Impact of natural or synthetic singletons in the capsid of human bocavirus 1 on particle infectivity and immunoreactivity

Julia Fakhiri[1,2,*], Kai-Philipp Linse[1,2,3,*], Mario Mietzsch[4], Man Xu[5], Marc A. Schneider[6,7], Michael Meister[6,7], Oliver Schildgen[8], Paul Schnitzler[1], Maria Soderlund-Venermo[5], Mavis Agbandje-McKenna[4], Dirk Grimm[1,2,3,#]

[1]Heidelberg University Hospital, Dept. of Infectious Diseases/Virology, 69120 Heidelberg, Germany

[2]BioQuant Center, University of Heidelberg, 69120 Heidelberg, Germany

[3]German Center for Infection Research (DZIF), partner site Heidelberg, Heidelberg, Germany

[4]Department of Biochemistry and Molecular Biology, Center for Structural Biology, McKnight Brain Institute, University of Florida, Gainesville, FL, 32610, USA

[5]Department of Virology, University of Helsinki, Helsinki, 00290, Finland

[6]Translational Research Unit, Thoraxklinik at Heidelberg University Hospital, 69126 Heidelberg, Germany

[7]Translational Lung Research Center Heidelberg (TLRC), German Center for Lung Research (DZL), 69120 Heidelberg, Germany

[8]Institute for Pathology, Kliniken der Stadt Köln gGmbH, Hospital of the Private University Witten/Herdecke, 51067 Cologne, Germany

*The first two authors contributed equally to the underlying research. Author order was determined by the most substantial contribution to the article draft.

[#]Correspondence should be addressed to:

D.G. (dirk.grimm@bioquant.uni-heidelberg.de), Heidelberg University Hospital, BioQuant BQ0030, Im Neuenheimer Feld 267, 69120 Heidelberg, Germany, Phone +49-6221-5451339, Fax +49-6221-5451481

Keywords: Bocavirus, BoV, capsid, mutations

Short title: Natural or synthetic variations in human bocavirus 1 capsids

ABSTRACT

Human bocavirus 1 (HBoV1) is a parvovirus that gathers increasing attention due to its pleiotropic role as a pathogen and emerging vector for human gene therapy. Curiously, albeit a large variety of HBoV1 capsid variants has been isolated from human samples, only one has been studied as a gene transfer vector to date. Here, we analyzed a cohort of HBoV1-positive samples and managed to PCR-amplify and sequence 29 distinct HBoV1 capsid variants. These differed from the originally reported HBoV1 reference strain in 32 nucleotides or four amino acids, including a frequent change of threonine to serine at position 590. Interestingly, this T590S mutation was associated with lower viral loads in infected patients. Analysis of the time course of infection in two patients for up to 15 weeks revealed a gradual accumulation of T590S, concurrent with drops in viral loads. Surprisingly, in a recombinant vector context, T590S was beneficial and significantly increased titers as compared to T590 variants but had no major impact on their transduction ability or immunoreactivity. Additional targeted mutations in the HBoV1 capsid identified several residues that are critical for transduction, capsid assembly or DNA packaging. Our new findings on the phylogeny, infectivity and immunoreactivity of HBoV1 capsid variants improve our understanding of bocaviral biology and suggest strategies to enhance HBoV1 gene transfer vectors.

IMPORTANCE

The family of *Parvoviridae* comprises a wide variety of members that exhibit a unique biology and that are concurrently highly interesting as a scaffold for the development of human gene therapy vectors. A most notable example is human bocavirus 1 (HBoV1), which we and others have recently harnessed to cross-package and deliver recombinant genomes derived from another parvovirus, the adeno-associated virus (AAV). Here, we expanded the repertoire of known HBoV1 variants by cloning 29 distinct HBoV1 capsid sequences from primary human samples and by analyzing their properties as AAV/HBoV1 gene transfer vectors. This led to our

discovery of a mutational hot spot at HBoV1 capsid position 590 that has accumulated in two patients during natural infection and that lowers viral loads but increases vector yields. Thereby, our study expands our current understanding of HBoV1 biology in infected human subjects and concomitantly provides avenues to improve AAV/HBoV1 gene transfer vectors.

INTRODUCTION

Parvoviruses are small, non-enveloped viruses that package a single-stranded (ss)DNA genome of ~5-6 kb. This genome contains two main open reading frames (ORF) that comprise the non-structural (*ns*) and the capsid genes (*cap* or *vp*). Bocaviruses (BoV), which belong to the autonomous parvoviruses, harbor an additional unique ORF that encodes the nucleophosphoprotein 1 (NP1). Intriguingly, a series of recent reports implies that ssDNA viruses including parvoviruses may evolve more rapidly than anticipated, evidenced by measurements of high nucleotide substitution rates (10^{-3} to 10^{-6} substitutions/site/year) that are comparable to the rate of RNA virus counterparts (1, 2). For example, high rates of 1×10^{-4} substitutions/site/year were inferred for some autonomous parvoviruses such as the carnivore parvoviruses (3, 4), human parvovirus B19 (5) and porcine parvovirus (6). Moreover, several studies have estimated a similar rate of both, structural and non-structural parvovirus gene evolution. For example, in human bocavirus 1 (HBoV1), the *np1* ORF shows the highest rate of mutations among the *ns* genes that correlates with significant changes in viral titer (7). This could be a result of the multiple roles of NP1 in viral replication (8) and capsid protein expression (9), in addition to its immunomodulatory function (10). Changes in the C-terminal part of the *ns1* ORF were shown to directly influence the role of NP1 in viral genome replication (8), which underlines the importance of a tightly regulated co-evolution of the non-structural genes. Viral titers are also influenced by mutations in the structural *vp* gene, especially in the VP1u region that is vital for the infectivity of the virus (7, 11).

The parvoviral capsid is an important determinant of virus tropism, host range and reaction to the immune system. It was shown that even small amino acid (aa) changes in the *vp* ORF can largely alter virus-cell interactions including cell-type preference. For instance, adeno-associated virus type 1 (AAV1) and AAV6 share 99% aa identity but exhibit a distinct polarity bias in primary airway epithelia (pHAE) (12) as well as different abilities to transduce human and

mouse hematopoietic stem cells (13). Even aa changes in the VP proteins within one AAV serotype can drastically alter viral features, as exemplified by the two sub-types of AAV3, AAV3a and AAV3b, which differ by only 6 aa but display distinct affinities to heparin (14, 15). Another example is the pair of rodent parvoviruses MVMp and MVMi (minute virus of mice) that share 97% sequence identity but differ in their *in vitro* (16) and *in vivo* cell tropisms (17). Similarly, the CPV2 and CPV-2a strains (canine parvovirus) differ only in four amino acids, which, however, lead to the extended feline tropism of CPV-2a (18).

This rich repertoire of parvoviruses with distinct properties has drawn enormous interest to their potential use as gene transfer vectors in cancer and gene therapy. In the recombinant genomes of these vectors, either all viral sequences or parts thereof are replaced by transgenes of interest. To package the recombinant genomes, the missing viral elements have to be supplied *in trans* during vector production (19-22). In particular, AAV has emerged as a promising viral vector following its extensive study for over five decades. A major reason for its popularity is the feasibility to pseudotype recombinant AAV2 genomes with AAV capsids from other natural serotypes or synthetic variants, which allows for transgene delivery to different target organs (23). Another recent example of interesting parvoviral vectors are chimeric rAAV/BoV vectors in which a rAAV genome is pseudotyped with the capsid proteins from primate BoV (22, 24). One of these BoV variants, HBoV1, has a unique tropism for the airways and has been utilized as gene delivery vehicle *in vitro* and *in vivo* (25). Curiously, despite the wealth of HBoV1 *vp* sequences that were isolated in many areas of the world, only one particular HBoV1 variant (GenBank: GQ925675) has so far been used as viral vector.

Accordingly, we aimed to study whether and how naturally occurring variations in the HBoV1 *vp* sequence affect properties of the virus. To this end, we constructed a battery of new pseudotyped viral vectors from HBoV1 variants that were *de novo* isolated from patient samples or described in previous studies. Next, we packaged a *Gaussia* luciferase (Gluc)-encoding rAAV

genome into each of the various HBoV1 capsids. This allowed us to characterize the effects of naturally occurring single-point mutations ("singletons") on viral DNA packaging, transduction and immunological reactivity with anti-HBoV1 antibodies. The results of our work reveal interesting, previously underappreciated aspects of HBoV1 biology and have important implications for the choice and use of rAAV/HBoV1 vectors in future gene therapy applications.

RESULTS

Analysis of capsid DNA and protein sequence diversity in naturally occurring HBoV1 variants

To analyze the natural sequence diversity of HBoV1 capsid genes and proteins, we collected a total of 64 samples from patients at the University Hospitals in Heidelberg and Cologne (both Germany) who had previously been tested positive for HBoV1. These samples comprised tracheal secretions, aspirates, pharyngeal washes, sputum as well as bronchioalveolar lavages collected from children and adults in the years 2014 to 2016. From these 64 samples, we were able to PCR-amplify and sequence the entire HBoV1 capsid-coding region (2016 bp) in 29 samples, *i.e.*, in 45.3% of all samples (exemplified in Fig. 1A; full DNA sequences are shown in the Supplementary Dataset). Typically, failure to amplify or fully sequence the capsid gene correlated with low viral titers in the original sample below 1×10^6 viral genomes (vg) per ml.

Interestingly, alignment of these 29 capsid DNA sequences with the HBoV1 reference sequence that was first reported by Allander *et al*. in 2005 (GenBank: DQ000495; note that this is not the sequence that is utilized in current HBoV1 vectors) showed differences in 32 nucleotide positions. As summarized in Supplementary Table 1 and Fig. 1B, the newly analyzed sequences carry between four and 19 mismatches with this reference sequence, corresponding to an average of 12.7 nucleotide differences per variant (368 variations in total, divided by 29 samples). Accordingly, their overall DNA sequence identity to DQ000495 is 99.1 to 99.8%, or

99.4% on average. Moreover, we noted that the mutations cluster in 14 of the 32 positions, namely, 441, 445, 714, 984, 1140, 1168, 1170, 1176, 1188, 1308, 1758, 1767, 1768 and 1785 (numbers are nucleotide positions in the HBoV1 *vp1* capsid gene), where more than half of the 29 sequences differed from the reference.

On the protein level, these point mutations translated into substitutions at four positions in the 671 aa HBoV1 VP1 capsid protein, as compared to the DQ000495 reference (Fig. 1C). Each of the 29 sequences differed in one to three positions, namely, 68, 149, 474 or 590 (numbers are aa positions in VP1), corresponding to identities of 99.6 to 99.9% (99.8% on average). Hence, the majority of nucleotide exchanges was silent on the VP1 protein level. Again, we observed a clustering of the mutations, most notable at aa position 149 (red in Fig. 1B-C) where 100% of all analyzed sequences carry a threonine instead of the alanine reported by Allander and colleagues (26). The second striking difference is seen at position 590 (orange in Fig. 1B-C) where we detected a serine instead of a threonine in 15 out of the 29 sequences. In addition, sample V1541706 carries asparagine instead of aspartate at position 68, and samples VK11443 and VK12783 have a replacement of serine with asparagine at position 474.

Next, we performed a phylogenetic analysis of all 29 DNA sequences together with 21 publicly available HBoV1 sequences from 14 countries and four continents. Additionally, we included the related viruses HBoV2, 3 and 4 (GenBank: NC012042, NC012564 and NC012729, respectively) as well as the non-primate bocaviruses, canine minute virus (NC004442) and bovine parvovirus 1 (NC001540), which we collectively defined as outgroup. This analysis confirmed that all 29 HBoV1 *vp* sequences cluster together with the 21 public sequences and are clearly distinct from the outgroup (bootstrap value of 99) (Fig. 1D).

Correlation of HBoV1 capsid sequence diversity and virus infectivity

In addition to the primary sequence, we analyzed the viral load in the original set of 64 HBoV1-positive patient samples that we had collected. By using quantitative (q)PCR, we succeeded at

177 measuring the viral load for 39 samples, comprising the 29 for which we had previously
178 obtained the full capsid sequence (see above) as well as 10 others where this information was
179 lacking. The values obtained ranged from 1.82×10^{3} to 6.13×10^{10} vg/ml, with a median of
180 3.91×10^{6} vg/ml (Fig. 2A). Prior work had defined a cutoff of 1×10^{6} vg/ml above which symptoms
181 of HBoV1 infection manifested in outpatients or inpatients (27). Accordingly, we can classify 15
182 of the 39 samples as having a viral load below this cutoff, while the other 24 are above.

183 As described above, HBoV1 VP1 aa position 590 is a hotspot for a change from threonine in
184 DQ000495 to serine. In line with this, we noted that roughly half of the samples whose viral load
185 we had determined carry a serine at this position. We thus correlated the occurrence of either
186 threonine or serine with viral load for 31 of the 39 samples. This subset was selected based on
187 the criteria that we could read over 90% of the complete capsid sequence and that we could
188 unanimously identify position 590 as threonine or serine. Remarkably, this analysis showed that
189 HBoV1 variants carrying a serine (T590S, n=17) have a ~18-fold lower viral load than those with
190 the originally reported threonine (T590, n=14) (Fig. 2B). In detail, the average viral load for the
191 T590S variant was 5.55×10^{8} vg/ml, in contrast to 1.02×10^{10} vg/ml for variant T590.

192 Among the analyzed samples, several had been collected from the same patient(s) at various
193 time points, which allowed us to study dynamic changes in the HBoV1 capsid sequence and
194 measure alterations in the viral load during the course of an infection. The results are depicted
195 in Fig. 2C for patient A (four time points over a period of 15 weeks) and Fig. 2D for patient B
196 (five time points over a period of three weeks). Changes at the nucleotide level over time were
197 observed at positions 804, 873, 1140, 1168, 1170, 1701, 1767, 1768 and 1785. Two examples
198 that were identified in both patients and that are illustrated in Fig. 2C-D are a gradual change of
199 thymine to cytosine at position 1140, or a change of guanine to adenine at position 1170.
200 Intriguingly, while these two mutations were silent on the protein level, we found that the gradual
201 replacement in both patients over time of the nucleotide sequence 5'-AA-3' at position

1767/1768 by 5'-CT-3' resulted in an exchange of threonine at aa position 590 to serine, *i.e.*, the same mutation that we had already observed and highlighted before. Congruent with the data in Fig. 2B, we measured a consistent drop in viral loads over time that concurred with the accumulation of the T590S mutation. This is evidenced by a reduction of viral loads in patient A, from 3.11×10^{9} vg/ml at the earliest time point to 7.46×10^{3} vg/ml at the latest (week 15) (Fig. 2C, E). Likewise, the viral loads in patient B dropped from 1.57×10^{10} vg/ml to 3.91×10^{6} vg/ml over the course of three weeks (Fig. 2D-E).

