Supplemental Fig. S1: Sample analysis timepoints.

Supplemental Text S2: Detailed methods

Supplemental Fig. S3: Horse D was persistently infected with the same EqHV population. No changes >5% from NZP1 consensus were found at any timepoint for horse D. Therefore, to interrogate whether low-level diversity could link sequential samples from the same horse, principle component analysis (PCA) of nucleotide diversity was performed. Positions with >0.5% divergence in early and late serum derived viral ORF sequences are included. Positions 5328, 5631, 9112, 2142, 1407, and 2751 contributed particularly to the separation of horse D (28 and 182 days post infection [dpi]) samples from the others.

Supplemental Fig. S4: Co-infection of EqHV and equine pegiviruses. Time course of viremia with EqHV and equine pegivirus (EPgV) is shown. Horse J was experimentally infected with EPgV-2 (TDAV) and horse R was naturally infected with EPgV-1 before enrollment in the current study. LLOQ: lower level of quantification.

Supplemental Table S5: Conserved sequences between EqHV and EPgV

Supplemental Fig. S6: Course of EqHV infection. (A) Serial serum chemistry and serum EqHV viral load for 26 weeks after i.v. inoculation is shown for all 7 horses. The time of seroconversion is indicated by an arrow. Horse D was persistently infected and was monitored for 30 weeks. **(B)** Correlation of serum miR-122 levels with AST during peak hepatitis evaluated by mixed effect analysis with horse as random effect. **(C)** Time course of serum miR-122 levels and other liver markers in two horses.

Supplemental Fig. S7: Anti-NS3 antibodies slowly wane over time and are boosted following productive rechallenge infections. Serial anti-NS3 antibody measurements by LIPS for four horses throughout the course of primary infection and two sequential rechallenge inoculations (indicated). Rechallenge inoculations that resulted in detectable viremia are indicated as productive. The first seropositive sample is indicated in red. RLU: relative light units.

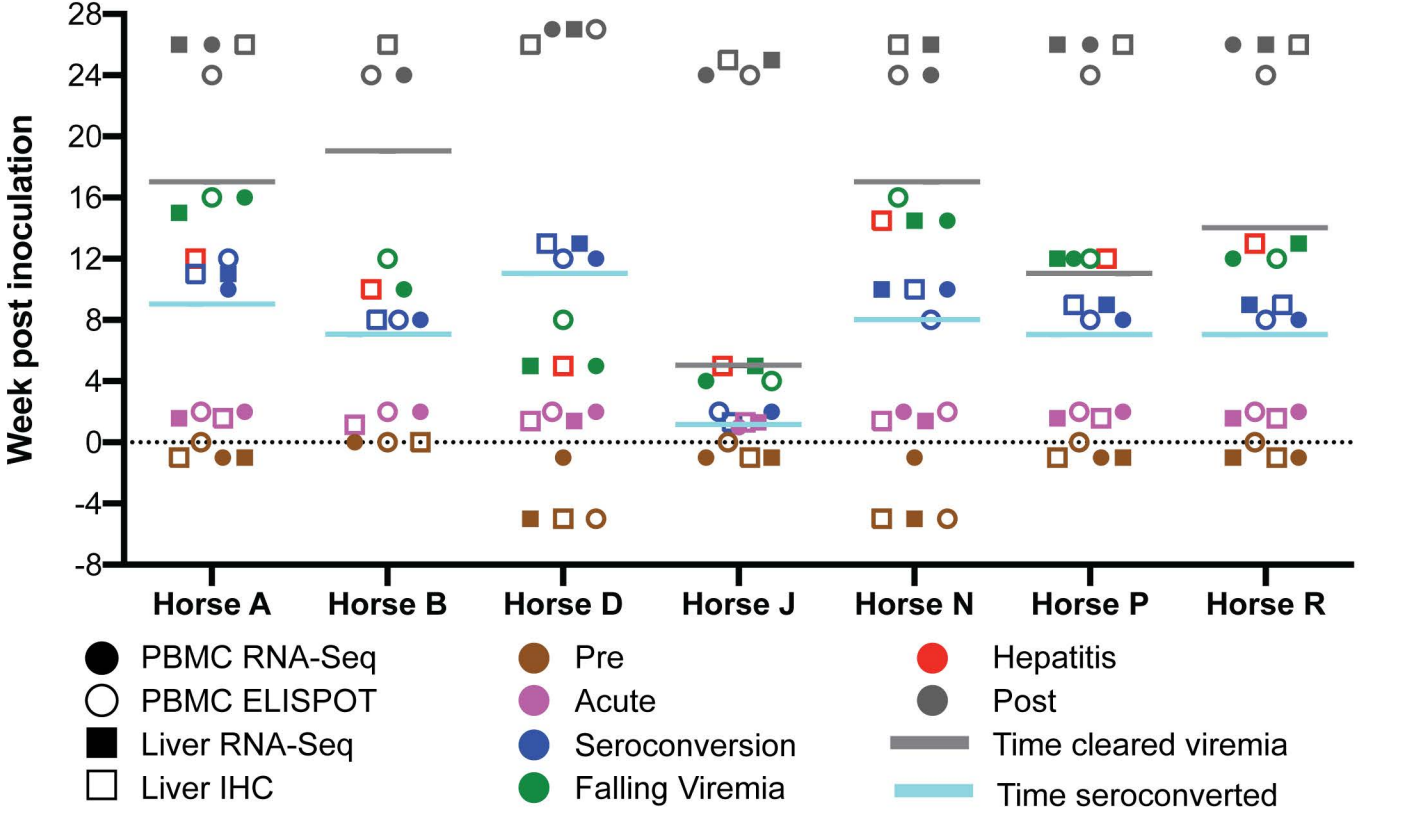
Supplemental Fig. S8: Viral evolution over the course of EqHV infection and sequence analysis of CU strain. (A-B) Viral genome diversity is shown for NZP1 (A) and CU (B) infections for ACUTE, FV, and rechallenge timepoints for remaining horses not shown in Fig. 3. Non-synonymous changes are in red, synonymous in grey, and changes to the UTRs in blue. Non-synonymous changes with a frequency >1% are labelled for NZP1 in (A). The dashed line indicates 50% (consensus change). **(C)** Sequence diversity in E1 and the N-terminus (proposed HVR1) of E2 for NZP1, the three CU sub-populations and for published sequences. Amino acid numbering according to NZP1 reference.

