Chapter 1

Classification and Evolution of Human Rhinoviruses

[AU1] Ann C. Palmenberg and James E. Gern

Abstract

The historical classification of human rhinoviruses (RV) by serotyping has been replaced by a logical system of comparative sequencing. Given that strains must diverge within their capsid sequenced by a reasonable degree (>12–13 % pairwise base identities) before becoming immunologically distinct, the new nomenclature system makes allowances for the addition of new, future types, without compromising historical designations. Currently, three species, the RV-A, RV-B, and RV-C, are recognized. Of these, the RV-C, discovered in 2006, are the most unusual in terms of capsid structure, receptor use, and association with severe disease in children.

Key words Rhinovirus, Evolution, Virus taxonomy, Immunology, Drug resistance

1 Historical RV Classification

The human rhinoviruses currently comprise the RV-A, RV-B, and RV-C species of the *Enterovirus* genus in the *Picornaviridae* family. This classification status was not always the case. For the RV-A and RV-B, several historic clinical panels archived by the American Type Culture Collection were originally combined and indexed into 100 RV types after assessment of antigenic cross-reactivity or serotyping in rabbits. From these data, and from physical characteristics of the virions (e.g., pH lability), it was obvious that the full list of composite isolates easily subdivided into two related species, the *HRV-A* and *HRV-B*. For many years, these units were assigned to their own genus (*human rhinoviruses or HRV*) because the disease presentations (common cold) were observably different from other classical enteroviruses, like poliovirus, coxsackie virus, or ECHO (enteric cytopathic human orphan) viruses. Moreover, they were also different from all other original picornavirus genera, the *Aphthoviruses*, *Cardioviruses*, and *Hepatoviruses*. Before 1985, most virus taxonomy systems were weighted heavily towards phenotypic parameters (i.e., virion stability properties or disease etiologies)

David A. Jans and Reena Ghildyal (eds.), *Rhinoviruses: Methods and Protocols*, Methods in Molecular Biology, vol. 1221, DOI 10.1007/978-1-4939-1571-2_1, © Springer Science+Business Media New York 2015

Fig. 1 Genome map of a typical RV shows the protein names and their involvement in capsid formation or replication processes

as it was commonly argued (at that time) that medical-based classification made it easier to teach in clinical settings.

In 1984–1985, the first HRV (B14) was sequenced in its entirety in parallel with the virion crystal structure determination [1, 2]. Surprisingly, the B14 genome organization, including the full array of functional genes (Fig. 1), proved nearly identical to that of poliovirus 1, one of the earliest determined picornavirus sequences [3, 4]. Indeed, as more genome sequences followed, the pattern became evermore apparent. The HRV-A and HRV-B, while distinct in their own groupings, were enterovirus-like in all measures of genome comparisons and probably should be considered as species within that genus. The Picornavirus Study Group Subcommittee (SG) eventually made this recommendation to the International Committee on the Taxonomy of Viruses (ICTV), where it was subsequently adopted [5]. The thorough, deep sequence-based classification precedent established by this decision has helped shape taxonomy protocols for all virus families. As part of the *HRV-Enterovirus* reassignments, the term "Human" was dropped from the species names and 99 of the original types became simply known as "Rhinoviruses," or "RV," retaining the RV-A and RV-B species letter designations of the previous system. Within this reclassification context, and after further evaluation of genetic, immunogenic, and receptor use (decay-accelerating factor as a receptor) properties, RV-A87 was reassigned to the *Enterovirus D* species (EV-D68) [6].

2 Current Classification

New RV isolates are now rarely tested for immunogenicity. The current classification scheme is based on overt similarities in genome organization, capsid properties, and primary sequence conservation [7]. Strains are assigned to the RV-A or RV-B if they share greater than 70 % amino acid identity in the P1, 2C, and 3CD regions with other members. Within the respective species, isolates are subdivided into numeric genotypes that respect the historic naming system, but now rely almost entirely on sequence comparisons of the VP1 protein or VP4/VP2. The preferred nomenclature [8] designates the species letter

(A, B, or C), and type number (e.g., A16). Strain designations are unique to each Genbank accession number and rarely indicated unless required for clarity.