Dissection of the impact of changes in the vp ORF using recombinant HBoV1 vectors

To further study the impact of the observed natural or, introduced later in this work, synthetic point mutations in the HBoV1 capsid, we harnessed a streamlined system for production of recombinant HBoV1 gene transfer vectors that we have established recently (24). Its hallmark is the ability to package rAAV vector genomes encoding a reporter into HBoV1 capsids, by triple-transfecting HEK293T cells with three plasmids expressing all necessary factors including the HBoV1 capsid proteins, and by purifying the resulting vector particles via iodixanol density gradient centrifugation. Specifically, here, we used an AAV vector genome expressing *Gaussia* luciferase (Gluc), which is a secreted protein that is easily detected in the supernatant of cultured cells. For the latter, we used primary human airway epithelia (pHAE) based on findings by others and us that these are highly susceptible to HBoV1 transduction (24, 28).

In total, we studied 18 HBoV1 capsid variants using this system, which comprised seven from our collection that had high viral loads between 3.4×10^{9} and 6.1×10^{10} vg/ml and that exemplified the roughly equal distribution of either serine or threonine at position 590 (Fig. 3A-B; samples V1500812, V1602382, V1502611, V1512195, V1541706, V1613007 and VK11443). As compared to DQ000495, they differ in five to 18 nucleotides, or one to three amino acids. As two references, we included DQ000495 as well as GQ925675, *i.e.*, the HBoV1 capsid variant that has been used in all recombinant HBoV1 vector preparations reported to date (24, 25, 29).

Furthermore, we selected nine variants that were described by Principi *et al*. in 2015 and that differ from DQ000495 in one to six amino acids (27). As the original sequences were not available to us as molecular clones, we recapitulated only non-synonymous substitutions that cause aa changes by successive overlap-extension PCR using DQ000495 as template (Supplementary Table 2 and 3). Of these nine capsids, seven were associated with high viral loads between 5×10^8 and 7.5×10^9 vg/ml in the study by Principi and colleagues (27). The remaining two variants from this study, named KPLGr1 and KPLGr2 here, represent two groups of patient-derived HBoV1 capsid sequences that were identical within each group and that were associated with low viral loads of 4.5×10^4 or 6.3×10^4 vg/ml, respectively. They were interesting since KPLGr1 carries the T590S mutation, while KPLGr2 has T590, thus representing the distribution of these two residues and making these two variants useful as additional controls. Altogether, this set of 18 capsid variants was composed of six carrying T590 and 12 with T590S.

Titration of recombinant vectors based on HBoV1 capsid mutants

To produce vectors based on the aforementioned 18 HBoV1 capsid variants, we performed three independent runs using three 15 cm^2 dishes each, following our triple-transfection and iodixanol purification scheme (24). Titration by quantitative (q)PCR revealed titers for the 54 (18×3) individual stocks between 6.51×10^8 to 6.61×10^{11} vg/ml (data not shown), with an average of 9.65×10^{10} vg/ml. The majority of capsid variants yielded titers in a range of 2×10^{10} to 2×10^{11} vg/ml, with the notable exception of KPLMI-30 and KPLMI-3503 whose average titers were 1.13×10^{10} or 2.15×10^9 vg/ml, respectively. In contrast, KPLGr1 and KPLGr2 that had been associated with low viral loads in patients before (27) consistently produced well in our hands.

We subsequently pooled all three independent preparations per capsid variant and further purified and concentrated them using Amicon Ultra-15 filter units. Final titers per nine 15 cm^2 dishes per capsid variant were between 4.11×10^{10} to 2.4×10^{11} vg/ml (average of 1.1×10^{11} vg/ml,

Fig. 3C), again with the exception of KPLMI-30 and KPLMI-3503, which yielded 1.7×10^9 or 4.67×10^8 vg/ml, respectively. Titers for the 16 capsids that produced well were determined twice since these stocks were used in later transduction experiments (see below). On average, concentration using the Amicon Ultra-15 filters had resulted in a vector particle recovery of 42%.

Interestingly, we noted a highly significant difference between the HBoV1 capsids depending on the presence of a threonine or serine at position 590. While the average titer of the six T590 variants was 6.66×10^{10} vg/ml, it was 1.59×10^{11} vg/ml for all capsids with the T590S mutation, *i.e.*, 2.39-fold higher (p=0.0003, Fig. 3D). Of note, for this calculation, we excluded the two low producers KPLMI-30 and KPLMI-3503 as they were obvious outliers. To facilitate the comparison of candidates and to highlight their differences throughout this work, five candidates were randomly chosen (including the two references DQ000495 and GQ925675) and colored.

Comparison of the transduction efficiency of all HBoV1 capsid variants in pHAE

To measure and compare the transduction efficiency of the 16 HBoV1 capsid variants that produced well, we used pHAE from five different donors (labeled 171469, 171476, 171834, 171905, 171975 in the following). Each capsid was tested twice per donor (n=10 for each variant) and compared to two negative controls (two wells of untransduced cells). Per transwell, we applied 3×10^8 vg, which corresponds to a multiplicity of infection (MOI) of 600 based on a count of roughly 5×10^5 cells per transwell. As the Gluc reporter that was encoded by all vectors is secreted from the cells, we could collect cell culture supernatant at three successive time points (day 6, 9 or 12 post-transduction) and thus analyze the kinetics of transduction.

Two general observations were, firstly, that the overall transduction efficiencies varied substantially between the donors, as best illustrated by the up to 10-fold differences in luciferase light units between donors #171476 and #171905 (Fig. 4A, also shown in Fig. 4D). Secondly, we noted donor-dependent relative differences in the performance of the various capsids. This is exemplified by variant KPLMI499 (yellow bars) that gave robust luciferase expression in

donors 171905 and 171975, but was less efficient in the other donors, 171469, 171476 and 171834 (Fig. 4A, also shown in Fig. 4D). For these two reasons, the raw data for all five donors are depicted individually in Fig. 4A-B.

Despite the donor variability, it was noteworthy that the two published variants, DQ000495 and GQ925675, were consistently among the top performers in all five donors. Furthermore, the HBoV1 reference sequence DQ000495 (green bars in Fig. 4A) that was originally reported in 2005 significantly outperformed GQ925675 (orange bars) in the pHAE derived from donor 171467 (Fig. 4A, also shown in Fig. 4D). This is intriguing considering that, to our best knowledge, the less efficient GQ925675 forms the basis for all recombinant AAV/HBoV1 vectors that are currently in use by us and others. Next to these two historic controls, also capsid variant V1512195 (blue bars, Fig. 4A and 4D) performed remarkably well in most cells, albeit we again noted some degree of donor dependency (*e.g.*, in donor 171975, where it ranked second to last). In contrast, capsid variants V1613007, KPLMI-246, KPLMI-253, KPLMI-311, V1500812 (all in gray) and V1502611 (pink bars) were frequently among the least efficient of all capsid variants, also with a few donor-specific exceptions (*e.g.*, the good performance of capsids V1500812 and V1502611 in donor 171469, Fig. 4A). Besides, we measured a steady increase in luciferase transgene expression for all 16 capsids and all five donors from day 6 to 12 post-transduction (data not shown), congruent with prior notions of AAV/HBoV1 kinetics in this cell culture model (24).

Next, we compared transduction efficiencies at day 12 based on the presence of T590 (six variants) versus T590S (twelve variants) (Fig. 4B). We found no significant differences between the two groups in their transduction abilities, but observed trends in donors 171834, 171469 and 171476, where variants with T590 had a slight advantage. *Vice versa*, the S590 variant tended to perform better in donors 171905 and 171975. Notably, as exemplified in Fig. 4C, we concurrently observed differences in the cellular composition of the pHAE cultures that might

have influenced the transduction efficiency and specificity of the HBoV1 variants. In particular, sample 171834 showed a much higher proportion of Mucin-positive goblet cells versus Tubulin-positive ciliated cells, as compared to samples 171905 and 171975 where ciliated cells predominated. An overview of the differential efficiency of selected variants in each of the five pHAE cultures is shown in Fig. 4D.

Finally, we separately tested the two HBoV1 variants that had yielded the lowest titers during vector production (Fig. 3C), KPLMI-30 and KPLMI-3503, in direct comparison to DQ000495 (Fig. 4E). For this, we used pHAE derived from two donors, one from the group of five that we had also used above (171476), while the other was used only in this experiment (171427). Moreover, owing to the limiting vector yields, we had to reduce the MOI from 600 to 200, corresponding to 1×10^8 vg per pHAE filter. Consistently, we found that these two HBoV1 variants mediated very low transduction that was 10- to 100-fold below the DQ000495 reference at all time points (Fig. 4E). Together, this implies that the multiple nucleotide and/or aa differences in these two capsids as compared to the HBoV1 reference (Fig. 3A-B and Supplementary Table 3) impact both, the ability to be produced as recombinant vector and to mediate robust transgene expression, at least in pHAE.

Packaging and transduction efficiency of HBoV1 tyrosine mutants

An alternative approach to identify HBoV capsids with improved properties that complements screening of natural HBoV variants is rational engineering of the viral capsid. To this end, the modification of surface-exposed tyrosine residues in the viral capsid is especially promising as these residues play important roles in assembly, ubiquitination and degradation, as well as in transcription and transduction of parvoviruses (30-36). Accordingly, we mutated six different tyrosine residues in the VP1 capsid protein to phenylalanine that we had originally predicted by structural modeling to be located on the HBoV1 capsid surface (Fig. 5A), namely, Y276, Y403, Y484, Y523, Y595 and Y657 (all VP1 aa numbering). To study the effect of each mutation on

particle assembly and transduction, we packaged two transgenes, *yfp* (yellow fluorescent protein) or Gluc, into the different capsid variants. All mutants yielded largely comparable vector amounts (data not shown), implying that none of the studied tyrosine mutations affects vector production.

Next, to assess the ability of these mutants to transduce pHAE, they were added to the apical surface of transwells at a MOI of 2×10^4. This higher MOI (versus 600 before) was used as we expected a lower infectivity for at least some of the tyrosine mutants, and as we wanted to be able to measure all transduction efficiencies in the same experiment and under identical conditions. Transductions were performed in the presence or absence of proteasome inhibitors, to study whether the Y-to-F mutations had circumvented capsid ubiquitination and thereby alleviated HBoV1's dependency on proteasome inhibition. Interestingly, we found that transduction of two of the six mutants was impaired at all time points (day 3 to 14, Fig. 5B), namely, Y484 (~6.8-fold reduction compared to wild-type HBoV1) and Y595 (~4.5-fold reduction). In contrast, the other four mutants were indistinguishable from the wild-type control. Moreover, the Y484F mutation had actually increased the dependency of the cognate capsid on proteasome inhibitors, because this capsid was inert in their absence. In this respect, it differs from the Y595F mutant that showed a similarly reduced potency in the presence of proteasome inhibitors, but that, unlike the Y484F variant, remained active also without these inhibitors (albeit at reduced efficiency, akin to all other variants).

Of note, after the completion of these experiments, the capsid structure of HBoV1 has been determined by cryo electron microscopy (37). The HBoV1 capsid structure revealed that of the six tyrosines that we studied here, only three are indeed located on the capsid surface, *i.e.*, Y276, Y403 and Y595 (Fig. 5C). In addition, the hydroxyl group of Y276 is inaccessible for potential phosphorylation. In contrast, Y484, Y523, and Y657 are not exposed on the capsid surface. Most notably, mutation of Y484, which is located on the inside of the capsid, had

yielded the strongest phenotype, implying a mode of action that differs from the anticipated phosphorylation and ubiquitination (see Discussion).

Differential inhibition of HBoV1 capsid variants by human immunoglobulins

To determine the effect of anti-HBoV1 antibodies on the functionality of the different HBoV1 capsid variants, we first performed an Enzyme Immuno Assay (EIA; Fig. 6A). For this purpose, microtiter plate wells were coated with the different capsid variants (antigens) at four different dilutions (only 8- and 16-fold dilutions are shown, as they resulted in a linear signal). Next, the reactivity of a human serum pool positive for HBoV1 antibodies was measured. We detected small differences typically within 1.3 OD between the variants in their ability to bind HBoV1-specific antibodies, except for GQ925675, which showed a 2- to 4-fold reduced binding. Surprisingly, for reasons as-of-yet-unknown, capsid variant VK11443 that differs only in one nucleotide from GQ925675 and has an identical aa composition did not show the same reduction in antibody binding. Also notable are the two capsid variants KPLMI-30 and KPLMI-3503, which resulted in an OD of 3.10 and 0.10, respectively (at 8-fold dilution). Notably, the low viral titers of both KPLMI-30 and KPLM-3503 (close to the background) limit the ability of qPCR-based analysis that requires encapsidated genomes to reliably estimate the amount of viral capsids. Thus, the slightly higher OD value measured for KPLMI-30 as compared to the other variants (differences between 0.72 and 2.54 ODs) might result from the technical challenge to quantify the proper amount of virus solution used in the EIA. The differential reactivity of these two variants with human antibodies is interesting in view of their strikingly reduced viral titers and transduction ability in pHAE (Fig. 3C and 4E, respectively). Based on the titer reduction, we speculated that these variants might have a defect either in assembly or genome packaging. To resolve these possibilities, we studied viral VP expression by Western blot analysis (Fig. 6B). All variants expressed VP1/VP2/VP3 proteins (without evidence for additional protein species) in the expected 1:1:10 stoichiometry, which shows that capsid protein expression is not limiting.

Thus, the absence of signal in the EIA for KPLMI-3503 viral particles implies that this variant might have an assembly defect. By contrast, the high signal for KPLMI-30 in combination with the low measured viral titers hints at an impairment in genome packaging. Notably, we do not rule out the possibility that subtle differences in VP protein composition not detected by Western blotting may have contributed to the variations in transduction or packaging ability, a theory that could be studied with other, more sensitive methods. Besides, we did not detect variations in the length of the encapsidated vector DNAs (data not shown), implying that genome integrity is not responsible for the differences in titer or transduction.