Supplemental Fig. S9: No change in PBMC populations during acute EqHV infection. (A) Complete blood counts performed monthly showed no notable changes in white blood cell counts. **(B-D)** Full time-course of PBMC subpopulations monitored by flow cytometry. **(B)** Major cell populations as percent of total PBMCs. **(C)** Subpopulations of T cells. **(D)** Proliferating cells (Ki67+) in each subpopulation of T cells. M, monocytes; NK, natural killer cells.

Supplemental Fig. S10: Liver histopathology. (A,B) Horse B pre-infection biopsy. Portal tracts were severely expanded by dense fibrous connective tissue. Prominent portal-to-portal bridging fibrosis was dispersed throughout the section. Sparse numbers of lymphocytes infiltrated the portal tracts. A mild to occasionally moderate sinusoidal lymphocytosis was dispersed throughout the parenchyma. Masson's trichrome stain showed large amounts of type I collagen accumulating in the portal tracts and bridging across lobules **(A)**. Reticulin staining accentuated the portal-to-portal bridging fibrosis and expansion of the portal tracts by extracellular matrix **(B)**. This horse had subclinical hepatitis with high GGT detected on screening serum biochemistry 7 months prior to enrollment in this study. Horse B was confirmed EqHV and EqPV-H serum RNA/DNA negative, and the hepatitis was presumptively attributed to Alsike clover (*Trifolium hybridum*) in the pasture. Hepatitis had resolved 5.5 months before enrollment, however as fibrosis is a common consequence of Alsike clover toxicity (42), this seems a likely explanation for horse B's pre-infection liver pathology. Fibrosis was also evident on horse B's ACUTE and HEP samples but not the POST sample, suggesting it might have been multifocal rather than diffuse. **(C,** HE**)** Horse N PRE liver showed a single focus of lymphocytes in a lobule, and **(D,** HE**)** moderate numbers of tightly packed small lymphocytes surrounding bile ducts in two portal tracts. Reticulin and Masson's Trichrome were within normal limits. **(E,F)** Horse J post liver biopsy showed fibrous extensions between portal tracts that were mildly accentuated with small amounts of accumulated type I collagen, which imparted a mild nodular appearance to the parenchyma **(E,** reticulin**)**. **(F,** HE**)** A mild lymphocytic infiltrate was present in the portal tracts and a mild sinusoidal lymphocytosis was dispersed throughout the parenchyma. HE, Hematoxylin and eosin; MT, Masson's trichrome.

Supplemental Fig. S11: Liver transcriptional responses reveal an ISG response.

Additional data from liver transcriptomic analyses. **(A)** Heat map of log-2-fold change (FC) in liver gene expression of genes differentially expressed at any timepoint across course of infection. EqHV (+/–) strand were significantly differentially expressed and are shown on a different scale. **(B)** Comparison of intrahepatic viral (+)RNA by RT-qPCR and RNA-seq analysis across timepoints. To preserve chronology, timepoints for horse D are plotted as PRE, ACUTE, FV, SC, FV. **(C)** Ratio between reads mapping to the viral (+)/(–) RNA strand at the ACUTE timepoint. **(D)** Gene ontology enrichment analysis for down-regulated genes in the liver at the ACUTE timepoint. **(E)** miRNA abundance profile from equine liver biopsies. Data are derived from AGO-CLIP miRNA chimeras, as these are less prone to bias compared to single miRNA reads. Data are averaged from liver biopsies of four horses, of which two were EPgV-2 infected (liver unrelated) (20), one was acutely infected and one persistently infected with EqHV (13). **(F-I)** Cumulative density function (CDF) of the log-2 FC in liver gene expression between timepoints as indicated. mRNAs are grouped by presence of miRNA 7- or 8-mer seed site in the 3'UTR.


Week post inoculation
28
24
20
16
12
8
4
0
-4
-8
Horse A
Horse B
Horse D
Horse J
Horse N
Horse P
Horse R
PBMC RNA-Seq
PBMC ELISPOT
Liver RNA-Seq
Liver IHC
Pre
Acute
Seroconversion
Falling Viremia
Hepatitis
Post
Time cleared viremia
Time seroconverted

S2 Text: Detailed methods

Serum biochemistry and hematology

Fresh serum samples were submitted to the New York State Animal Health Diagnostic Center (AHDC) for serum biochemistry of liver markers. Complete blood count was performed on fresh EDTA-anticoagulated blood monthly by the AHDC.

Hemogram parameters and reference intervals were: hematocrit, 34 – 46%; hemoglobin, 11.8 – 15.9 g/dL; red blood cell count, 6.6 – 9.7 $x10^6$/µL; mean corpuscular volume, 43 – 55 fL; mean corpuscular hemoglobin, 15 – 20 pg; mean corpuscular hemoglobin concentration 34 – 37 g/dL; red cell distribution width, 16.3 – 19.3%; nucleated red blood cells, 0/100 white blood cells; white blood cells, 5.2 – 10.1 $x10^3$/µL; segmented neutrophils, 2.7 – 6.6 $x10^3$/µL; band neutrophils 0.0 – 0.1 $x10^3$/µL; lymphocytes, 1.2 – 4.9 $x10^3$/µL; monocytes, 0.0 - 0.6 $x10^3$/µL; eosinophils, 0.0 – 1.2 $x10^3$/µL; basophils, 0.0 – 0.2 $x10^3$/µL; platelet count, 94 – 232 $x10^3$/µL; mean platelet volume, 5.3 – 8.4 fL; total protein, 5.2 – 7.8 g/dL. Blood smears were examined manually to confirm automated results.