Assignment of a new strain to a known genotype generally requires >86–87 % aligned nucleic acid identity in either or both of the key capsid regions. Type assignments are considered tentative until at least the full VP1 sequence is completed and verified [8]. Full-genome sequencing revealed that some historic types were really more closely related than this (e.g., A54 and A98, or A29 and A44), and others such as A8 and A45, defining "clade D," were in fact so different from all other RV-A; they perhaps warrant eventual designation as another species [9]. Part of the ongoing mission of the Picornavirus Special Group (SG) is to continually sort out such discontinuities and attempt to provide a common code for new isolates and types as they are discovered. For example, in the past few years, six new types have been added to the RV-A (A101–A106) and five new types have been added to the RV-B (B100–B104). Isolates for A8 and A95 have been merged into a single type (A8), as have A54/A98 (to A54), and A29/A44 (to A29). Other types were split (e.g., B52 into B52 plus B104), or their isolates rearranged (e.g., A36 and A89). All these changes now more accurately reflect strain/type commonalities required by the overlying classification scheme.

An excellent recent review on this topic by McIntyre et al. summarizes the current state of the field [8]. Recent taxonomy proposals approved, or under consideration by the Picornavirus SG or by the ICTV, can be publically reviewed at http://www.ictvonline.org/virusTaxonomy.asp. Presently, the RV-A have 77 recognized types and the RV-B have 30 types. Type RV-A1 is unique in that it has assigned isolates that are sufficiently different as to warrant special distinction, as A1A and A1B subtypes. If these units are counted separately, it brings the RV-A to 78 types. Because of the recently recommended mergers among several closely related types, a few of the historic type numbers have been dropped from the current system and are no longer used (A44, A87, A95, A98). If a researcher should discover an isolate sufficiently different to warrant consideration as a new type, they should consult the website curated by the Picornavirus SG (http://www.picornaviridae.com). Via links on this site, comparative sequences can be submitted (preferably for the full capsid, but for the full VP1 gene at a minimum) for SG consideration. New type numbers are awarded sequentially. New species designations (*see* below for RV-C) require full ICTV approval.

3 Receptor and Drug Groups

The classic panel of 99 original RV-A and RV-B are the canonical agents of the "common cold." Many are well studied at the structural and clinical levels. All these isolates use either ICAM-1

(89 "major" types) or LDLR (10 "minor" types) as their cellular receptors. The molecular nuances of these interactions have been described by many co-crystallization and EM studies. The set of full-genome sequences, including at least one representative of each historic type, was completed in 2009 [10]. From this work, it became clear that the RV-A + B included in the major and minor groups conserve particular surface footprints that explain how and why these isolates use their respective receptors to interact with cells [11].

This same virus panel has been subjected to extensive characterizations according to composite strain sensitivities to a slate of potential therapeutics targeting their capsids [8]. The basic strategy is aimed at inhibiting the virus before infection by intercalating drugs into the unique surface "pockets" characteristic of all enterovirus virions. The type-specific sensitivities were found to subdivide, roughly along species lines, into two experimental groups [12]. The structures of 28 virus-drug complexes have been determined to atomic resolution. The Group-1 viruses (all RV-B plus A8, A13, A32, A43, A45, and A54) have long, narrow pockets interior to their VP1 proteins, which accommodate matching long-chain hydrophobic drugs. The Group-2 viruses (all other RV-A) have shorter, wider VP1 hydrophobic pockets, and therefore accept an alternate cohort of drugs. These points are important to any discussion of rhinovirus classification because there is frequent semantic confusion when dividing the historic strains into their species (RV-A or RV-B), or their receptor units (major or minor) or their drug Groups (1 or 2). It should be remembered that each term designates separate, non-overlapping properties. None of the most recently added RV types (i.e., A101–106 or B100–104) have ever been directly tested for receptor binding or drug sensitivity. Their respective activities, based on sequence comparison alone, *predict* them to be "major" in terms of receptor, but divided between Groups 1–2 (along species lines) for drug reactivity.