Concerning the T590S variation, the EIA assay did not reveal obvious differences between the T and S groups in their binding to human anti-HBoV1 antibodies, which implies that aa position 590 does not confer a differential reactivity to HBoV1 capsid antibodies in human sera.

It is known for AAV vectors that antibody binding does not always result in neutralization of virus transduction (38). To test whether this applies to HBoV1 as well and to compare our different variants in a functional assay (transduction ability), we performed *in vitro* neutralization assays using commercially available, pooled human immunoglobulins (IVIg). This mix of IgG antibodies from healthy individuals has previously been shown to potently reduce the activity of the standard GQ925675 vector, in line with its large seroprevalence in the human population (24).

Due to the limited availability of pHAE, we selected seven HBoV1 variants for the assay with equal T/S590 distribution, including the GQ925675 and DQ000495 reference strains. To this end, we mixed 5×10^9 vg per capsid variant (corresponding to a MOI of 1×10^4 vg per well) with six different IVIg concentrations and incubated these mixtures for 1 h at 37°C, before adding them to pHAE from two different donors (Fig. 6C). These IVIg concentrations were shown in pilot studies to result in a reduction or complete inhibition of transduction with GQ925675 (data not shown). Comparison of Gluc expression from the different HBoV1 variants at day 5 and 10 showed no obvious differences between the two groups (T or S at position 590) in their

transduction abilities in the presence or absence of IVIg, which supports the previous notion that this variation does not confer increased resistance to neutralizing antibodies. Still, contradicting this trend is the KPLGr2 variant, which differs only in aa position 590 from KPLGr1 but shows a higher resistance to IVIg, despite its stronger binding to antibodies in the EIA assay (Fig. 6A).

Analysis of evolutionary selection pressures on HBoV1

Intrigued by our finding that the natural T/S590 variation has a profound impact on viral titer, we asked whether this site was subject to a positive selection pressure. To address this question, we used several methodologies comprising MEME, SLAC, FEL and FUBAR (see Methods). Indeed, we found one site (using MEME) under "positive selection" (moderately significant with p=0.51), namely, the abovementioned aa 590 (Fig. 7A). However, a positive selection of T/S590 was not supported by the other methods used, which showed a neutral selection pressure at this position. Thus, albeit it is implied by the MEME results, we cannot firmly conclude that the observed substitution at this site has an impact on intra-species transmission and adaptation of HBoV1.

Finally, we analyzed the conservation of the aa 590 residue in the VR-VIIIB by comparing the aa composition of this region to other primate BoV (Fig. 7B). Notably, the Gorilla bocavirus (GBoV) described by Kapoor *et al*. (39) is genetically most closely related to HBoV1 (on both, the nucleotide and amino acid level). Accordingly, the VR-VIIIB of HBoV1 is also most homologous to the one in GBoV, with both carrying a threonine at position 590, in contrast to an asparagine in HBoV2-4. Also interesting in this context is the profound ability of GBoV to transduce human airway epithelial cells that we have reported recently (24). Together, this may hint towards an inter-species transmission of HBoV1/GBoV.

DISCUSSION

The present work was fueled by a string of recent publications showcasing the great potential of recombinant gene transfer vectors derived by packaging of AAV vector genomes into HBoV1 capsids (22, 24, 25, 40). Most recently, these reports have inspired us to engineer similar vectors based on alternative BoV, *i.e.*, HBoV2-4 and Gorilla Bocavirus (GBoV), leading to our discovery of their favorable properties for gene transfer into various primary human cells (24). At the same time, this work by others and us has revealed a series of gaps in our current understanding of fundamental BoV biology, whose resolution will not only benefit our knowledge of the bocaviral life cycle but also promises to foster the development of next-generation BoV gene therapy vectors.

In the first part of this work, we aimed to study the natural variation in HBoV1 isolates. Specifically, we focused our attention on the capsid (*vp*) ORF, which is the determinant of virus tropism and the subject of extensive research in viral vector development. Therefore, we amplified and sequenced 29 full-length *vp* sequences from patient samples collected in Heidelberg and Cologne. Despite the high degree of sequence conservation among the HBoV1 isolates, which is in line with previous reports (41), we detected an interesting hotspot for variation in VP1 at aa position 590 in around 50% of analyzed patient samples. This change results from a conversion of two nucleotides (5'-AA-3') at positions 1767-1768 in *vp1* to 5'-CT-3'. Interestingly, a previous study by Principi and co-workers, who analyzed samples from Milan (Italy), also showed this high prevalence of the T590S variation (27). The additional collection of several samples from the same individual over time allowed us to also follow the course of infection in two patients. Surprisingly, we found dynamic changes at the 590 aa position, starting with a clear abundance of threonine that was gradually replaced by a serine (see Fig. 2C-D). The emergence and persistence of the T590S switch could be explained by different events: (i) a *de-novo* change of two adjacent nucleotides during virus evolution, which is, however,

unexpected in view of the estimated mutation rate of primate bocaviruses (9×10^{-4} mutations/site/year) (42, 43), or (ii), as previously proposed by Martin and colleagues (44), a secondary infection with another strain, which might have a replication or immunological advantage and thus dominated over time. Another, rather rare scenario also reported in the above-mentioned study is the co-infection with two strains during the same primary event that have different kinetics or immunoreactivity. To conclusively identify the proper scenario, deep sequencing analysis of samples from different time points has to be performed to detect even minor quantities of specific variants. Interestingly, as opposed to the study by Principi *et al.*, the T590S change was concomitant with a decrease in overall virus load (27). However, these varying study outcomes may have resulted from different time and end points in sample collection. At this point, it thus remains equally possible that the observed decrease at the endpoint of sample collection truly reflects a biological property of T590S or that a secondary infection has happened in these patients. Consequently, a firm conclusion regarding the persistence or spread of 590S *versus* T590 variants and on their possible positive selection cannot be drawn until more samples have been collected and analyzed during symptomatic and asymptomatic periods.

In addition to the prominent T590S mutation, we detected additional aa changes resulting from single-nucleotide polymorphisms (SNP): D to N (aa position 86), S to N (aa position 474) and a dominant A-to-T mutation (aa position 149) in all variants (as compared to the DQ000495 reference strain). To experimentally unravel the role of these naturally occurring SNPs or singletons in the *vp* ORF on the producibility and transduction ability of HBoV1, independent of the exact collection time points and sample types, we made use of a previously established recombinant vector system in which rAAV genomes are packaged into BoV capsids (22, 24). To this end, we selected variants from patient samples with approximately equal T/S distribution (Fig. 3A-B) and packaged a rAAV-Gluc vector into each capsid. The high aa identity of the

HBoV1 variants in our study prompted us to include additional synthetic mutants into our screen that either (i) recapitulate several natural HBoV1 variants reported by Principi *et al.* (27) or (ii) were rationally designed, based on a previously published HBoV1 structure by Gurda *et al.* (45). During packaging of the rAAV-Gluc genomes into the different HBoV1 *vp* variants, we gained significantly higher titers with the ones harboring a serine at aa position 590 (see Fig. 3C-D), which was surprising as it is diametrically opposed to the decrease observed in the patient material. Hence, these findings allow us to conclude that the decrease in 590S viral load in patient material does not result from a reduced ability to produce viral progeny (Fig. 2B). Moreover, the consistently higher viral titers obtained using 590S variants indicate a direct effect of the variation on capsid assembly and/or genome packaging. Interesting in this context and supporting this hypothesis is that T/S 590 lies in the capsid VR-VIIIB (also called “HI loop”, Fig. 8A-B). This region belongs to the surface-exposed, hypervariable regions and was linked to particle assembly and genome packaging in AAV (46), hinting at a similar function in the BoV context. In addition to its role in particle assembly, the VR-VIIIB is crucial for the externalization of the VP1u region during endosomal escape and thus contributes to particle infectivity (46). Accordingly, to test whether the *vp* variations studied in this work affected the transduction ability of HBoV1, pHAE grown on transwells were transduced from the apical side with an equal number of viral particles. Transgene expression was followed over time by measuring the secreted reporter Gluc in the medium. Intriguingly, nearly all tested variants displayed high transduction abilities (except for KPLMI-3503, KPLMI-30, HBoV1 Y595F and HBoV1 Y484F). Moreover, their activity was dependent on the transwell composition and thus varied among pHAE cultures, consistent with our prior observations with various primate BoV vectors (24). When we compared transduction abilities in the context of the T590S variation, we did not observe significant differences between the two groups. This is in line with a previous study using recombinant AAVs, which showed that swapping of the complete HI loop between serotypes with high sequence identity affects their producibility but not transduction ability (46).

Hence, we concluded that the T/S variation affects the virus titers by playing a role in particle assembly and/or genome packaging but does not determine infectivity. By contrast, mutation of the tyrosine residue to phenylalanine at aa position 595, which flanks the VR-VIIIB in the HBoV1 capsid, resulted in a 4.5-fold reduction in virus infectivity. Importantly, this residue is highly conserved among primate BoVs except for HBoV4, where it can be naturally replaced by a phenylalanine (*e.g.*, HBoV4 strain FJ973561).

An even stronger phenotype (6.8-fold reduction in infectivity) was obtained when Y484, located between VR-VI and VR-VII, was mutated to a phenylalanine. This residue lies inside the capsid and is also highly conserved among primate BoVs, indicating an important role in the BoV infection pathway, most likely perhaps for capsid assembly. Further analysis of Y484 in the wild-type HBoV1 capsid structure showed a potential hydrogen bond of the side chain's hydroxyl group to the backbone (A435) of VR-V situated above this residue. Thus, an intriguing working hypothesis for future work is that the removal of the hydroxyl group by mutation to phenylalanine could alter the conformation of VR-V, which is a potential determinant of host tropism, and result in reduced virus infectivity or vector transduction efficiency.

The best performers in our transduction assays were DQ000495 and HBoV1 Y523F, which consistently mediated comparable or higher transduction than the standard BoV vector GQ925675 and thus represent promising candidates for a future application as viral vectors. In contrast to the favorable effect of S590, two of the reconstructed mutants (KPLMI-30 and KPLMI-3503) resulted in 64- and 239-fold lower average viral titers, respectively, despite the presence of the S590 residue. The EIA assay revealed the presence of assembled particles for KPLMI-30 but not KPLMI-3503, despite the presence of free VP proteins (compare Fig. 6A and B), which implies a defect in genome packaging, particle assembly and/or antibody recognition, respectively. This hypothesis is supported by the localization of the different residues in both variants. In KPLMI-30, the two residues SA (at aa 396-397) are located on the capsid surface at

the 5-fold canyon (Fig. 8C-D), *i.e.*, a region surrounding the 5-fold axis channel, which serves as a portal for viral genome packaging (37, 47). By contrast, in KPLMI-3503, aa residues 534-536 (KPD) are on the inside of the capsid between the HI loop and VR-VIII that is located on the sides of the 3-fold protrusions, *i.e.*, the determinants of antibody recognition and infectivity in other parvoviruses (37) (Fig. 8E-F). Thus, we speculate that these residues might have led to structural changes that either negatively influenced capsid assembly or interfered with antibody binding. The latter would, however, not explain the 239-fold reduced viral titers. Consequently, at this point, the exact mechanisms underlying our observations remain unknown and our assumptions require experimental validation. For instance, to unanimously determine whether aa 534-536 affect particle assembly or impact genome packaging, a different antibody recognizing a conformational epitope should be used in future experiments. Finally, it is interesting that the transduction ability of both mutants was severely compromised in pHAE (Fig. 4E), which seems to be at odds with the high titers (~5×10^9 vg/ml) at which these mutants were detected in the original report (27). Here, it is important to mention again that only nucleotide changes that cause an aa change were transferred into our expression constructs. Thus, the effect of silent mutations that might have led to changes in different, as-of-yet unknown ORFs within *vp2* cannot be assessed in this setting.

In the last part of this work, we asked whether the T590S variation affects the serological reactivity of the HBoV1 capsid, which might have resulted in the emergence and/or persistence of this mutation in the clinical samples. This question is also important for the application of viral vectors in gene therapy, where pre-existing neutralizing antibodies significantly lower the therapeutic benefit. To this end, we pursued two independent approaches: (i) EIA to assess the binding of antibodies in HBoV1-positive sera to the HBoV1 *vp* variants, and (ii) a functional assay in which the impact of a pool of human antibodies (IVIg) on the transduction abilities of the HBoV1 variants was studied.

Interestingly, we found only small differences between the variants in the EIA, except for the GQ925675 variant that showed 2- to 4-fold reduced binding to serum antibodies (Fig. 6A). This, however, did not directly correlate with the results from the transduction assays, where GQ925675 had no benefit compared to the other tested variants (Fig. 6C). This discrepancy was also observed in previous reports using AAV vectors and indicated that antibody binding is not always linked to particle functionality (38). The *in vitro* neutralization assay also did not reveal any correlation between the T/S590 variation and the susceptibility of the HBoV1 capsid to neutralization by pooled human IgG. One exception is the Gr2 variant, which was more resistant to neutralization as compared to the other variants and which differed from Gr1 only in the presence of S590. This led us to the conclusion that T/S590 mostly affects the efficiency of virus production, yet it remains to be determined how this variation affects virus spread or latency.

The higher resistance of the KPLGr2 variant in this work is remarkable and would be of advantage for the future application of this variant as a viral vector. Thus, it would be now interesting to (i) validate the immune-escaping ability of KPLGr2 by testing its reactivity to different patient sera, and (ii) to include other variants that have the same aa sequence as KPLGr2, namely, V1512195, V1502611 and V1500812 but that differ substantially in their nucleotide composition. These variations have led to different transduction abilities in pHAE (Fig. 4A and 4D), which might be a result of alternative, as-of-yet undiscovered ORFs that could also influence the immunoreactivity. Thus, it should be highly rewarding to additionally study the abovementioned variants and the other BoV capsids reported in this work as it will further enrich our knowledge of bocaparvovirus biology and support efforts to breed optimal viral vectors for human gene therapy.