Biochemical markers examined were: aspartate aminotransferase (AST), sorbitol dehydrogenase (SDH), glutamate dehydrogenase (GLDH), gamma glutamyltransferase (GGT), bile acids, total, direct, and indirect bilirubin, creatine kinase (CK), triglycerides, serum amyloid A (SAA), serum iron, total iron binding capacity (TIBC), and ferritin (FE) saturation. Reference intervals were: AST, 222-489 U/L; SDH, 1-6 U/L; GLDH, 2-10 U/L; GGT, 8-33 U/L; bile acids, 2-10 µmol/L; total bilirubin, 8.55-35.9 µmol/L; direct bilirubin, 1.71-5.13 µmol/L; indirect bilirubin, 5.13-34.2 µmol/L; CK, 171-567 U/L; triglycerides, 14-65 mg/dL; SAA, 0-8 µg/mL; serum iron, 95-217 µg/dL; TIBC, 289-535 µg/dL; and FE saturation, 27-56%.

Serum EqHV and miR-122 PCR and serology

For viremia, RNA was extracted from serum using MagMAX-96 Viral RNA Isolation Kit (Ambion) and viral load was quantified by RT-qPCR using TaqMan Fast Virus 1-Step Master Mix (Applied Bioscience) on a LightCycler (Roche) with the following cycling parameters: 50°C for 30 min, 95°C for 5 min followed by 40 cycles of 95°C for 15 s, 56°C for 30 s and 60°C for 45 s. Primers/probes were TS-O-00891/892/893 for EqHV, TS-O-00373/374/375 for EPgV-1 and TS-O-00130/131/132 for EPgV-2. For miRNAs, RNA was extracted using Trizol and cDNA prepared using the miScript kit with HiSpec buffer (Qiagen) for 1 hr at 37°C and 5 min at 95°C. Takara One step SYBR primescript RT-PCR (perfect RT) kit was used for qPCR by excluding the Primescript RT enzyme mix II and using primers TS-O-00226 (miR-122 specific) and TS-O-00309 (miScript Universal primer). Samples were diluted 10-fold before qPCR, and a synthetic miR-122 standard was included. Anti-NS3 antibody levels were determined by LIPS as previously described (1).

Enzyme-linked immunosorbent spot assay (ELISPOT)

Antigen-specific T-cells secreting IFN-γ were enumerated using equine ELISpot kit (Mabtech). Cells were plated in 96 well PVDF plates at $2x10^5$ per well in duplicate, and stimulated separately with three peptide pools comprised of peptides of 18 amino acids with 11 residue overlap covering the complete NS3 coding sequence of EqHV reference strain NZP1 (GenBank NC_038425), or Concanavalin-A 5 µg/mL (Sigma), or media alone, as positive and negative controls, respectively. The plates were incubated for 42-48 hours and developed according to instructions. The number of spot-forming cells were measured using an automatic counter (Immunospot). A positive response was considered only when the mean of peptides-stimulated wells was more than mean of negative wells + 3 standard deviation. The total number of spot-forming cells were calculated by subtracting the mean number of spots in negative control wells from the mean of peptides containing wells.

Flow cytometric phenotyping

Fresh PBMC were labeled for flow cytometric phenotyping. For 2 horses (Horses A, R), PBMC were cryopreserved and flow cytometry was performed in batch after panel development was finalized. Others were done fresh. Two antibody panels, Panel M and Panel T, were used and analyzed as previously described (2). Briefly, Panel M identified populations of T cells (CD3+CD14-CD16-), B cells (CD3-CD14-PanIg+), NK-like cells (CD3+CD14-CD16+), monocytes (CD3-CD14+) and alternatively activated monocytes (CD3-CD14+CD16+). Panel T resolved populations of CD4 T cells (CD3+CD14-CD21-CD4+CD8-), CD8 T cells (CD3+CD14-CD21-CD4-CD8+), and Ki67+ proliferating CD4 and CD8 T cells.

Liver histopathology

Slide preparation and labeling was performed by the AHDC. Sections were labeled with HE, Masson's Trichrome, and reticulin. Immunohistochemistry was performed to detect T cells (anti-CD3, clone LN10, Leica, Buffalo Grove, IL, USA), B cells (anti-Pax5, clone 1EW, Leica, Buffalo Grove, IL, USA), and macrophages/dendritic cells (m/DC) (anti-Iba1, Wako Pure Chemical Industries, Richmond, VA, USA). Slides were read by one author (S.M.) who was blinded to horse and infection status. The number of T cells in sinusoids was counted over five 40x fields and an average number of cells per mm^2 was calculated as previously described (3). The number of T cells per portal tract was averaged over 5 portal tracts.

RNA isolation, quantification, and sequencing

For complete ORF amplification, RNA was extracted from 250 µL serum using TRIzol LS Reagent (Thermo Fisher Scientific). After addition of chloroform and centrifugation, the aqueous phase was mixed with 450 µL anhydrous ethanol and transferred to an RNA Clean & Concentrator-5 column (Zymo Research) for downstream RNA purification and concentration. RT was performed with Maxima H Minus Reverse Transcriptase (Thermo Fisher Scientific). Samples were pre-incubated in the presence of RNase inhibitors (Promega) at 65°C for 2 min prior to addition

of the RT enzyme, followed by 50°C for 2 h using 0.1 μM primers NPHV-9379-RT and CU-9379-RT for NZP1 and CU strain, respectively. Amplification of the complete ORF was performed using Q5 High-Fidelity Hot start DNA Polymerase (NEB), including high GC Enhancer, and the primer pairs NPHV-288-F/NPHV-9314-R and TS-O-01230/NPHV-9314-R for NZP1 and CU strain, respectively. PCR cycling parameters were 98°C for 30 s, followed by 35 cycles of 98°C for 10 s, 65°C for 10 s and 72°C for 8 min, with a final extension at 72°C for 10 min. Amplified DNA was purified using DNA Clean & Concentrator 25 columns (Zymo Research), and libraries for deep sequencing were prepared using the NEBNext Ultra II FS DNA Library Prep Kit for Illumina (NEB). Quality of DNA libraries was validated using a 2100 Bioanalyzer Instrument (Agilent). The Qubit dsDNA High-Sensitivity Assay Kit (Thermo Fisher Scientific) was used to quantify DNA libraries in order to pool these in equimolar concentrations prior to denaturation. Pooled libraries were loaded on a MiSeq v3 150 cycle flow-cell, and sequencing performed on a MiSeq benchtop sequencer (Illumina).