4 Rhinovirus C

In 2006 the discovery of a new RV species surprised the molecular and clinical communities [13]. The RV-C are clearly rhinoviruses, but unlike RV-A + B, they are not readily propagated in typical cell culture systems, including WI-38, WisL, BEAS-2B, A549, and HeLa lines [11]. These isolates are not "new" in terms of evolution, but rather they were physically undetected by all typical characterization methods that required cultured virus growth, such as plaque assays [11]. The current 51 recognized RV-C types (as binned by sequence analysis) were instead identified by PCR while fishing through patient samples for other RV. As with the RV-A + B, each RV-C type includes those isolates whose VP1 sequences

exceed 87 % pairwise identity at the nucleotide level [8, 14]. The RV-C have special clinical relevance since it is now recognized that these strains are associated with up to half of infections in young children [11]. They grow readily in both the lower and upper airways and tolerate higher growth temperatures in culture [15]. Moreover, the RV-C use cell receptors that are not common to the RV-A + B [11]. Unfortunately, these receptors are apparently lost whenever primary tissue snippets are transitioned to undifferentiated monolayers. RV-C can be grown in mucosal organ cultures, but this technique requires the availability of primary human donor samples [11]. Parallel work with differentiated sinus or bronchial epithelial cells at air-liquid interface (ALI) is promising [15, 16], but neither technique has yet produced enough virus for extensive biological studies. Instead, RV-C information relies heavily on comparative sequence analysis to maximize data from limited experimental samples.

To this end, a great many RV-C capsid fragments have been sequenced, and for about 32 types there are (nearly) full-length genome data [17]. Common to all known isolates in this species are unusually large relative deletions (indels) in the VP1 capsid protein. The fundamental VP1 protein cores superimpose among all RV, but the loops that connect the internal β-strands of the RV-C VP1 are shorter by ~22 amino acids relative to the RV-A, and ~28 amino acids relative to the RV-B. The composite structural loops containing these elements supply virtually all of the mass to the fivefold virion plateau. Therefore, the physical RV-C capsid structures are predicted to be very different from the RV-A + B over at least 1/3 of the virion surface [17]. The changes profoundly affect the receptor-binding platform, (predicted) type immunogenicity, and capsid-drug reactivity [17, 18].

5 Physical Characteristics

By way of review, all RV have genome organizations and (general) capsid structures similar to those of other *Enteroviruses* (Fig. 1). But unlike isolates in the other species of this genus, which remain viable at pH 3.0, RV particles (RV-A + B + C) are unstable below pH 5–6. The icosametric capsid (~30 nm diameter) has 60 copies each of proteins VP1, VP2, VP3, and VP4, named in order of descending electrophoretic mobility. The protein shell surrounds a densely packed, single-stranded, positive-sense, RNA genome of 7079 (RV-C1) to 7233 (RV-B92) bases, a count which does not include the variable length 3′ poly(A) tail. Like poliovirus, the surfaces of RV-A + B + C capsids are dominated by the three largest proteins. VP4 is internal to the structure, centered near the fivefold axis. Around the exterior fivefold plateau, a symmetrical "canyon" provides receptor-binding sites and immunogenic surfaces.

All RV-A + B are major (ICAM-1) or minor (LDLR) with regard to their receptor preference [19], but the cellular receptors used by the RV-C are certain to be different [11] and are currently unidentified. Bioinformatics predicts that the RV-C deletions in the VP1 regions will produce species-specific topologies for the canyon region and the fivefold plateau and the (common?) RV-C receptor is sure to be compatible with these dramatic changes.

The RV genomes are messenger sense, encoding the polyprotein reading frame (ORF) and multiple important RNA structural motifs. Adjacent to the 5′ cloverleaf, a regulatory feature for translation and replication, each RV encodes a strain-specific pyrimidine-rich tract that may be involved in suppressing innate immunity triggers [10]. The type-1 IRES is 3′ to this tract and includes a variable-length stem structure pairing the ORF start site (AUG) with an upstream AUG. Unlike poliovirus, intervening sequences between these AUGs are probably not scanned by initiating ribosomes [20]. The picornavirus VPg uridylylation reaction, required for RNA synthesis, is templated by a special structure called the *cre* (*cis*-acting replication element) whose location varies in every species of picornavirus. For the RV-A, the *cre* is in the 2A gene [21]. For the RV-B, the *cre* is in the 2C gene [21]. The RV-C *cre* has been proposed as one of the two sites in the 1B gene [10, 21, 22]. Neither has been confirmed experimentally. The short, 3′ untranslated sequences (UTR) are highly variable. Invariably, they configure as an inclusive stem motif displaying at least one bogus termination codon in the terminal loop. This codon may be in-frame or out-of-frame with the authentic ORF stop site, and has been proposed to play a role in the recruitment of translation termination factors [10].