MATERIALS AND METHODS

Plasmids and cloning procedures. The HBoV1 helper plasmid pCMVNS*Cap (GenBank: GQ925675) was previously described (9, 48) and kindly provided by Ziying Yan. The DQ000495 *cap* sequence was ordered as gene block from IDT (Leuven, Belgium). The gene block and HBoV1 *cap* sequences from clinical samples were amplified using primers #1 and #2 in Supplementary Table 2 (both with overhangs containing BsmBI restriction sites). The PCR product was cloned using a Golden Gate reaction into a previously described acceptor plasmid (pCMVNS*ΔVP-2×BsmBI) (24) lacking the *cap* sequence.

All other synthetic variants reported in this work were generated by introducing mutations using overlap-extension (OE-)PCR as previously described (49). For each change, two PCR reactions were performed using overlapping primers (forward and reverse) containing the mutation(s) of interest (#5 to #36 in Supplementary Table 2) and two common external primers (#3 and #4 in Supplementary Table 2) with restriction sites (BstBI / EagI) that allow for the cloning of the end products into pCMVNS*Cap (see Supplementary Tables 3 and 4 for an overview of the nucleotide changes introduced).

Phylogenetic analysis. Phylogenetic analysis of HBoV1 *cap* sequences was performed in MEGA7.0.26 (Pennsylvania State University, PA, USA). The evolutionary history was inferred using the Maximum Likelihood method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. Only bootstrap values above 70% are displayed.

Cell culture and patient material. HEK293T were grown in Dulbecco's Modified Eagle's Medium (DMEM) with GlutaMAX™ (Thermo Fisher Scientific, Waltham, MA, USA), supplied with 10% fetal bovine serum (FBS) and 100 U/ml penicillin-streptomycin (both Merck/Sigma-

Aldrich, Darmstadt, Germany). Polarized human airway epithelial cells were generated as previously described (24) from resected bronchial tissue and were obtained from Lung Biobank Heidelberg (member of the German Center for Lung Research, DZL), at University Hospital Heidelberg, Germany. The cells were grown on ThinCerts (Greiner Bio-One, Frickenhausen, Germany) and differentiated at an air-liquid interface in PneumaCult ALI Basal medium supplemented with PneumaCult ALI 10× Supplement (both from StemCell, Vancouver, Canada).

Recombinant virus production. Pseudotyped BoV/AAV vectors were produced in HEK293T cells as previously described (24) using a triple plasmid transfection of (i) one of the BoV helper plasmids, (ii) a self-complementary (sc)AAV plasmid encoding Gluc (1.912 kb insert size) and (iii) pDGΔVP, a plasmid encoding *rep* from AAV2 and adenovirus helper genes (21). Cells were harvested 72 h post-transfection and the crude cell lysate was processed for iodixanol gradient centrifugation as previously described (24, 50). The virus-containing 40% iodixanol fraction was pulled from the gradient, mixed with 15 mL PBS and applied to an Amicon Ultra-15 (Merck, Darmstadt, Germany) centrifugal filter unit (100,000 nominal molecular weight limit). Several centrifugation steps at 500-1,000×g allowed for buffer exchange and concentration of viral preparations.

qPCR analysis of patient samples and recombinant virus titers. To determine viral titers, alkaline lysis was performed by mixing 10 µL of each virus stock with 10 µL TE buffer and 20 µL 2 M NaOH. The mixture was heated up to 56ºC for 30 min and then neutralized using 38 µL of 1 M HCl. Next, a 1:1,000 working solution was prepared and 5 µL were used in a TaqMan real-time PCR reaction as previously described (50), using the 2× SensiMix II Probe Kit (Bioline, Luckenwalde, Germany) and a probe binding in the promoter region (see Supplementary Table 5 for probe/primer combinations).

628 To determine the viral load in patient samples, 2.5 µL of extracted DNA (QIAsymphony kit;
629 Qiagen, Hilden, Germany), were directly mixed with 22.5 µL qPCR mix containing: (i) 12.5 µL
630 SensiMix SYBR No-Rox Kit (Bioline), (ii) 0.25 µL of each forward and reverse primer
631 (Supplementary Table 5) and (iii) 9.5 µL H_2O. The qPCR reaction were measured in duplicates
632 using a Rotor-Gene Q cycler (Qiagen) and the following conditions: Initial activation (95ºC, 10
633 s), followed by 40 cycles of (i) denaturation (95ºC, 15 s), (ii) annealing (58ºC, 20 s) and (iii)
634 extension (72ºC, 20 s). To ensure the detection of different HBoV1 strains, the forward and
635 reverse primers were designed to bind in the relatively constant promoter region.
636
637 ***Gaussia* luciferase assay.** Gluc activity was determined in the cell culture medium as
638 previously described (24). Briefly, 20 µL of the cell medium were incubated with 100 µL assay
639 buffer supplied with Coelenterazine (PJK, Kleinblittersdorf, Germany) at a final dilution of 11.7
640 µM. Gluc activity was detected in a GloMax96 microplate luminometer equipped with an
641 automatic injector (Promega, Madison, WI, USA).
642
643 **Enzyme immunoassay (EIA).** HBoV1 variants selected from the transduction experiments
644 were tested for their reactivity to human serum pools using an in-house IgG EIA. Therefore, 96-
645 well microtiter plates were coated overnight at 4°C in duplicates with 4-, 8-, 16- and 32-fold
646 diluted (in PBS) bocaviral stocks. Then, plates were coated with diluent LOY-X (Labsystems
647 Diagnostics, Vantaa, Finland) three times for 10 min each. A HBoV1-IgG-positive serum pool
648 (1:200) diluted in RED buffer (Kaivogen, Turku, Finland) was pipetted into each well and the
649 plate was incubated at 37°C for 1 h. As control, a 1:200 dilution of a HBoV1-IgG-negative serum
650 pool was used. After five washes with 0.05% Tween-20 in PBS, an HRP-(horseradish
651 peroxidase-) conjugated anti-human IgG (1:2,000; DAKO/Agilent, Glostrup, Denmark) diluted in
652 LOY-X was applied. For the detection of HRP-labeled antibodies, TMB substrate (Merck/Sigma-
653 Aldrich) was added and incubated at room temperature for 20 min. Then, H_2SO_4 (0.5 M) was

added to stop the reaction, and absorbances were measured at 450 nm using a MultiScan EX (Thermo Fischer Scientific).

Western blot analysis. Western blotting was performed as previously described (49). Briefly, 5×10^5 HEK293T cells were transfected with 2 µg of BoV helper plasmid. Three days post-transfection, the cells were harvested in 300 µL PBS, mixed with an equal volume of 2× SDS sample loading buffer and boiled for 5 min at 95ºC. The cell lysates were then centrifuged at 13,000 rpm and 10 µL from each lysate were separated on 8% SDS-PAGE gels (Biorad, Hercules, CA, USA). Next, the proteins were transferred to a nitrocellulose membrane (NeoLab, Heidelberg, Germany) using semi-dry transfer. For detection of the three capsid proteins (VP1, VP2 and VP3), an in-house produced rabbit polyclonal anti-HBoV1 antibody was used at a 1:1,000 dilution. To produce the anti-HBoV1 antibody, rabbits were inoculated with HBoV1 VP3-VLPs (virus-like particles) that were produced in a baculovirus expression system. Serum was obtained on day 120 and tested for HBoV1 IgG by EIA. Immunization of rabbits was performed by GenScript (Piscataway Township, NJ, USA). For further details, see reference (51).

For detection of the primary antibody, a HRP-conjugated secondary donkey anti-rabbit antibody (#NA934V; GE Healthcare, Chicago, IL, USA) was used (1:10,000). To visualize protein bands, the Lightning Plus-ECL reagent (PerkinElmer, Waltham, MA, USA) was added and a chemiluminescence imager (Intas ChemoStar, Göttingen, Germany) was used to detect the signal.

Transduction of pHAE in the presence or absence of IVIgs. Primary HAE were incubated from the apical side with the different pseudotyped HBoV1 variants at a MOI of 1×10^4. To enhance transduction, two proteasome inhibitors were applied to the medium on the basolateral side, as previously described (22, 24), namely, 5 µM doxorubicin hydrochloride (Santa Cruz

679 Biotechnology, Dellas, TX, USA) and 40 µM ALLN calpain Inhibitor I (G-Biosciences, St. Louis,
680 MO, USA). After 16 h, the virus was removed from the apical compartment and the medium
681 replaced with fresh medium without inhibitors.

682 To perform transductions in the presence of IVIg (Kiovig; Baxalta, Bannockburn, IL, USA), the
683 IVIg solution was diluted to working concentrations of 2, 10, 20, 40, 60 or 200 mg/dl. Next, equal
684 volumes of virus and IVIg solution were mixed and placed at 37ºC for 1 h. Positive (PBS,
685 +HBoV1) or negative (+IVIg, -HBoV1) controls for transduction were also included. Transgene
686 expression was measured at 5 or 10 days post-transduction.
687
688 **Flow cytometry analysis.** Characterization of pHAE cell composition using flow cytometry was
689 performed as previously described (22, 24). Briefly, to stain goblet and cilitated cells, the
690 following primary antibodies were used: goblet cell marker MUC5A/C (#ab3649; Abcam,
691 Cambridge, UK) diluted 1:100 and ciliated cell marker β-Tubulin IV (#T7941; Merck/Sigma-
692 Aldrich) diluted 1:100. Primary antibodies were incubated for 1 h at 4°C, followed by treatment
693 with secondary anti-mouse antibody (AF-647 goat anti-mouse #A21235, Thermo Fisher
694 Scientific) for 30 min at room temperature. Cells were measured on a FACSVerse (BD
695 Biosciences, Franklin Lakes, NJ, USA) and analysis was performed using Flowing Software
696 (version 2.5.1; Turku Centre for Biotechnology, Turku, Finland). Only living cells were used for
697 the analysis.
698
699 **Statistical analysis.** The statistical analysis was performed in PRISM Version 8.0 (GraphPad
700 Software Inc.; https://www.graphpad.com). Two data sets in Fig. 2B or 3D were compared using
701 an unpaired two-tailed t-test analysis (for 2B, Welch's correction was applied to account for
702 highly significant differences in variances). The data sets in Fig. 4B were analyzed using a
703 multiple t-test. Statistical significance was determined using the Holm-Sidak method, with alpha

= 0.05. Computations assume that all rows are samples from populations with the same scatter (SD). In Fig. 4D, a one-way ANOVA with Dunnett's test was used to compare each data set with the reference (DQ000495.1). Significance is denoted by asterisks above the SD or range bar. *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$, ****, $p < 0.0001$, ns, non-significant.

Structural analysis. For the structural analysis, the model of the HBoV1 VP monomer (PDB-ID: 5URF) and the HBoV1 capsid 60-mer was downloaded from the VIPERdb online server (http://viperdb.scripps.edu) (52). Surface exposed tyrosines were identified by generation of a gray surface representation of the capsid 60-mer followed by the coloration of tyrosine residues in Chimera (53). Similarly, the location of specific residues on the capsid surface such as serine 590 were identified. The depiction of the ribbon diagrams of the HBoV1 VP monomer were generated using the 'smooth loop' option in PyMol (54).

Selective pressure analysis. Tests for negative or positive selections were conducted on the Datamonkey server (55). Methods used involve single-likelihood ancestor (SLAC) (56), the fixed-effects likelihood (FEL) (56), the fast, unconstrained Bayesian approximation (FUBAR) (57) and the mixed effects model of evolution (MEME) (58). To reduce the probability of false-positive events, a p-value threshold of 0.1 in SLAC, FEL and MEME and a FUBAR posterior probability threshold of 0.9 were used to identify sites for selection.

Ethical approval. This study was carried out in accordance with the recommendations of the University Hospital Heidelberg with written informed consent from all subjects in accordance with the Declaration of Helsinki. All samples were received and maintained in an anonymized manner. The protocol was approved by the ethics commission at University Hospital Heidelberg under the protocol numbers S-270/2001 (collection of surgical material for lung research) and S-

782/2018 (amplification of BoV sequences from patient material and application of derived BoV vectors in pHAE).

AUTHOR CONTRIBUTIONS

J.F., K.-P.L. and D.G. conceived the study. J.F. and K.-P.L. generated constructs and performed the majority of experiments. M.A.-M. and Ma.M. generated capsid structures and residue information. M.S.-V. and M.X. designed and performed EIA assays. M.A.S. and Mi.M. provided pHAE. O.S. and P.S. provided extracted DNA from patient material. J.F. and D.G. wrote the manuscript. All authors read the manuscript and approved the final version.

CONFLICTS OF INTEREST

D.G. is a co-founder and shareholder of AaviGen GmbH. All other authors declare no competing financial interests.

ACKNOWLEDGMENTS

J.F. and D.G. are grateful for funding from the Cystic Fibrosis Foundation (CFF, grant number GRIMM15XX0), the German Research Foundation (DFG, Cluster of Excellence CellNetworks, EXC81) as well as from the Heidelberg Biosciences International Graduate School HBIGS at Heidelberg University. K.-P.L. and D.G. are thankful for a MD stipend to K.-P.L. from the German Center for Infection Research (DZIF, BMBF). D.G. acknowledges additional funding by the DZIF (TTU-HIV 04.803 and 04-815). M.A.S. and M.M. acknowledge funding by the German Center for Lung Research (DZL, BMBF, 82DZL00402). D.G. appreciates further support from the Collaborative Research Center SFB1129 (DFG; TP16; Project number 240245660). M.A.M. and M.M. were supported by NIH R01 GM082946. M.S.V. and M.X. acknowledge funding by the Sigfried Jusélius Foundation and the Life and Health Medical Association, and M.X. by the China Scholarship Council (CSC) and the Finnish-Norwegian Medical Foundation.