Library preparation and RNA-seq for equine PBMC and liver samples

PBMCs were lysed in TRIzol and total RNA was extracted as described above. The integrity of extracted RNA was analyzed on a 2100 bioanalyzer instrument using an RNA 6000 nano assay kit (Agilent Technologies). High-quality RNA (avg. RIN score 9.8, min. 9.0) was quantified using the Qubit RNA broad-range assay kit and 500 ng were used as input for TruSeq stranded mRNA library preparation (Illumina).

Liver biopsies (~10 mg) fixed in RNAlater were transferred into a MagNA Lyser Green Beads tube containing 1 mL cold TRIzol. Liver tissue was homogenized using a MagNA Lyser instrument (Roche) at 6500 rpm for 60 s with continuous cooling on ice every 20 s. RNA extraction was done as outlined above and the integrity was assessed on a 2100 bioanalyzer instrument (some samples N/A, avg. RIN score of others 7.9). For the liver samples, total RNA-seq was prioritized over poly(A) selection in order to recover viral reads. Depletion of ribosomal RNA from 500 ng input RNA was achieved using the Ribo-off rRNA depletion kit (Vazyme) for which horse rRNA removal has been validated. Following purification of rRNA-depleted total RNA using RNAClean XP beads (Beckman Coulter), the RNA was eluted with FPF buffer (Illumina) for subsequent TruSeq stranded mRNA library preparation that, except for the use of FPF buffer, is identical to the total RNA protocol, starting from the RNA fragmentation step.

Quality of DNA libraries was checked using a 2100 bioanalyzer instrument. The Qubit dsDNA High-Sensitivity Assay Kit (Thermo Fisher Scientific) was used to quantify DNA libraries in order to pool these in equimolar concentrations prior to denaturation. Pooled libraries were loaded on a NextSeq 500/550 v2.5 75 cycle flow-cell, and sequencing performed on a NextSeq benchtop sequencer (Illumina) at the Department of Clinical Microbiology (Hvidovre Hospital, Copenhagen). Data derived from two (PBMCs) or three (liver) independent deep-sequencing runs were pooled to ensure sufficient depth of sequencing coverage.

RNA-seq data analysis

For transcriptome analysis, reads were mapped to the EquCab3.0 genome (Ensembl v98) supplemented with the EqHV NZP1 genome (GenBank: KP325401.1) with HISAT2 (4) and quantified with featureCounts (5) including positive and negative strand annotation of NZP1 using default settings. Previous equine transcriptome analysis (6) identified duplicate entries for numerous immunologically relevant genes (CD14, IL7R, etc.) in the EquCab3.0 reference. To recover quantification of genes with duplicate entries (n = 77 genes, which were discarded in initial quantifications due to read multimapping), featureCounts was run a second time on a filtered reference containing only genes with duplicate entries; in this stage, multimapped reads were counted fractionally (each alignment carries 1/*x* counts where *x* is the total number of alignments reported for the read) and then collapsed at the gene level by summation (featureCounts with -m and –fraction explicitly). The resulting counts matrix for genes with duplicate entries was appended to the initial count matrix. Differential expression was analyzed in R using Limma and Voom. The PRE, ACUTE, SC, FV and POST time points were included in analysis. Horse D was excluded due to persistent infection. Horse B was excluded from liver but not PBMC analysis for technical reasons. Gene ontology analysis was done in R using GoSeq. miRNA abundance was determined as the fraction of chimeric reads for each miRNA in Argonaute (AGO) cross-linking immunoprecipitation (CLIP) libraries from four equine liver biopsies (7,8). For analysis of miR-122 de-repression, miRNA targeting of mRNAs were defined by presence of 7-mer seed sites in EquCab3.0 3'UTRs.

Library preparation and data processing for scRNA-seq

scRNA-seq was performed using 10x Genomics Chromium Single Cell 3' Reagent Kit (v2). Libraries were pooled and sequenced on Illumina NextSeq 500 in paired-end configuration (Read 1, cell barcode: 26 nt; Read 2, transcript: 98 nt) to a target read depth of approximately 35,000 paired-end reads per cell. The EquCab3.0 reference genome (9) was used in all scRNA-seq analyses. Reads were assigned to cell barcodes, mapped and quantified per gene using CellRanger (v 3.0.1, 10X Genomics) with default parameters ("standard workflow"). Raw BAM files were extracted and processed with the End Sequence Analysis Toolkit (10) and a workflow optimized for equine scRNA-seq analysis as described (6). Putative "multiplet" cell barcodes were identified and removed from downstream analyses with the DoubletDetection tool (11).

scRNA-seq data analysis

Processed gene-cell matrices were analyzed in R (v3.5.1) using the Seurat package (v3.1.1). Data were filtered to exclude genes detected in less than 3 cells (per sample timepoint; PRE, ACUTE), to exclude cells with less than 750 unique molecular identifiers (UMIs) or greater than 20,000 UMIs (putative doublets), and to exclude cells with greater than 5% UMIs assigned to mitochondrial genes (putative dead or dying cells). Gene-cell count matrices were independently normalized with SCTransform (12), and the top 5000 most variable genes (variance-stabilizing transformation) were selected for each sample timepoint. Data for the PRE timepoint were previously analyzed as part of establishing a reference dataset of equine PBMCs (6); assigned cell

type annotations were used in the present analysis. To classify cell types in the ACUTE timepoint, we projected reference cell type annotations using Seurat's FindTransferAnchors and TransferData function (dims: 1:30; weight.reduction: "pcaproject") (13). For UMAP visualizations, data from both timepoints were integrated using Seurat's FindIntegrationAnchors and IntegrateData function (dims: 1:30), and cells colored by reference annotation (PRE timepoint) or reference-based classification (ACUTE timepoint).