6 Genetic Relationships

As might be expected from the original RV typing system, a large degree of sequence diversity among the RV manifests as amino acid changes in capsid surface regions mapped as neutralizing immunogenic epitopes (Nims). The high frequency of mutational fixation in these Nims, particularly for VP1, is one of the key reasons for the plethora of recognized RV genotypes. Although it is possible to measure and define comparative relationships among any set of extant isolates, it is virtually impossible to retrace the exact lineages that gave rise to them. “Evolutionary” trees created from VP1 data are quite different from those using VP2/4, 3D, 3C, the IRES, or other regions of the genome [10]. In part this is because nonstructural genes (except for 2A) fix mutations at more variable rates. But recombination (*see* below) is also frequent within and between strains from different RV species. Few if any of even the most characteristic lineages are known to breed true. At best a representative phylogram (Fig. 2) can illustrate some measure of relationships among the major clades and highlight

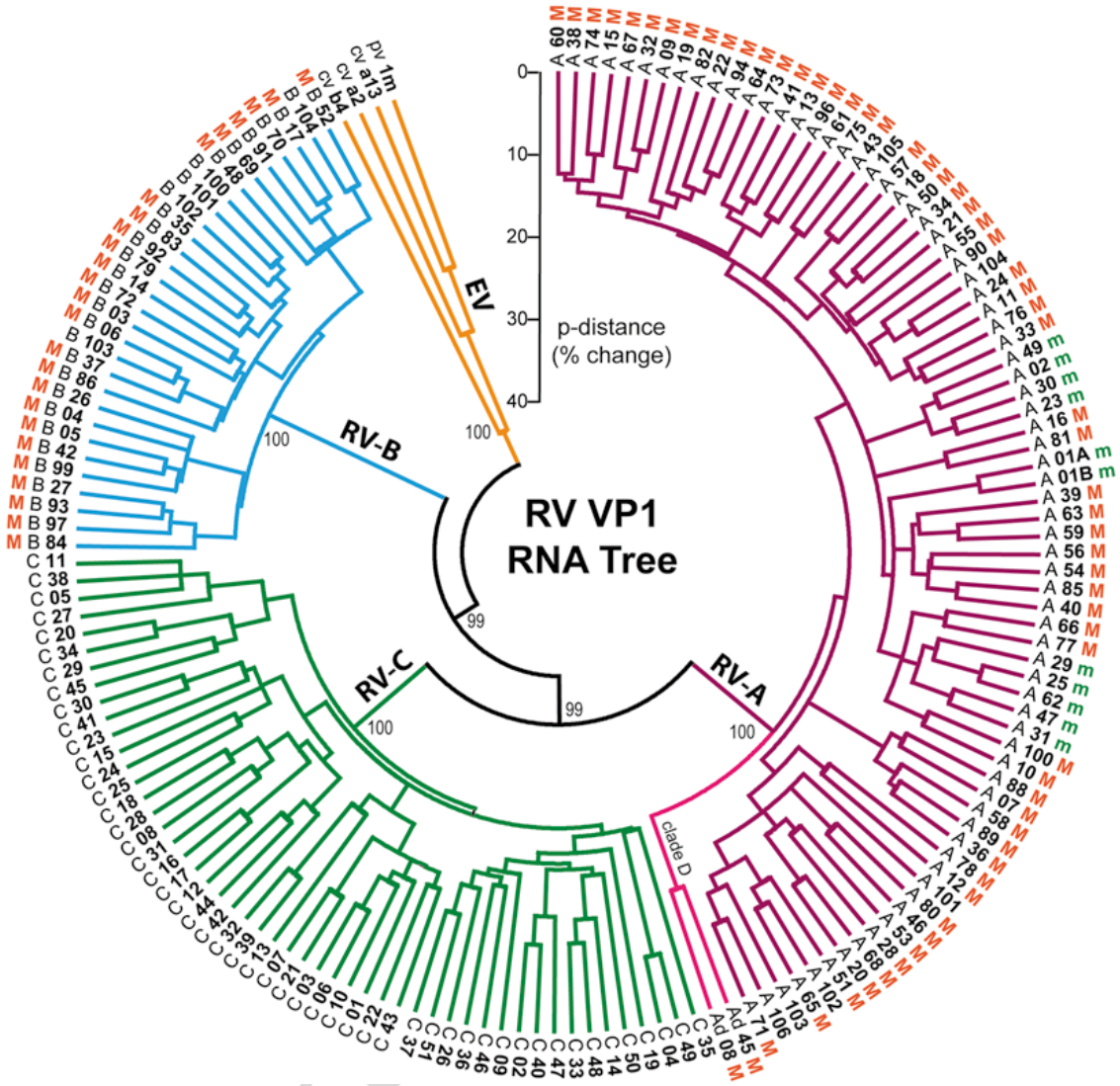

Fig. 2 Circle phylogram of relationships for currently recognized genotypes [8] of RV-A, RV-B, and RV-C. The tree was calculated with neighbor-joining methods from aligned, VP1 RNA sequences, and rooted with data from four enteroviruses (EV) of the EV-A, EV-B, and EV-C species, similar to ref. 10. The Major ("M," ICAM-1) and minor ("m," LDLR) receptor groups are indicated if determined experimentally. The RV-C receptor is unknown. Bootstrap values (percent of 200 replicates) are indicated at key nodes

those genotypes that are most similar to each other. When parsing new clinical isolates into their appropriate types, it is always important to remember that the larger the sequence that is compared, the more accurate the putative classification. Deep sequence alignments [8] covering ~1,000 VP1 datasets are especially valuable when discriminating, say, A25/A62, B52/B104, or other very similar types. As is characteristic of most such trees, no matter how they are calculated, this current depiction places the RV-A and RV-C more closely together on the tree than either is to the RV-B. Moreover, within the RV-A, a distinctive "clade D" (A8/A45) always branches off on its own from the other genotypes [10, 23].

At present, there are too few isolates within this clade to change the classification (RV-D?), but as the taxonomy system continues to evolve, that idea remains a possibility.

7 Recombination

In addition to the multitude of available VP1 sequences, completion of the full cohort of RV-A+B genome sequences [10] identified extensive evidence for historic recombination which, de facto, created several of the existing genotype clades (Fig. 3). A18, A34, A54,