REFERENCES

1. Sanjuán R. 2012. From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. PLoS Pathog 8:e1002685.
2. Aiewsakun P, Katzourakis A. 2016. Time-Dependent Rate Phenomenon in Viruses. J Virol 90:7184-95.
3. Horiuchi M, Yamaguchi Y, Gojobori T, Mochizuki M, Nagasawa H, Toyoda Y, Ishiguro N, Shinagawa M. 1998. Differences in the evolutionary pattern of feline panleukopenia virus and canine parvovirus. Virology 249:440-52.
4. Shackelton LA, Parrish CR, Truyen U, Holmes EC. 2005. High rate of viral evolution associated with the emergence of carnivore parvovirus. Proc Natl Acad Sci U S A 102:379-84.
5. Shackelton LA, Holmes EC. 2006. Phylogenetic evidence for the rapid evolution of human B19 erythrovirus. J Virol 80:3666-9.
6. Ren X, Tao Y, Cui J, Suo S, Cong Y, Tijssen P. 2013. Phylogeny and evolution of porcine parvovirus. Virus Res 178:392-7.
7. Hao R, Ni K, Xia Q, Peng C, Deng Y, Zhao X, Fu Z, Liu W, Liu E. 2013. Correlation between nucleotide mutation and viral loads of human bocavirus 1 in hospitalized children with respiratory tract infection. J Gen Virol 94:1079-85.
8. Zhang J, Bai Y, Zhu B, Hao S, Chen Z, Wang H, Guan W. 2017. Mutations in the C-terminus of HBoV NS1 affect the function of NP1. Sci Rep 7:7407.
9. Zou W, Cheng F, Shen W, Engelhardt JF, Yan Z, Qiu J. 2016. Nonstructural Protein NP1 of Human Bocavirus 1 Plays a Critical Role in the Expression of Viral Capsid Proteins. J Virol 90:4658-4669.
10. Zhang Z, Zheng Z, Luo H, Meng J, Li H, Li Q, Zhang X, Ke X, Bai B, Mao P, Hu Q, Wang H. 2012. Human bocavirus NP1 inhibits IFN-beta production by blocking association of IFN regulatory factor 3 with IFNB promoter. J Immunol 189:1144-53.
11. Qu XW, Liu WP, Qi ZY, Duan ZJ, Zheng LS, Kuang ZZ, Zhang WJ, Hou YD. 2008. Phospholipase A2-like activity of human bocavirus VP1 unique region. Biochem Biophys Res Commun 365:158-63.
12. Yan Z, Lei-Butters DC, Keiser NW, Engelhardt JF. 2013. Distinct transduction difference between adeno-associated virus type 1 and type 6 vectors in human polarized airway epithelia. Gene Ther 20:328-37.
13. Song L, Kauss MA, Kopin E, Chandra M, Ul-Hasan T, Miller E, Jayandharan GR, Rivers AE, Aslanidi GV, Ling C, Li B, Ma W, Li X, Andino LM, Zhong L, Tarantal AF, Yoder MC, Wong KK, Jr., Tan M, Chatterjee S, Srivastava A. 2013. Optimizing the transduction efficiency of capsid-modified AAV6 serotype vectors in primary human hematopoietic stem cells in vitro and in a xenograft mouse model in vivo. Cytotherapy 15:986-98.
14. Handa A, Muramatsu S, Qiu J, Mizukami H, Brown KE. 2000. Adeno-associated virus (AAV)-3-based vectors transduce haematopoietic cells not susceptible to transduction with AAV-2-based vectors. J Gen Virol 81:2077-84.
15. Rabinowitz JE, Bowles DE, Faust SM, Ledford JG, Cunningham SE, Samulski RJ. 2004. Cross-dressing the virion: the transcapsidation of adeno-associated virus serotypes functionally defines subgroups. J Virol 78:4421-32.
16. Maxwell IH, Spitzer AL, Maxwell F, Pintel DJ. 1995. The capsid determinant of fibrotropism for the MVMp strain of minute virus of mice functions via VP2 and not VP1. J Virol 69:5829-32.
17. Itah R, Tal J, Davis C. 2004. Host cell specificity of minute virus of mice in the developing mouse embryo. J Virol 78:9474-86.

18. Calatayud O, Esperón F, Cleaveland S, Biek R, Keyyu J, Eblate E, Neves E, Lembo T, Lankester F. 2019. Carnivore Parvovirus Ecology in the Serengeti Ecosystem: Vaccine Strains Circulating and New Host Species Identified. J Virol 93.
19. Maxwell IH, Long CJ, Carlson JO, Rhode SL, 3rd, Maxwell F. 1993. Encapsidation of a recombinant LuIII parvovirus genome by H1 virus and the fibrotropic or lymphotropic strains of minute virus of mice. J Gen Virol 74 (Pt 6):1175-9.
20. Ponnazhagan S, Weigel KA, Raikwar SP, Mukherjee P, Yoder MC, Srivastava A. 1998. Recombinant human parvovirus B19 vectors: erythroid cell-specific delivery and expression of transduced genes. J Virol 72:5224-30.
21. Grimm D, Kay MA, Kleinschmidt JA. 2003. Helper virus-free, optically controllable, and two-plasmid-based production of adeno-associated virus vectors of serotypes 1 to 6. Mol Ther 7:839-50.
22. Yan Z, Keiser W, Song Y, Deng X, Cheng F. 2013. A Novel Chimeric Adenoassociated Virus 2/Human Bocavirus 1 Parvovirus Vector Efficiently Transduces Human Airway Epithelia. Mol Ther 21:2181-94.
23. Zincarelli C, Soltys S, Rengo G, Rabinowitz JE. 2008. Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. Mol Ther 16:1073-80.
24. Fakhiri J, Schneider MA, Puschhof J, Stanifer M, Schildgen V, Holderbach S, Voss Y, El Andari J, Schildgen O, Boulant S, Meister M, Clevers H, Yan Z, Qiu J, Grimm D. 2019. Novel Chimeric Gene Therapy Vectors Based on Adeno-Associated Virus and Four Different Mammalian Bocaviruses. Mol Ther Methods Clin Dev 12:202-222.
25. Yan Z, Feng Z, Sun X, Zhang Y, Zou W, Wang Z, Jensen-Cody C, Liang B, Park SY, Qiu J, Engelhardt JF. 2017. Human Bocavirus Type-1 Capsid Facilitates the Transduction of Ferret Airways by Adeno-Associated Virus Genomes. Hum Gene Ther 28:612-625.
26. Allander T, Tammi MT, Eriksson M, Bjerkner A, Tiveljung-Lindell A, Andersson B. 2005. Cloning of a human parvovirus by molecular screening of respiratory tract samples. Proc Natl Acad Sci U S A 102:12891-6.
27. Principi N, Piralla A, Zampiero A, Bianchini S, Umbrello G, Scala A, Bosis S, Fossali E, Baldanti F, Esposito S. 2015. Bocavirus Infection in Otherwise Healthy Children with Respiratory Disease. PLoS One 10:e0135640.
28. Huang Q, Deng X, Yan Z, Cheng F, Luo Y, Shen W, Lei-Butters DC, Chen AY, Li Y, Tang L, Soderlund-Venermo M, Engelhardt JF, Qiu J. 2012. Establishment of a reverse genetics system for studying human bocavirus in human airway epithelia. PLoS Pathog 8:e1002899.
29. Ziying Y, Keiser W, Song Y, Deng X, Cheng F. 2013. A Novel Chimeric Adenoassociated Virus 2/Human Bocavirus 1 Parvovirus Vector Efficiently Transduces Human Airway Epithelia. Mol Ther 21:2181-94.
30. Salganik M, Aydemir F, Nam HJ, McKenna R, Agbandje-McKenna M, Muzyczka N. 2014. Adeno-associated virus capsid proteins may play a role in transcription and second-strand synthesis of recombinant genomes. J Virol 88:1071-9.
31. Reguera J, Grueso E, Carreira A, Sanchez-Martinez C, Almendral JM, Mateu MG. 2005. Functional relevance of amino acid residues involved in interactions with ordered nucleic acid in a spherical virus. J Biol Chem 280:17969-77.
32. Markusic DM, Nichols TC, Merricks EP, Palaschak B, Zolotukhin I, Marsic D, Zolotukhin S, Srivastava A, Herzog RW. 2017. Evaluation of engineered AAV capsids for hepatic factor IX gene transfer in murine and canine models. J Transl Med 15:94.
33. Petrs-Silva H, Dinculescu A, Li Q, Min SH, Chiodo V, Pang JJ, Zhong L, Zolotukhin S, Srivastava A, Lewin AS, Hauswirth WW. 2009. High-efficiency transduction of the mouse retina by tyrosine-mutant AAV serotype vectors. Mol Ther 17:463-71.

855 34. Zhong L, Li B, Mah CS, Govindasamy L, Agbandje-McKenna M, Cooper M, Herzog RW,
856 Zolotukhin I, Warrington KH, Jr., Weigel-Van Aken KA, Hobbs JA, Zolotukhin S,
857 Muzyczka N, Srivastava A. 2008. Next generation of adeno-associated virus 2 vectors:
858 point mutations in tyrosines lead to high-efficiency transduction at lower doses. Proc Natl
859 Acad Sci U S A 105:7827-32.
860 35. Martini SV, Silva AL, Ferreira D, Rabelo R, Ornellas FM, Gomes K, Rocco PR, Petrs-
861 Silva H, Morales MM. 2016. Tyrosine Mutation in AAV9 Capsid Improves Gene Transfer
862 to the Mouse Lung. Cell Physiol Biochem 39:544-53.
863 36. Ling C, Li B, Ma W, Srivastava A. 2016. Development of Optimized AAV Serotype
864 Vectors for High-Efficiency Transduction at Further Reduced Doses. Hum Gene Ther
865 Methods 27:143-9.
866 37. Mietzsch M, Kailasan S, Garrison J, Ilyas M, Chipman P, Kantola K, Janssen ME, Spear
867 J, Sousa D, McKenna R, Brown K, Söderlund-Venermo M, Baker T, Agbandje-McKenna
868 M. 2017. Structural Insights into Human Bocaparvoviruses. J Virol 91.
869 38. Fitzpatrick Z, Leborgne C, Barbon E, Masat E, Ronzitti G, van Wittenberghe L, Vignaud
870 A, Collaud F, Charles S, Simon Sola M, Jouen F, Boyer O, Mingozzi F. 2018. Influence
871 of Pre-existing Anti-capsid Neutralizing and Binding Antibodies on AAV Vector
872 Transduction. Mol Ther Methods Clin Dev 9:119-129.
873 39. Kapoor A, Mehta N, Esper F, Poljsak-Prijatelj M, Quan PL, Qaisar N, Delwart E, Lipkin
874 WI. 2010. Identification and characterization of a new bocavirus species in gorillas. PLoS
875 One 5:e11948.
876 40. Yan Z, Zou W, Feng Z, Shen W, Park SY, Deng X, Qiu J, Engelhardt JF. 2018.
877 Establishment of a High Yield rAAV/HBoV Vector Production System Independent of
878 Bocavirus Non-structural Proteins. Hum Gene Ther doi:10.1089/hum.2018.173.
879 41. Kenmoe S, Vernet MA, Njankouo-Ripa M, Penlap VB, Vabret A, Njouom R. 2017.
880 Phylogenic analysis of human bocavirus detected in children with acute respiratory
881 infection in Yaounde, Cameroon. BMC Res Notes 10:293.
882 42. Babkin IV, Tyumentsev AI, Tikunov AY, Kurilshikov AM, Ryabchikova EI, Zhirakovskaya
883 EV, Netesov SV, Tikunova NV. 2013. Evolutionary time-scale of primate bocaviruses.
884 Infect Genet Evol 14:265-74.
885 43. Zehender G, De Maddalena C, Canuti M, Zappa A, Amendola A, Lai A, Galli M, Tanzi E.
886 2010. Rapid molecular evolution of human bocavirus revealed by Bayesian coalescent
887 inference. Infect Genet Evol 10:215-20.
888 44. Martin ET, Kuypers J, McRoberts JP, Englund JA, Zerr DM. 2015. Human Bocavirus 1
889 Primary Infection and Shedding in Infants. J Infect Dis 212:516-24.
890 45. Gurda BL, Parent KN, Bladek H, Sinkovits RS, DiMattia MA, Rence C, Castro A,
891 McKenna R, Olson N, Brown K, Baker TS, Agbandje-McKenna M. 2010. Human
892 bocavirus capsid structure: insights into the structural repertoire of the parvoviridae. J
893 Virol 84:5880-9.
894 46. DiPrimio N, Asokan A, Govindasamy L, Agbandje-McKenna M, Samulski RJ. 2008.
895 Surface loop dynamics in adeno-associated virus capsid assembly. J Virol 82:5178-89.
896 47. Kailasan S, Garrison J, Ilyas M, Chipman P, McKenna R, Kantola K, Soderlund-
897 Venermo M, Kucinskaite-Kodze I, Zvirbliene A, Agbandje-McKenna M. 2016. Mapping
898 Antigenic Epitopes on the Human Bocavirus Capsid. J Virol 90:4670-4680.
899 48. Yan Z, Zou W, Feng Z, Shen W, Park SY, Deng X, Qiu J, Engelhardt JF. 2019.
900 Establishment of a High-Yield Recombinant Adeno-Associated Virus/Human Bocavirus
901 Vector Production System Independent of Bocavirus Nonstructural Proteins. Hum Gene
902 Ther 30:556-570.
903 49. Grosse S, Penaud-Budloo M, Herrmann AK, Borner K, Fakhiri J, Laketa V, Kramer C,
904 Wiedtke E, Gunkel M, Menard L, Ayuso E, Grimm D. 2017. Relevance of Assembly-

Activating Protein for Adeno-associated Virus Vector Production and Capsid Protein Stability in Mammalian and Insect Cells. J Virol 91.
50. Fakhiri J, Nickl M, Grimm D. 2019. Rapid and Simple Screening of CRISPR Guide RNAs (gRNAs) in Cultured Cells Using Adeno-Associated Viral (AAV) Vectors. Methods Mol Biol 1961:111-126.
51. Li X, Kantola K, Hedman L, Arku B, Hedman K, Söderlund-Venermo M. 2015. Original antigenic sin with human bocaviruses 1-4. J Gen Virol 96:3099-108.
52. Carrillo-Tripp M, Shepherd CM, Borelli IA, Venkataraman S, Lander G, Natarajan P, Johnson JE, Brooks CL, Reddy VS. 2009. VIPERdb2: an enhanced and web API enabled relational database for structural virology. Nucleic Acids Res 37:D436-42.
53. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25:1605-12.
54. DeLano, WL. 2002. Pymol: An open-source molecular graphics tool, p 82–92. *In* CCP4 Newsletter on protein crystallography.
55. Delport W, Poon AF, Frost SD, Kosakovsky Pond SL. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26:2455-7.
56. Kosakovsky Pond SL, Frost SD. 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol 22:1208-22.
57. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K. 2013. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol Biol Evol 30:1196-205.
58. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. 2012. Detecting individual sites subject to episodic diversifying selection. PLoS Genet 8:e1002764.