Statistics

Statistical tests of clinical parameters were performed in GraphPad Prism version 8.3.0 for Mac (GraphPad Software, San Diego, CA, USA) or JMP Pro 14.0.0 (SAS Institute Inc., Cary, NC, USA). Mixed effects analysis with horse as a random effect and Dunnett's post hoc tests were performed. Significance was set at $p < 0.05$.

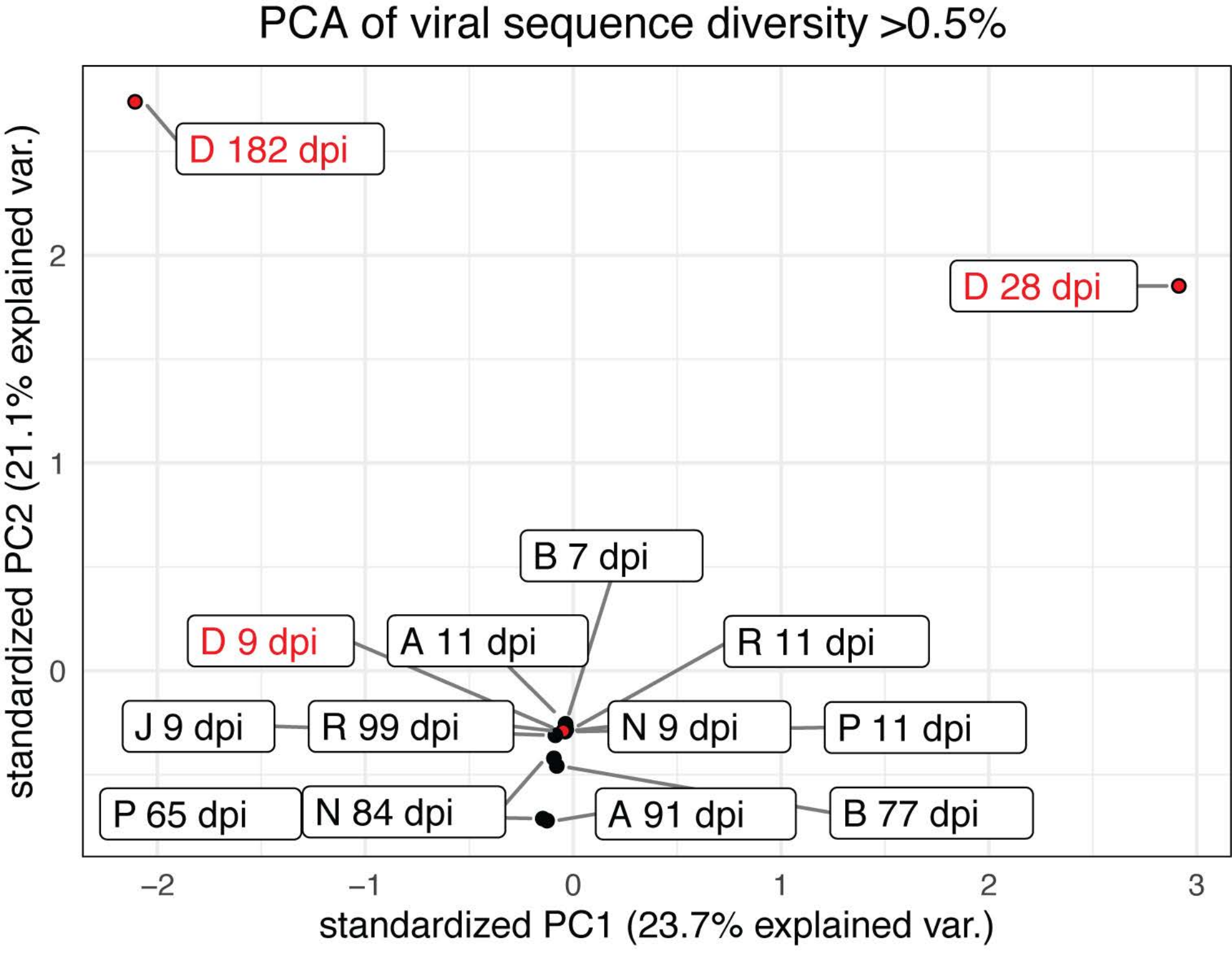
Primer sequences

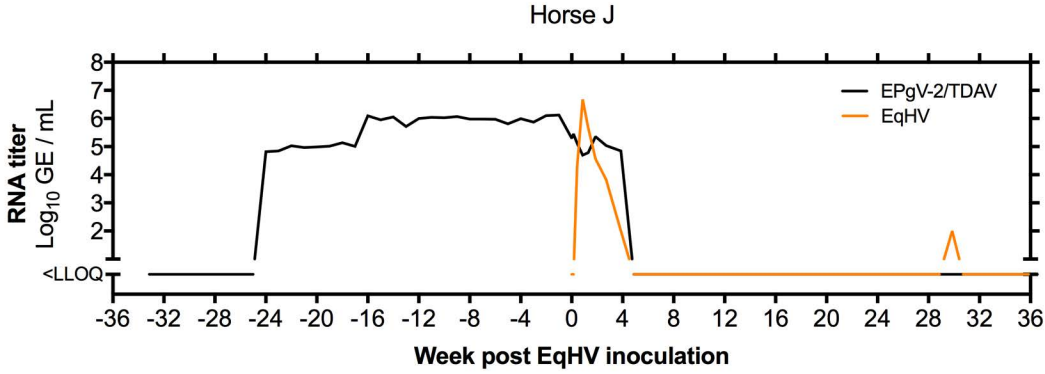
Primer no.	Primer name	Primer sequence
1	TS-O-00130	CCCAAACCGAGCCRCCCT
2	TS-O-00131	CGGACTGAATTATAGGCGTCG
3	TS-O-00132	/FAM/CCGGGATTTACCCGAAGAACCCTG/BHQ_1/
4	TS-O-00226	TGGAGTGTGACAATGGTGTTT
5	TS-O-00309	GAATCGAGCACCAGTTACGC
6	TS-O-00373	ACGCAGAGCAAGATTACCTATGC
7	TS-O-00374	CCTGGTGGAGTAGCAGTAGC
8	TS-O-00375	/FAM/ACGCTGACGTCGTGATTTGCGACGA/BHQ_1/
9	TS-O-00891	CACATCACCATGTGTCACTCC
10	TS-O-00892	CGCGATTTTCGTGTACTCAC
11	TS-O-00893	/FAM/TCACGAATTCCAGCTCCCT/BHQ_1/
12	TS-O-01230	ATCGCGGCTTGAACGTCCTA
13	NPHV-288-F	CACGAAGGAAGGCGGGGG
14	NPHV-9314-R	CCATCACCCACCCTCCTGTT
15	NPHV-9379-RT	CCATAGGGGCGGAACAGG
16	CU-9379-RT	CCATAGGGGCAGAACAGG

