.* **32**, 1792 (2004).
- S16. M. Suyama, D. Torrents, P. Bork, *Nucleic Acids Res.* **34**, W609-W612 (2006).
- S17. T. C. Bjorndahl, H. Monzavi, D. S. Wishart, *J. Biomol. NMR* **26**, 85 (2003).
- S18. E. Arnold, M. G. Rossmann, *J. Mol. Biol.* **211**, 763 (1990).
- S19. D. W. Gohara *et al.*, *J. Biol. Chem.* **275**, 25523 (2000).

- S20. J. F. Petersen *et al.*, *EMBO J.* **18**, 5463 (1999).
- S21. D. M. Strauss, L. W. Glustrom, D. S. Wuttke, *J. Mol. Biol.* **330**, 225 (2003).
- S22. B. Brown, M. S. Oberste, K. Maher, M. A. Pallansch, *J. Virol.* **77**, 8973 (2003).
- S23. V. R. Racaniello, D. Baltimore, *Proc. Natl. Acad. Sci. U. S. A.* **78**, 4887 (1981).
- S24. P. J. Hughes, C. North, C. H. Jellis, P. D. Minor, G. Stanway, *J. Gen. Virol.* **69 (Pt 1)**, 49 (1988).
- S25. T. Skern *et al.*, *Nucleic Acids Res.* **13**, 2111 (1985).
- S26. A. L. Kistler *et al.*, *Virol. J.* **4**, 40 (2007).
- S27. C. Tapparel *et al.*, *BMC Genomics* **8**, 224 (2007).
- S28. W. M. Lee, W. Wang, R. R. Rueckert, *Virus Genes* **9**, 177 (1995).
- S29. J. R. Harris, V. R. Racaniello, *J. Virol.* **79**, 5363 (2005).
- S30. M. Duechler *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **84**, 2605 (1987).
- S31. P. McErlean *et al.*, *J. Clin. Virol.* **39**, 67 (2007).
- S32. N. Renwick *et al.*, *J. Infect. Dis.* **196**, 1754 (2007).
- S33. D. Lamson *et al.*, *J. Infect. Dis.* **194**, 1398 (2006).
- S34. S. K. Lau *et al.*, *J. Clin. Microbiol.* **45**, 3655 (2007).
- S35. F. Hofer *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **91**, 1839 (1994).
- S36. G. Stanway, P. J. Hughes, R. C. Mountford, P. D. Minor, J. W. Almond, *Nucleic Acids Res.* **12**, 7859 (1984).
- S37. H. Shimodaira, M. Hasegawa, *Bioinformatics* **17**, 1246 (2001).