FIGURE LEGENDS

FIG 1 Isolation and sequencing of natural HBoV1 capsid variants. **(A)** Representative agarose gel showing the results of PCR amplification of HBoV1 capsid genes in patient samples. The expected size of the full-length PCR amplicon was 2016 bp. POS, positive control (HBoV1 helper plasmid); NEG, negative control (H_2O). **(B-C)** Nucleotide (B) and protein (C) sequences of the HBoV1 capsid variants shown on the left in each panel. The second column depicts the number of nucleotide (B) or aa (C) differences as compared to variant DQ000495 (shown at the top). The numbers above each column have to be read vertically and indicate the position of each nucleotide (B) or aa (C) in the HBoV1 capsid gene or protein (VP1), respectively. A dot indicates no change. The four colors highlight the corresponding nucleotides and AA in the two panels. **(D)** Phylogenetic tree of the shown HBoV1 capsid variants derived by applying

maximum-likelihood methodology and 500 bootstrap repeats. Bootstrap repeats with a support of over 60% are shown at the nodes. HBoV2 (GenBank NC012042), HBoV3 (NC012564), HBoV4 (NC012729), Bovine Parvovirus (NC001540) and Canine Minute Virus (NC004442) were defined as outgroup. Capsid variants that resulted from this work are marked with a dot.

FIG 2 Correlation of HBoV1 capsid sequence and viral load in patient material. **(A)** Results of qPCR titration of viral loads in 39 different patient samples. Shown is the median value (3.91×10^6 vg/mL). **(B)** Correlation of aa at position 590 (threonine or serine) with viral load in 31 selected (see main text for criteria) samples. Shown are means plus range. ns, non-significant. **(C)** Sequencing results of selected regions of the HBoV1 capsid in material taken from patient A (sputum and tracheal secretions) at the four indicated time points (total collection period was about 15 weeks). Also shown are the corresponding viral loads. Letters and colors indicate the aa or nucleotide at each position (green = adenine, blue = cytosine, black = guanine, red = thymine). **(D)** Same as panel C, but data for patient B and material (tracheal secretions) collected over a period of three weeks. **(E)** Viral loads measured for patients A and B (vg/mL) and shown in panels (**C**) and (**D**) are plotted (Y-axis) against the measurement time-points (days; x-axis).

FIG 3 Selection and production of recombinant HBoV1 capsid variants. **(A-B)** Nucleotide (A) and protein (B) sequences of the 18 HBoV1 capsid variants shown on the left in each panel. The second column depicts the number of nucleotide (A) or aa (B) differences as compared to variant DQ000495 (shown at the top). The numbers above each column have to be read vertically and indicate the position of each nucleotide (A) or aa (B) in the HBoV1 capsid gene or protein (VP1), respectively. A dot indicates no change. The colors highlight the corresponding nucleotides and aa in the two panels. Panel B also shows the viral loads that were measured for each of the corresponding patient samples. Values marked with an asterisk were taken from

969 Principi *et al*. 2015 (27) and rounded. **(C)** Vector titers (means ± SD) were determined by qPCR
970 for the shown 18 HBoV1 variants (each produced in nine 15 cm^2 plates and purified). All
971 titrations were performed twice except for those marked with an asterisk that were performed
972 once. The dotted line represents the average value of 1.1×10^{11} vg/ml. Colored bars represent
973 randomly chosen candidates that are highlighted in the next section of this work. **(D)** Same
974 results as in panel C but sorted by the presence of a threonine (n=6) or serine (n=12) at position
975 590. Shown are means ± SD. ***, $p < 0.001$ (unpaired t-test).
976
977 **FIG 4** Functional characterization of HBoV1 variants. **(A)** The indicated HBoV1 variants were
978 tested for their transduction abilities in pHAE derived from five different patients (#171834 to
979 #171476). To this end, a rAAV-Gluc genome was packaged into each HBoV1 capsid and pHAE
980 were transduced apically at a MOI of 600. Gluc activity (means ± SD) was measured in the
981 medium 12 days post-transduction and plotted on the y-axis as arbitrary light units (ALU).
982 Colored bars represent candidates shown in Fig. 3C and are intended to facilitate comparison of
983 candidates between pHAE derived from different patients. **(B)** Same results as in panel C but
984 sorted by the presence of a threonine (n=6, 2 transwells each) or serine (n=12, 2 transwells
985 each) at position 590. Shown are means ± SD. ns, non-significant (multiple t-test). **(C)** Flow
986 cytometry analysis of untransduced pHAE derived from the indicated patient samples (n=2
987 independent transwells per patient). Cells were stained for cell type-specific markers, *i.e.*, ß-
988 Tubulin IV (ciliated cells) and MUC5AC (goblet cells). Blue areas show the percentages of
989 positive cells for the indicated markers. **(D)** Side-by-side comparison of selected HBoV1 variants
990 highlighted in Fig. 3C and 4A. Shown is their transduction ability in pHAE derived from different
991 patients (indicated by the numbers below the x-axis). Transduction efficiency was estimated
992 from the Gluc reporter activity in the medium, which is plotted as ALU on the y-axis (mean ±
993 SD). *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$, ****, $p < 0.0001$, ns, non-significant (one-way
994 ANOVA). **(E)** Transduction of pHAE with the shown HBoV1 variants at a MOI of 200. Gluc

activity (means ± SD) was measured in the medium 3, 9 and 12 days post-transduction and plotted on the y-axis as ALU. NEG, negative control (untransduced cells).

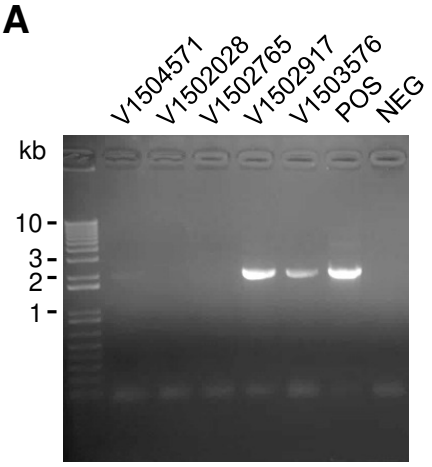
FIG 5 Synthetic tyrosine mutants of the HBoV1 capsid and their transduction properties. **(A)** Shown tyrosine residues (Y) were mutated to phenylalanines (F). Numbers indicate the position of the aa in VP1. Underlined nucleotides represent mutated residues that result in the corresponding aa change. **(B)** Gluc activity measured at the indicated time points. pHAE were transduced with the different tyrosine HBoV1 mutants from panel (**A**) at a MOI of 2×10^4 in the presence (+) or absence (-) of proteasome inhibitors. N, negative control (untransduced cells); wt, wild-type HBoV1 capsid (positive control). The dotted line represents the assay background. **(C)** HBoV1 capsid surface representation with surface-exposed tyrosine residues highlighted in green and the hydroxyl group colored in red. In variants with an asterisk, the hydroxyl group is partially inaccessible. This image was generated using Chimera.

FIG 6 Analysis of the immunoreactivity of HBoV1 capsid variants. **(A)** EIA using pooled human sera positive for HBoV1. Iodixanol-purified viral stocks (adjusted to an average of 5×10^{10} vg/mL) of the indicated HBoV1 capsid variants were used to coat the wells of a microtiter plate (8- and 16-fold dilutions). EIA absorbance values (optical density; OD) are plotted on the y-axis. **(B)** Western blot analysis of the variants shown in panel (A). HEK293T cells were transfected with the different HBoV1 plasmids to analyze the expression of the three capsid proteins VP1, VP2 and VP3. **(C)** *In vitro* neutralization assay using commercially available human immunoglobulins (IVIg). Gluc-expressing vectors were pre-incubated with the indicated IVIg concentrations for 1 h at 37°C and then used to apically transduce pHAE at a MOI of 1×10^4 (5×10^9 vg per well). All assays were performed in duplicates. Shown are mean Gluc activity levels (light units) plus range measured at days 5 and 10 post-transduction.

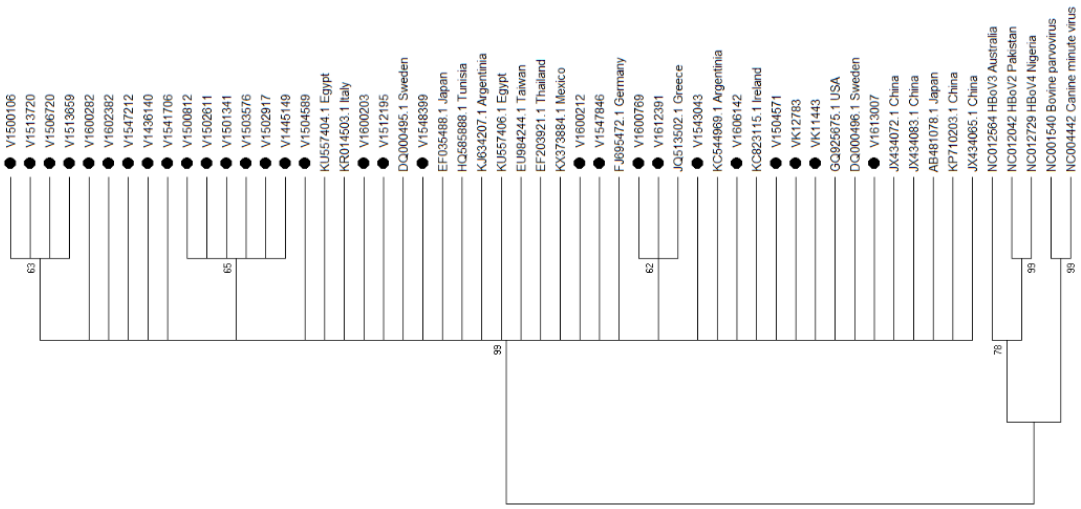
FIG 7 Selection pressure acting on the intra-host level and comparison of the inter-host genetic diversity of the VR-VIIIB. **(A)** Positive or negative selection pressure acting on the *vp* codons of the HBoV1 strains in this study. MEME: mixed effects model of evolution; SLAC: single-likelihood ancestor; FUBAR: fast unconstrained Bayesian approximation methods; FEL: fixed-effects likelihood. **(B)** Amino acid 590 variation among different primate bocaviruses.

FIG 8 Structural representation of the HBoV1 T590S variant (GQ925675) as well as of the HBoV1 variants KPLMI-30 and KPLMI-3503. **(A)** Shown is the HBoV1 capsid surface with the 2-, 3- and 5-fold symmetry axes as well as the location of the S590 residue (in green). **(B)** Ribbon diagrams of the VP3 monomer of the HBoV1 S590 variant. The S590 residue is represented by a green sphere and localizes to the VR-VIIIB (HI loop). **(C,E)** Surface representation of the HBoV1 KPLMI-30 (**C**) or KPLMI-3503 (**E**) capsid. The aa residues 396-397 in KPLMI-30 are located on the capsid surface and are shown by green spheres, whereas the critical residues in KPLMI-3503 (534-536) reside inside the capsid and are not visible in the image in panel (**E**). **(D,F)** Ribbon diagrams of the VP3 monomer of the HBoV1 variants in panels (**C**) and (**E**), respectively. Residue changes as compared to DQ000495 mentioned in panels (**C**) and (**E**) are shown as green spheres. Images in panels (**A**), (**C**) and (**E**) are radially colored from blue to white and red, representing capsid center to surface regions. These images were generated using Chimera. In panels (**B**), (**D**) and (**F**), the capsid VRs from I to IX are indicated. These images were generated using PyMOL (http://www.pymol.org/).