References

1. Scheel TKH, Kapoor A, Nishiuchi E, Brock K V., Yu Y, Andrus L, et al. Characterization of nonprimate hepacivirus and construction of a functional molecular clone. Proc. Natl. Acad. Sci. U. S. A. [Internet]. 2015;112:2192–2197. Available from: http://www.pnas.org/lookup/doi/10.1073/pnas.1500265112
2. **Tomlinson JE, Wolfisberg R**, Fahnøe U, Sharma H, Renshaw RW, Nielsen L, et al. Equine pegiviruses cause persistent infection of bone marrow and are not associated with hepatitis. PLoS Pathog. [Internet]. 2020;16:e1008677. Available from: https://doi.org/10.1371/journal.ppat.1008677
3. Meuten DJ, Moore FM, George JW. Mitotic Count and the Field of View Area: Time to Standardize. Vet. Pathol. 2016;53:7–9.
4. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. [Internet]. 2019;37:907–915. Available from: http://dx.doi.org/10.1038/s41587-019-0201-4
5. Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923–930.
6. **Patel RS, Tomlinson JE,** Divers TJ, van de Walle GR, Rosenberg BR. Single-cell resolution landscape of equine peripheral blood mononuclear cells reveals diverse cell types including T-bet+ B cells. BMC Biol. 2021;19:1–18.
7. Yu Y, Scheel TKH, Luna JM, Chung H, Nishiuchi E, Scull MA, et al. miRNA independent hepacivirus variants suggest a strong evolutionary pressure to maintain miR-122 dependence. PLoS Pathog. 2017;13:1–21.
8. Scheel TKH, Moore MJ, Luna JM, Nishiuchi E, Fak J, Darnell RB, et al. Global mapping of miRNA-target interactions in cattle (Bos taurus). Sci. Rep. [Internet]. 2017;7:1–13. Available from: http://dx.doi.org/10.1038/s41598-017-07880-8
9. Kalbfleisch TS, Rice ES, DePriest MS, Walenz BP, Hestand MS, Vermeesch JR, et al. Improved reference genome for the domestic horse increases assembly contiguity and composition. Commun. Biol. 2018;1:1–8.
10. Derr A, Yang C, Zilionis R, Sergushichev A, Blodgett DM, Redick S, et al. End Sequence Analysis Toolkit (ESAT) expands the extractable information from single-cell RNA-seq data. Genome Res. 2016;26:1397–1410.
11. Gayoso A, Shor J, Carr AJ, Sharma R, Pe'er D. GitHub: DoubletDetection. 2019;
12. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019;20:1–15.
13. **Stuart T, Butler A**, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive Integration of Single-Cell Data. Cell [Internet]. 2019;177:1888-1902.e21. Available from: https://doi.org/10.1016/j.cell.2019.05.031

Author names in bold designate shared co-first authorship.


PCA of viral sequence diversity >0.5%
D 182 dpi
D 28 dpi
B 7 dpi
D 9 dpi
A 11 dpi
R 11 dpi
J 9 dpi
R 99 dpi
N 9 dpi
P 11 dpi
P 65 dpi
N 84 dpi
A 91 dpi
B 77 dpi
standardized PC2 (21.1% explained var.)
2
1
0
−2
−1
0
1
2
3
standardized PC1 (23.7% explained var.)


Horse J
RNA titer
Log10 GE / mL
8
7
6
5
4
3
2
<LLOQ
EPgV-2/TDAV
EqHV
-36 -32 -28 -24 -20 -16 -12 -8 -4 0 4 8 12 16 20 24 28 32 36
Week post EqHV inoculation