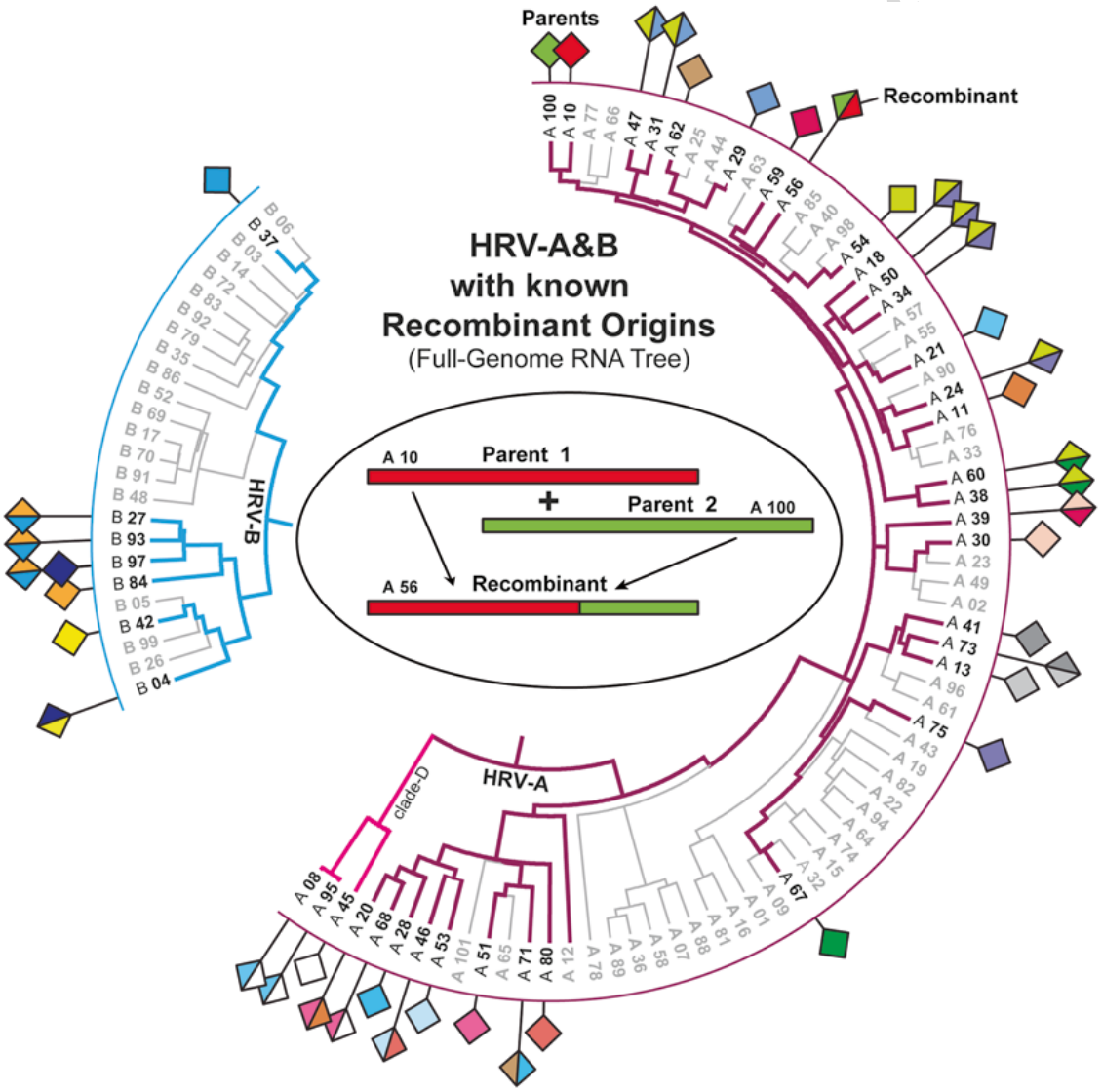

Fig. 3 Recombinant origins for many RV-A&B were uncovered by full-genome sequencing [10]. Parents (*solid boxes*) or progeny (*two-color boxes*) are founders of many extant clades. This illustration is modified from "Field's Virology" (2013), Ch 18, "Rhinoviruses," Wolters Kluwer, publishers

and A24 are independent derivatives of events between A54 and A75, for example. A54 is also a parent of A38 and A60. Similarly, B27, B93, and B97 have common parents in B37 and B42. Some of these viruses are promiscuous (!) and evidently, simultaneous infections must be a common event. Surprisingly though, none of these known recombinants have exchanged capsid regions. The most common trades include the 5′ UTR, primarily upstream of the IRES, or less frequently, fragments from P2–P3 regions. More recent, deep RV data from multiple field isolates has confirmed this idea, and now show clearly that the RV-A and RV-C frequently recombine between themselves, and when they do, they usually exchange not the expected capsid Nims, but 5′ UTR regions, and (often) their respective 2A protease genes [24, 25]. Comparative 2A work is under way to document why these particular recombinants are apparently favored. Possibly, divergent protease specificities may help these viruses regulate the overall cell response to infection.

Acknowledgements

This work was supported by NIH grant U19 AI104317. The authors thank Wolters Kluwer, publishers of Field's Virology (2013), for permission to include certain text and figures from Ch 18, "Rhinoviruses" by the same authors.