A



D



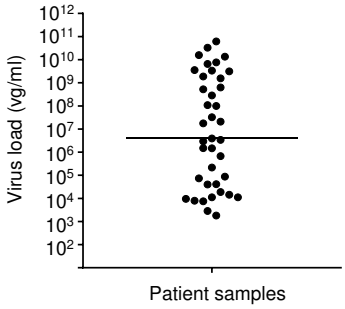
B

		0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		0	1	2	2	4	4	4	7	7	7	8	8	9	0	1	1	1	1	1	3	4	5	5	6	7	7	7	7	7	8	9	9
		3	5	0	6	2	4	4	1	6	7	0	7	8	4	4	6	7	7	8	0	2	6	7	6	0	5	6	6	8	3	5	7
		6	0	2	4	9	1	5	4	8	4	4	3	4	4	0	8	0	6	8	8	1	0	8	2	1	8	7	8	5	3	9	7
DQ000495	**(d) nt**	**G**	**G**	**G**	**G**	**G**	**G**	**G**	**A**	**A**	**G**	**T**	**G**	**A**	**G**	**T**	**T**	G	C	**T**	**G**	**G**	**G**	**G**	**G**	**C**	**A**	**A**	**A**	**T**	**A**	**G**	**G**
V1500106	17	.	.	.	.	.	A	A	G	.	.	.	A	C	.	C	C	A	T	C	A	.	.	.	.	A	T	C	T	C	G	.	.
V1600203	4	A	.	.	.	.	.	A	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.
V1600212	5	.	.	.	.	.	.	A	G	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	A
V1600282	16	.	.	.	.	.	A	A	G	.	.	.	.	C	.	C	C	A	T	C	A	.	.	.	.	A	T	C	T	C	G	.	.
V1600769	7	.	.	.	.	.	.	A	G	.	.	.	.	C	.	C	C	A	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
V1500812	11	.	.	.	.	.	A	A	G	.	.	G	.	C	.	.	C	.	T	C	A	.	.	.	.	.	T	.	.	.	G	.	.
V1501341	11	.	.	.	.	.	A	A	G	.	.	G	.	C	.	.	C	.	T	C	A	.	.	.	.	.	T	.	.	.	G	.	.
V1602382	17	.	.	.	.	.	A	A	G	.	.	.	.	C	.	C	C	A	T	C	A	.	.	A	.	A	T	C	T	C	G	.	.
V1502611	11	.	.	.	.	.	A	A	G	.	.	G	.	C	.	.	C	.	T	C	A	.	.	.	.	.	T	.	.	.	G	.	.
V1502917	11	.	.	.	.	.	A	A	G	.	.	G	.	C	.	.	C	.	T	C	A	.	.	.	.	.	T	.	.	.	G	.	.
V1503576	11	.	.	.	.	.	A	A	G	.	.	G	.	C	.	.	C	.	T	C	A	.	.	.	.	.	T	.	.	.	G	.	.
V1504571	17	.	.	.	.	.	A	A	G	.	.	G	A	C	.	C	C	A	T	C	A	.	.	.	.	A	T	C	T	.	G	.	.
V1504589	13	.	.	.	.	.	A	A	G	.	.	G	.	C	.	C	C	.	T	C	A	.	.	.	.	.	T	C	.	.	G	.	.
V1606142	14	.	.	.	.	.	.	A	G	.	.	.	.	C	.	C	C	A	T	.	A	.	.	.	.	A	T	C	T	C	G	.	.
V1506720	17	.	.	.	.	.	A	A	G	.	.	.	A	C	.	C	C	A	T	C	A	.	.	.	.	A	T	C	T	C	G	.	.
V1512195	5	.	A	.	A	.	.	A	G	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
V1513659	15	.	.	.	.	.	A	A	G	.	.	.	A	C	.	C	C	.	T	C	A	.	.	.	.	.	T	C	T	C	G	.	.
V1513720	17	.	.	.	.	.	A	A	G	.	.	.	A	C	.	C	C	A	T	C	A	.	.	.	.	A	T	C	T	C	G	.	.
V1436140	15	.	.	.	.	.	A	A	G	.	.	G	.	C	.	C	C	A	T	C	A	.	.	.	.	.	T	C	T	.	G	.	.
V1541706	19	.	.	A	.	.	A	A	G	G	.	.	.	C	A	C	C	A	T	C	A	.	.	.	.	A	T	C	T	C	G	.	.
V1543043	15	.	.	.	.	.	.	A	G	.	A	.	A	C	.	C	C	A	T	.	.	.	.	.	.	A	T	C	T	C	G	.	.
V1445149	11	.	.	.	.	.	A	A	G	.	.	G	.	C	.	.	C	.	T	C	A	.	.	.	.	.	T	.	.	.	G	.	.
V1547212	16	.	.	.	.	.	A	A	G	.	.	.	.	C	.	C	C	A	T	C	A	.	.	.	.	A	T	C	T	C	G	.	.
V1547846	5	.	.	.	.	.	.	A	G	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	A
V1548399	6	.	.	.	.	.	.	A	G	.	A	.	.	C	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	A
V1612391	7	.	.	.	.	.	.	A	G	.	.	.	.	C	.	C	C	A	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.
V1613007	18	.	.	.	.	.	A	A	G	.	A	.	.	C	.	C	C	A	T	.	A	.	A	.	.	A	T	C	T	C	G	A	.
VK11443	18	.	.	.	.	A	.	A	G	.	.	.	.	C	.	C	C	A	T	C	A	A	A	.	.	A	T	C	T	C	G	.	.
VK12783	18	.	.	.	.	A	.	A	G	.	.	.	.	C	.	C	C	A	T	C	A	A	A	.	.	A	T	C	T	C	G	.	.

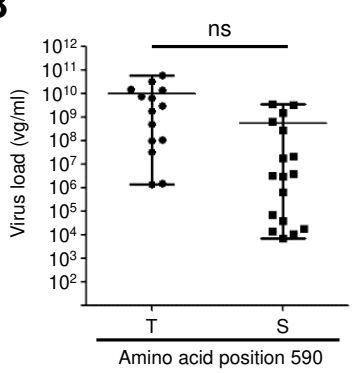
C

		0	1	4	5
		6	4	7	9
		8	9	4	0
DQ000495	**(d) aa**	**D**	**A**	**S**	**T**
V1500106	2	.	T	.	S
V1600203	1	.	T	.	.
V1600212	1	.	T	.	.
V1600282	2	.	T	.	S
V1600769	1	.	T	.	.
V1500812	1	.	T	.	.
V1501341	1	.	T	.	.
V1602382	2	.	T	.	S
V1502611	1	.	T	.	.
V1502917	1	.	T	.	.
V1503576	1	.	T	.	.
V1504571	2	.	T	.	S
V1504589	1	.	T	.	.
V1606142	2	.	T	.	S
V1506720	2	.	T	.	S
V1512195	1	.	T	.	.
V1513659	2	.	T	.	S
V1513720	2	.	T	.	S
V1436140	2	.	T	.	S
V1541706	3	N	T	.	S
V1543043	2	.	T	.	S
V1445149	1	.	T	.	.
V1547212	2	.	T	.	S
V1547846	1	.	T	.	.
V1548399	1	.	T	.	.
V1612391	1	.	T	.	.
V1613007	2	.	T	.	S
VK11443	3	.	T	N	S
VK12783	3	.	T	N	S

A



B



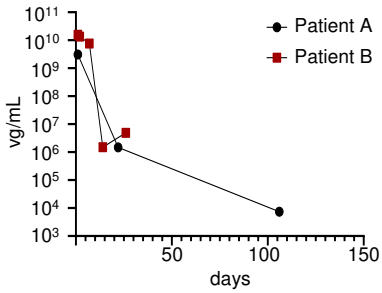
C

Sample	V1445149	V1501341	V1504571	V1513659
Time point	#1	#2 (+ 21 d)	#3 (+ 42 d)	#4 (+ 105 d)
Virus load (vg/ml)	3.11×10^{9}	1.47×10^{6}	-	7.46×10^{3}
Amino acids: 380–381	NN	NN	NN	NN
Nucleotides: 1138–1143	AA**T**AAT	AA**T**AAT	AA(**C/T**)AAT	AA(**C/T**)AAT
Amino acids: 390–391	LM	LM	LM	LM
Nucleotides: 1168–1173	CT**G**ATG	CT**G**ATG	CT(**A/G**)ATG	CT(**A/G**)ATG
Amino acids: 589–590	A**T**	A**T**	A(**S/T**)	A(**S/T**)
Nucleotides: 1765–1770	GC**AA**CA	GC**AA**CA	GC(**CT/AA**)CA	GC(**CT/AA**)CA

D

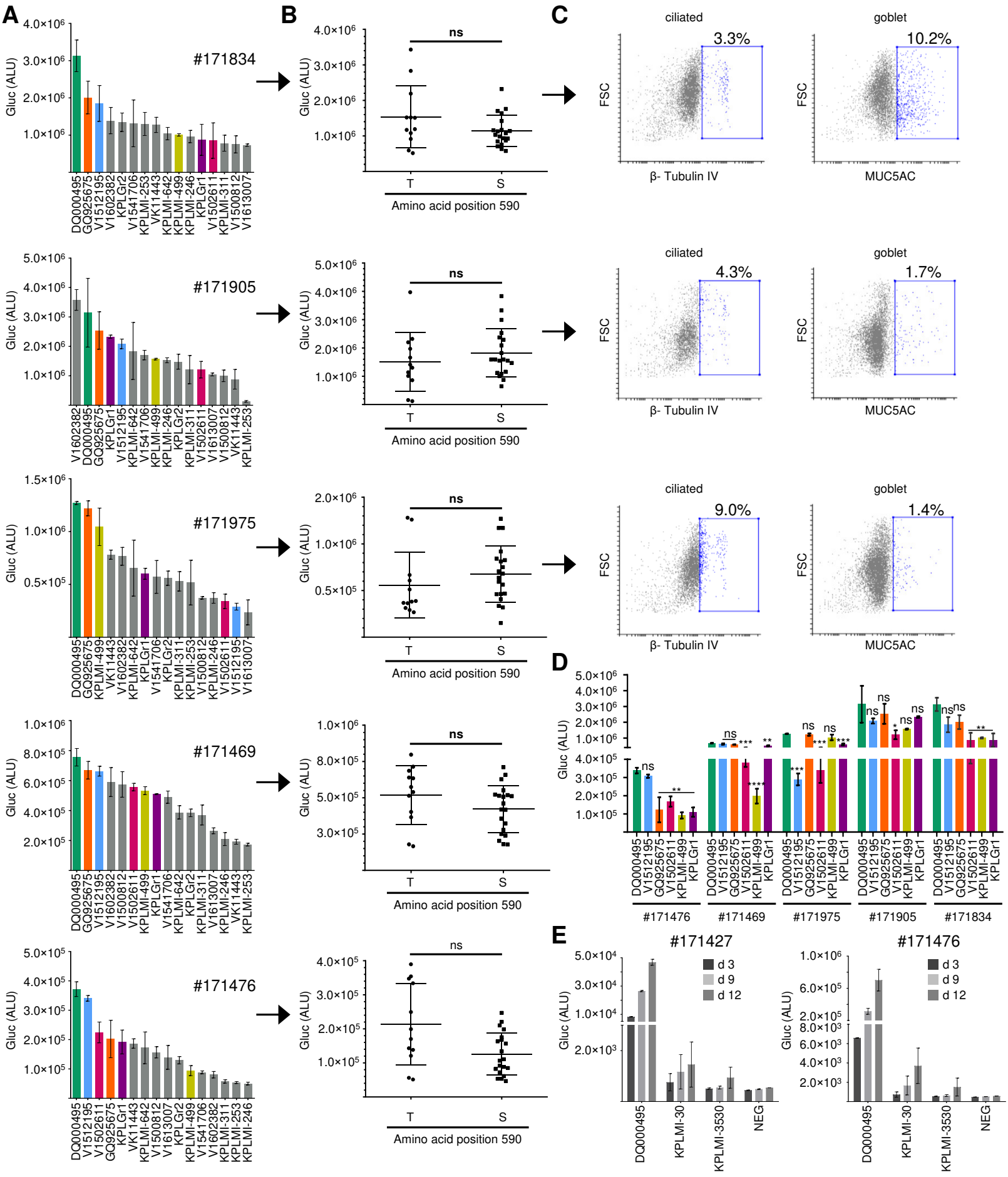
Sample	V1502611	V1502917	V1503576	V1504589	V1506218
Time point	#1	#2 (+ 1 d)	#3 (+ 6 d)	#4 (+ 13 d)	#5 (+ 25 d)
Virus load (vg/ml)	1.57×10^{10}	1.35×10^{10}	7.65×10^{9}	1.50×10^{6}	3.91×10^{6}
Amino acids: 380–381	NN	NN	NN	NN	NN
Nucleotides: 1138–1143	AA**T**AAT	AA**T**AAT	AA**T**AAT	AA(**C/T**)AAT	AA(**C/T**)AAT
Amino acids: 390–391	LM	LM	LM	LM	LM
Nucleotides: 1168–1173	**C**T**G**ATG	**C**T**G**ATG	**C**T**G**ATG	**C**T(**A/G**)ATG	(**C/T**)T(**A/G**)ATG
Amino acids: 589–590	A**T**	A**T**	A**T**	A(**S/T**)	A(**S/T**)
Nucleotides: 1765–1770	GC**AA**CA	GC**AA**CA	GC**AA**CA	GC(**CT/AA**)CA	GC(**CT/AA**)CA

E



A

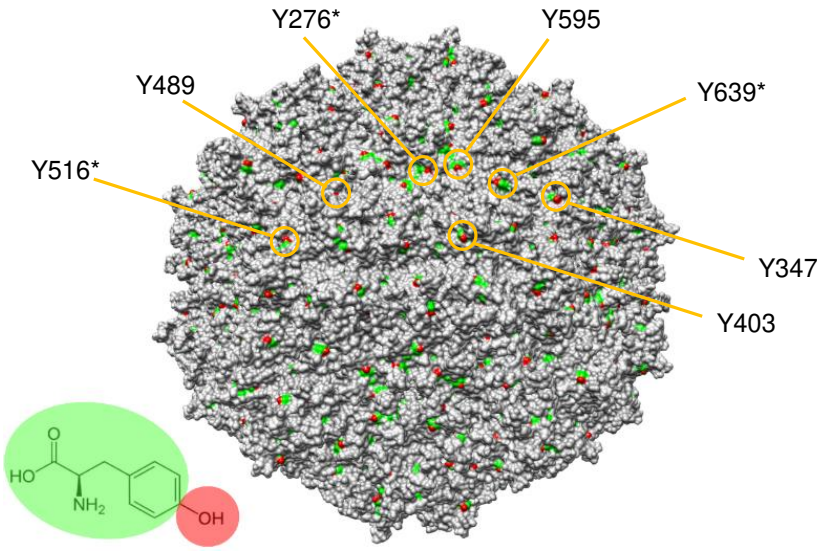
| | (d) nt | 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1


A
B
C
D
E
#171834
#171905
#171975
#171469
#171476
Gluc (ALU)
ns
T
S
Amino acid position 590
ciliated
goblet
3.3%
10.2%
4.3%
1.7%
9.0%
1.4%
FSC
β- Tubulin IV
MUC5AC
#171427
d 3
d 9
d 12
DQ000495
KPLMI-30
KPLMI-3530
NEG