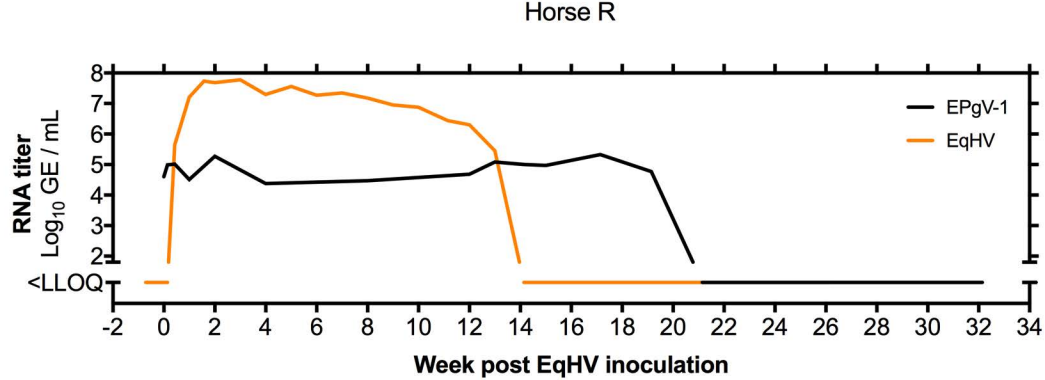
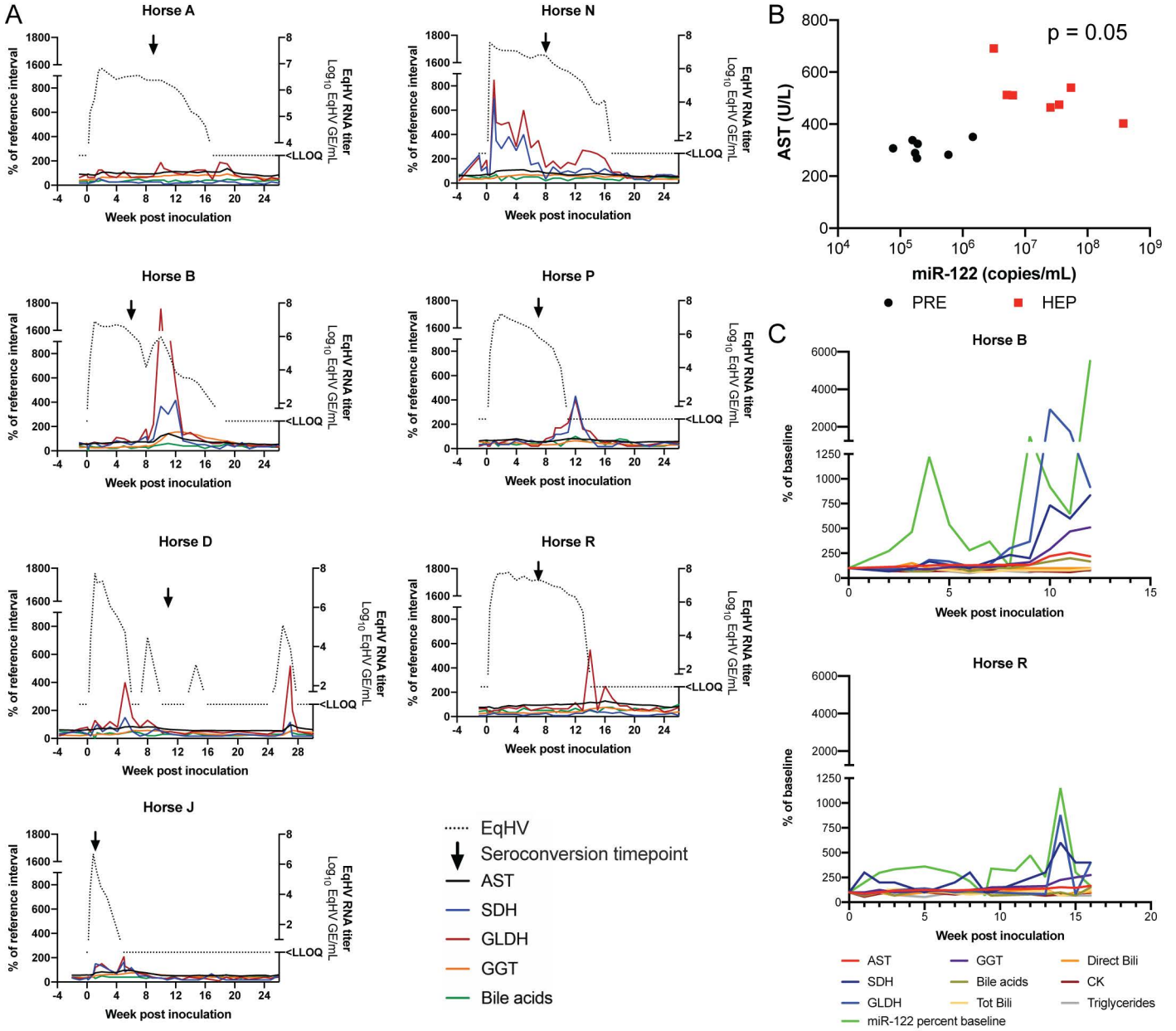
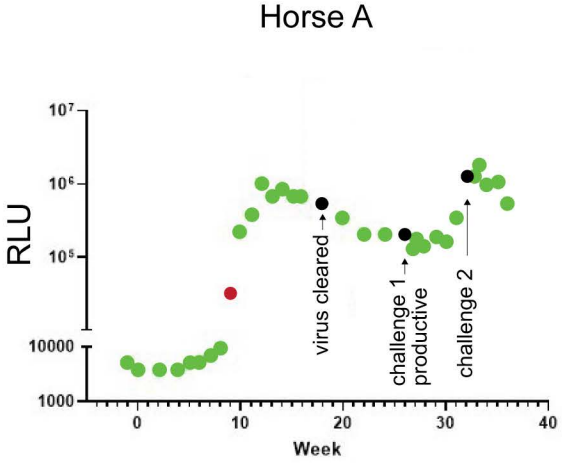
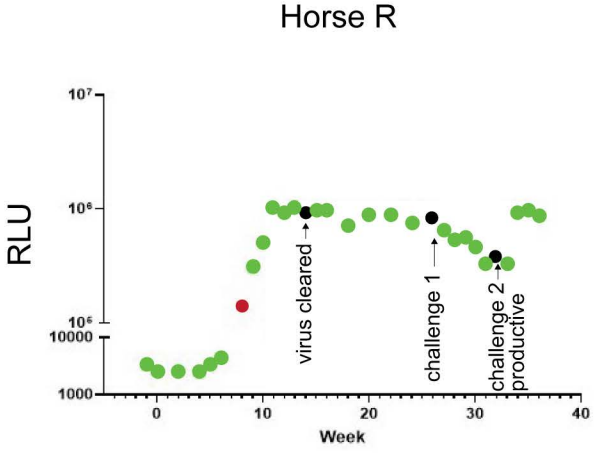

Horse R
RNA titer
Log10 GE / mL
8
7
6
5
4
3
2
<LLOQ
EPgV-1
EqHV
-2 0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34
Week post EqHV inoculation

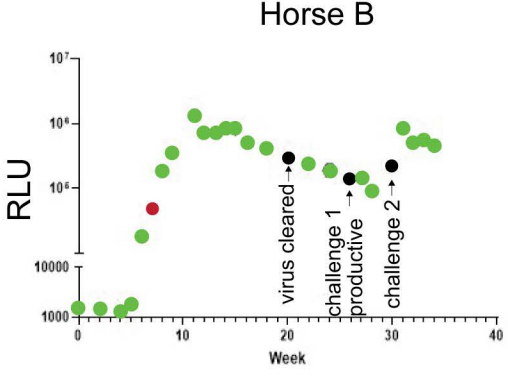
Table S5: Conserved epitopes between NPHV, EPgVs. Completely conserved epitopes of at least 7 amino acid residues between EqHV (NZP1) and EPgV-1 (C35) and EPgV-2 (TDAV) are shown.