References

1. Stanway G, Hughes PJ, Mountford R et al (1984) The complete nucleotide sequence of a common cold virus: human rhinovirus 14. Nucleic Acids Res 12:7859–7875
2. Erickson JW, Frankenberger EA, Rossmann MG et al (1983) Crystallization of a common cold virus, human rhinovirus 14: isomorphism with poliovirus crystals. Proc Natl Acad Sci U S A 80:931–934
3. Kitamura N, Semler BL, Rothberg PG et al (1981) Primary structure, gene organization and polypeptide expression of poliovirus RNA. Nature 291:547–553
4. Racaniello VR, Baltimore D (1981) Molecular cloning of poliovirus cDNA and determination of the complete nucleotide sequence of the viral genome. Proc Natl Acad Sci U S A 78: 4887–4891
5. Carstens EB, Ball LA (2009) Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2008). Arch Virol 154:1181–1188
6. Savolainen C, Blomqvist S, Mulders MN et al (2002) Genetic clustering of all 102 human rhinovirus prototype strains: serotype 87 is close to human enterovirus 70. J Gen Virol 83:333–340
7. Palmenberg AC, Rathe J, Liggett S (2010) Analysis of the complete genome sequences of human rhinovirus. J Allergy Clin Immunol 125:1190–1199
8. McIntyre CL, Knowles NJ, Simmonds P (2013) Proposals for the classification of human rhinovirus species A, B and C into genotypically assigned types. J Gen Virol 94: 1791–1806
9. Rathe JA, Liu X, Tallon LJ et al (2010) Full-genome sequence and analysis of a novel human rhinovirus strain within a divergent HRV-A clade. Arch Virol 155:83–87
10. Palmenberg AC, Spiro D, Kuzmickas R et al (2009) Sequencing and analysis of all known human rhinovirus genomes reveals structure and evolution. Science 324:55–59
11. Bochkov YA, Palmenberg AC, Lee W-M et al (2011) Molecular modeling, organ culture and reverse genetics for a newly identified human rhinovirus C. Nat Med 17:627–632

12. Andries K, Dewindt B, Snoeks J et al (1990) Two groups of rhinoviruses revealed by a panel of antiviral compounds present sequence divergence and differential pathogenicity. J Virol 64:1117–1123
13. Dominguez SR, Briese T, Palacios G et al (2008) Multiplex MassTag PCR for respiratory pathogens in pediatric nasopharyngeal washes by conventional diagnostic testing shows a high prevalence of viruses belonging to a newly recognized rhinovirus clade. J Clin Virol 43: 219–222
14. Simmonds P, McIntyre C, Savolainen-Kopra C et al (2010) Proposals for the classification of human rhinovirus species C into genotypically assigned types. J Gen Virol 91:2409–2419
15. Ashraf S, Brockman-Schneider R, Bochkov YA et al (2013) Biological characteristics and propagation of human rhinovirus-C in differentiated sinus epithelial cells. Virology 436:143–149
16. Hao W, Bernard K, Patel N et al (2012) Infection and propagation of human rhinovirus C in human airway epithelial cells. J Virol 86:24–32
17. Basta HA, Sgro J-Y, Palmenberg AC (2013) Modeling of the human rhinovirus C capsid suggests a novel topology with insights on receptor preference and immunogenicity. Virology 448:176–184
18. Basta HA, Ashraf S, Bochkov YA et al (2013) Modeling of the human rhinovirus C capsid and antiviral drug resistance. Virology 448:82–90
19. Vlasak M, Roivainen M, Reithmayer M et al (2005) The minor receptor group of human rhinovirus (HRV) includes HRV23 and HRV25, but the presence of a lysine in the VP1 H1 loop is not sufficient for receptor binding. J Virol 79:7389–7395
20. Jackson RJ (1996) Initiation site selection mechanisms. In: Hershey JWB, Mathews MB, Sonenberg N (eds) Translational control. Cold Spring Harbor Laboratory Press, New York, NY, pp 71–112
21. Steil BP, Barton DJ (2009) Cis-active RNA elements (CREs) and picornavirus RNA replication. Virus Res 139:240–252
22. Cordey S, Gerlach D, Junier T et al (2008) The cis-acting replication elements define human enterovirus and rhinovirus species. RNA 14: 1568–1578
23. Ledford RM, Patel NR, Demenczuk TM et al (2004) VP1 sequencing of all human rhinovirus serotypes: insights into genus phylogeny and susceptibility to antiviral capsid-binding compounds. J Virol 78:3663–3674
24. Huang T, Wang W, Bessaud M et al (2009) Evidence of recombination and genetic diversity in human rhinoviruses in children with acute respiratory infection. PLoS One 4:e6355
25. McIntyre CL, McWilliam Leitch EC, Savolainen-Kopra C et al (2010) Analysis of genetic diversity and sites of recombination in human rhinovirus species C. J Virol 84:10297–10310