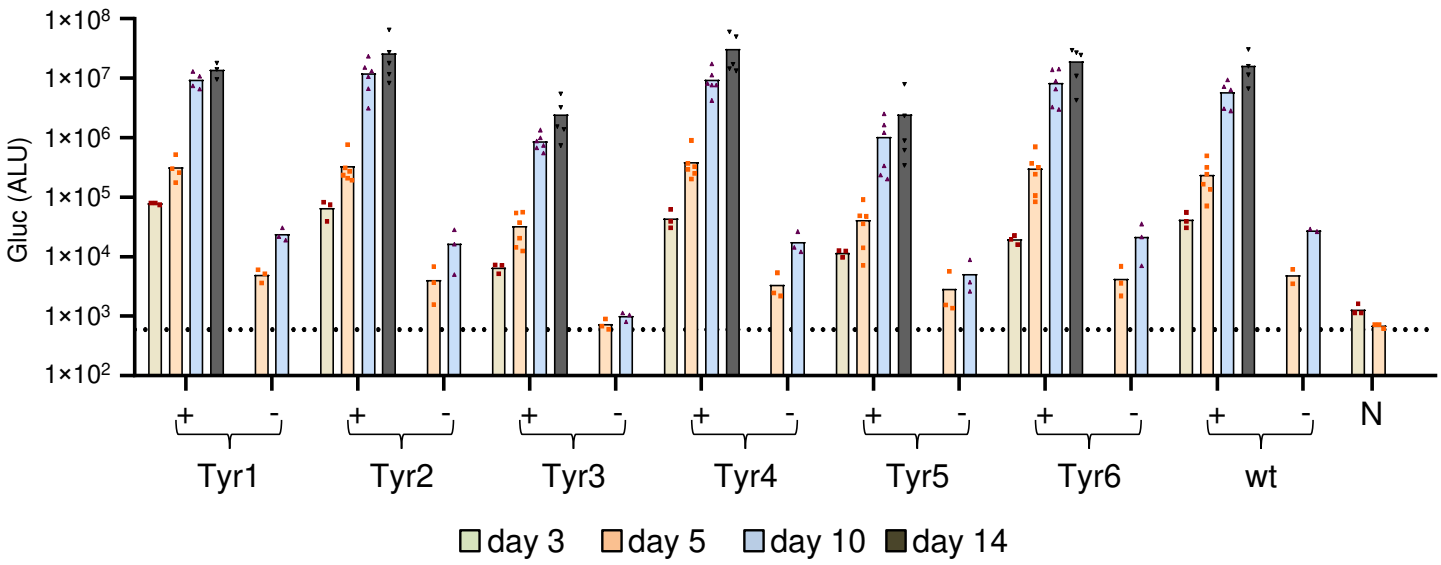
A

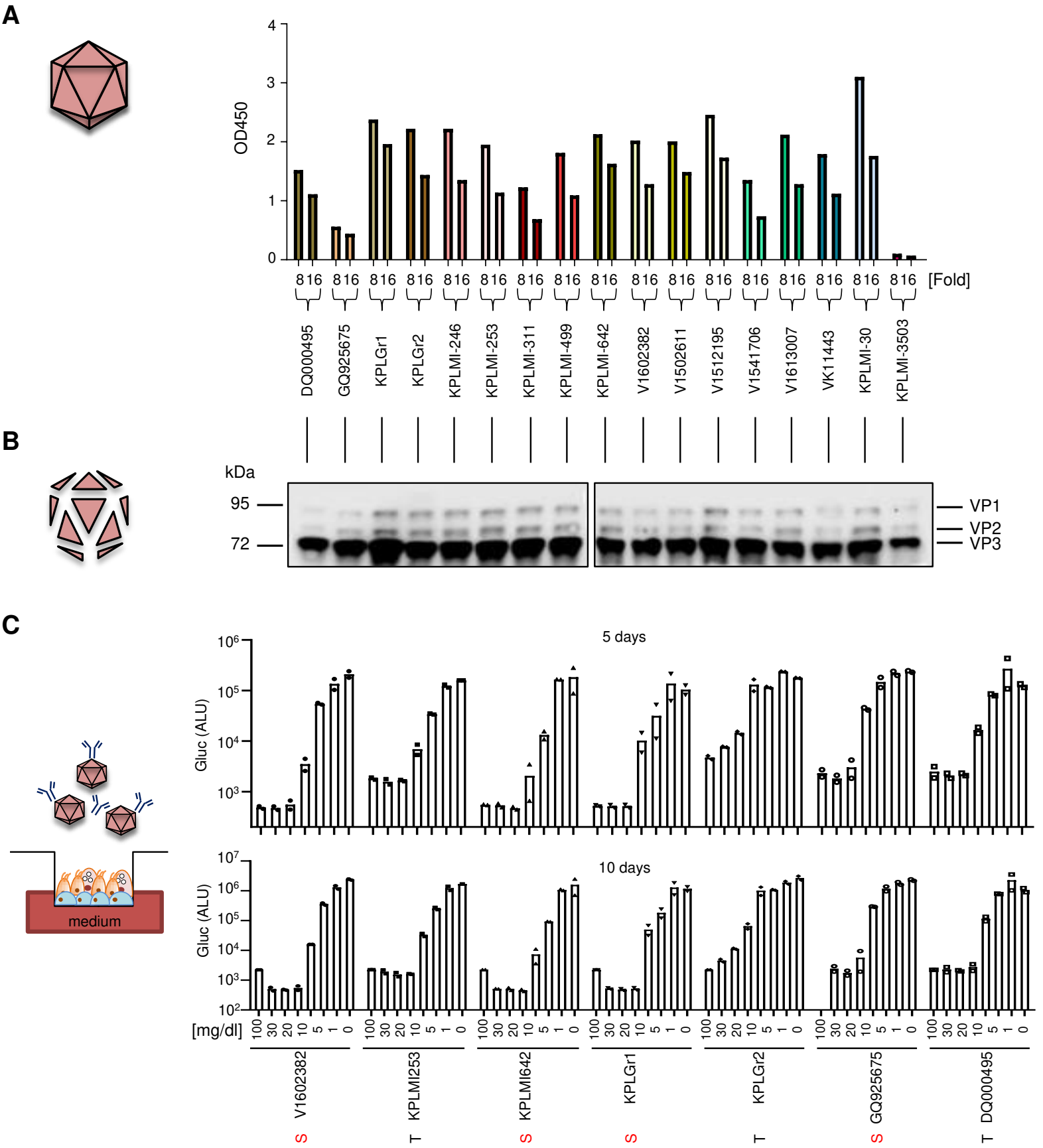
AA change	5′ → 3′ nucleotide
Tyr1 (Y276F)	CACAACATTCAACAAT
Tyr2 (Y403F)	AGTTCAGTTCATAAGA
Tyr3 (Y484F)	CCTAGAATTCAAACTT
Tyr4 (Y523F)	AACCACATTCAATCTA
Tyr5 (Y595F)	AGACTCATTCCTAAAC
Tyr6 (Y657F)	CACGTCATTCGATCAG

C



B

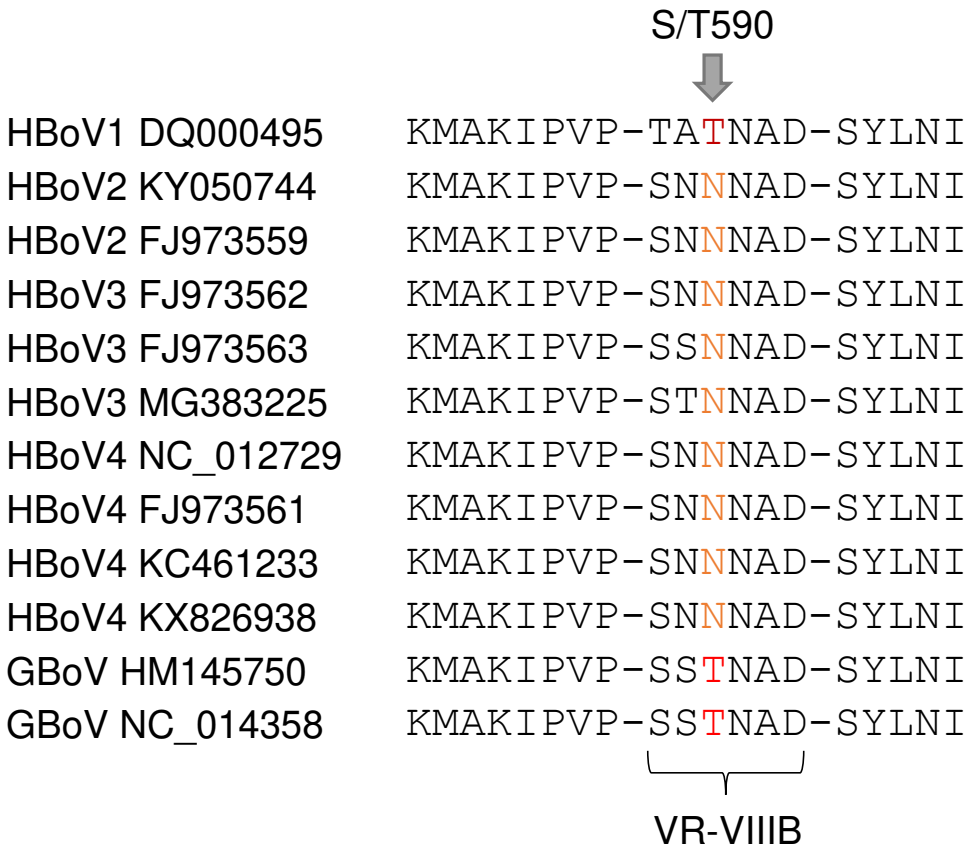



A
OD450
4
3
2
1
0
8 16 [Fold]
DQ000495
GQ925675
KPLGr1
KPLGr2
KPLMI-246
KPLMI-253
KPLMI-311
KPLMI-499
KPLMI-642
V1602382
V1502611
V1512195
V1541706
V1613007
VK11443
KPLMI-30
KPLMI-3503
B
kDa
95
72
VP1
VP2
VP3
C
medium
5 days
10 days
Gluc (ALU)
[mg/dl]
100 30 20 10 5 1 0
S V1602382
T KPLMI253
S KPLMI642
S KPLGr1
T KPLGr2
S GQ925675
T DQ000495

A

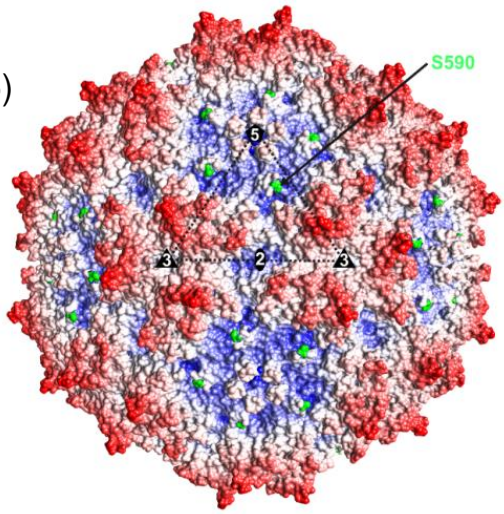
Method	positive	negative
MEME	1768-TCA-1770 (p=0.051); aa 590	none
SLAC	none	6
FUBAR	none	26
FEL	none	14

B

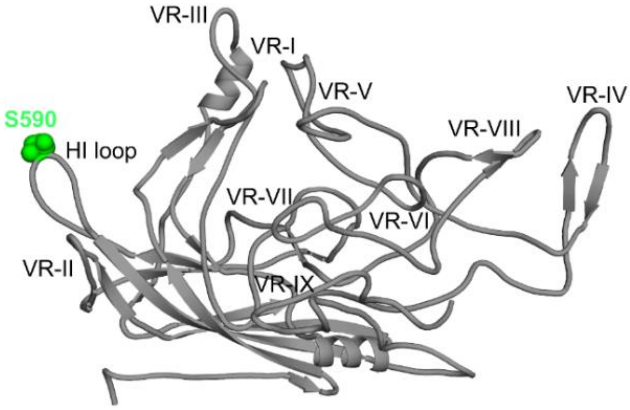


A

HBoV1
(GQ925675)

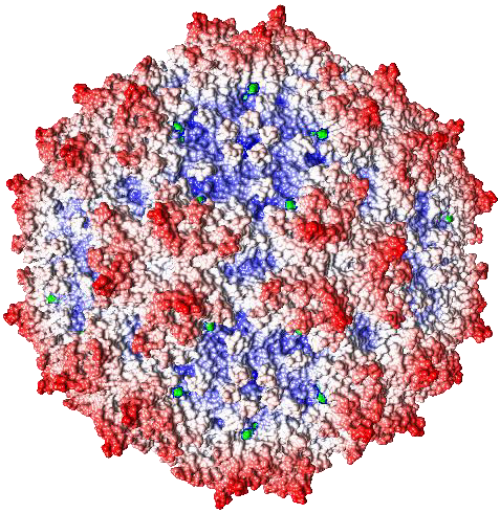


B

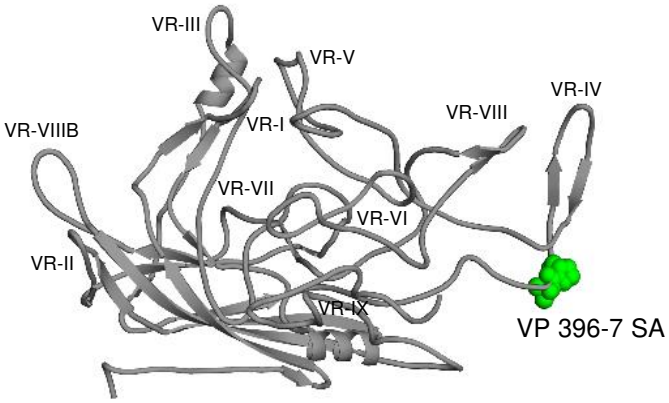


C

KPLMI-30

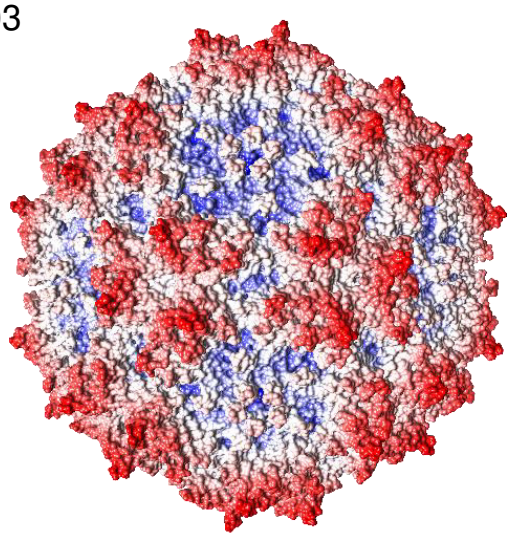
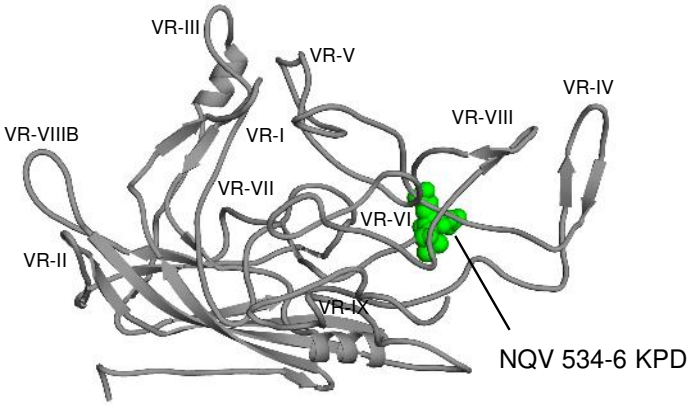


D



E

KPLMI-3503



F