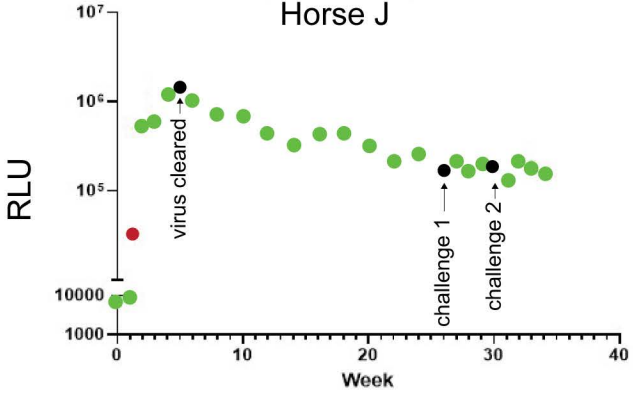
Epitope	NZP1 pos	Gene	EPgV-1 (C35)	EPgV-2 (TDAV)
PTGSGKST	1211	NS3	x	
VLVLNPSVA	1231	NS3	x	
LVLNPSVA	1232	NS3		x
ATATPPG	1327	NS3	x	
GDIPFYG	1351	NS3	x	x
VVATDAL	1414	NS3		x
GNFDTVTDC	1426	NS3	x	
FDTVTDCN	1428	NS3		x
MPPLEGE	2328	NS5A		x
AYGFQYTP	2542	NS5B		x

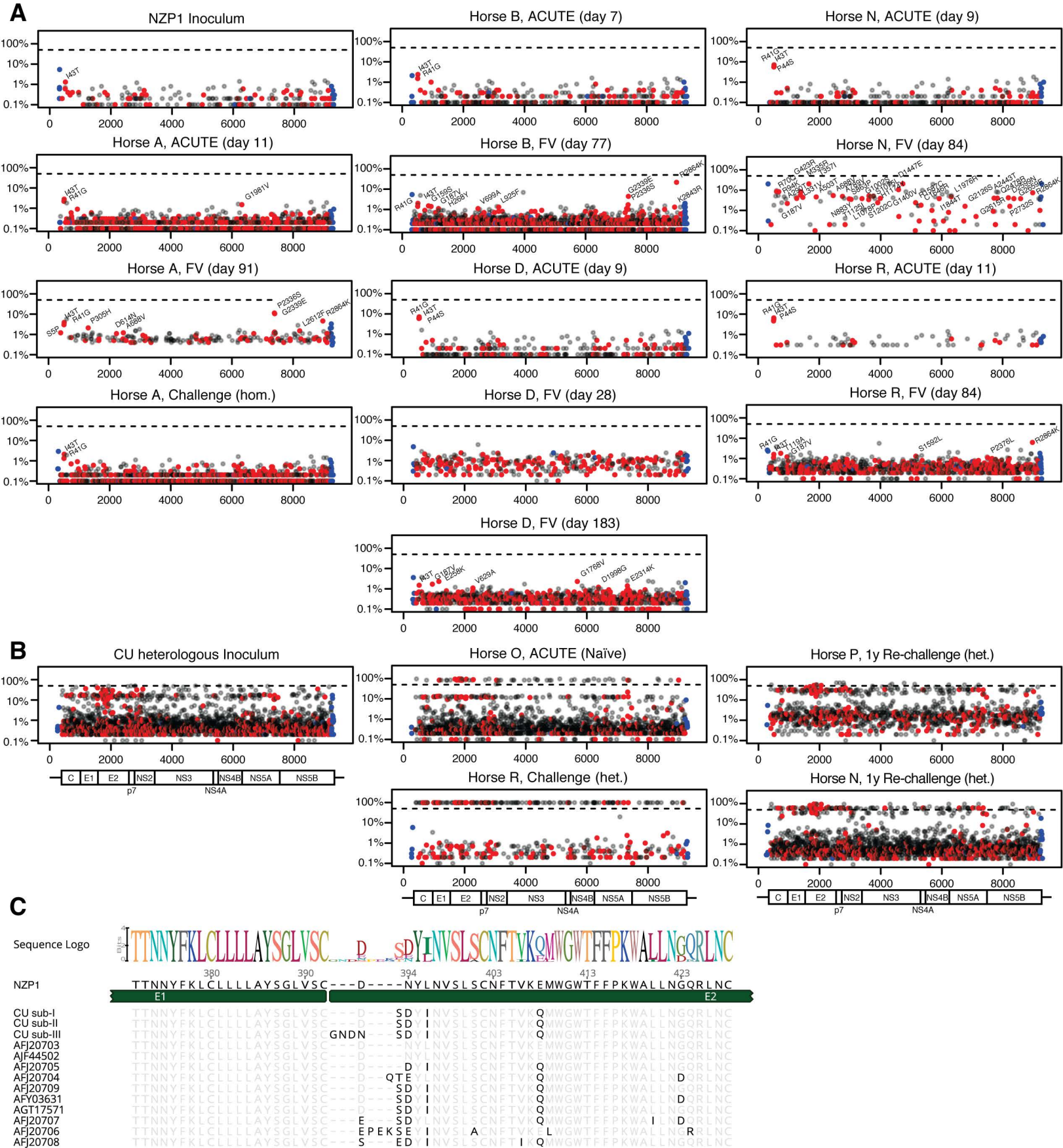

A
Horse A
Horse N
Horse B
Horse P
Horse D
Horse R
Horse J
% of reference interval
EqHV RNA titer
Log10 EqHV GE/mL
<LLOQ
Week post inoculation
EqHV
Seroconversion timepoint
AST
SDH
GLDH
GGT
Bile acids
B
p = 0.05
AST (U/L)
miR-122 (copies/mL)
PRE
HEP
C
Horse B
Horse R
% of baseline
Week post inoculation
AST
SDH
GLDH
GGT
Bile acids
Tot Bili
Direct Bili
CK
Triglycerides
miR-122 percent baseline

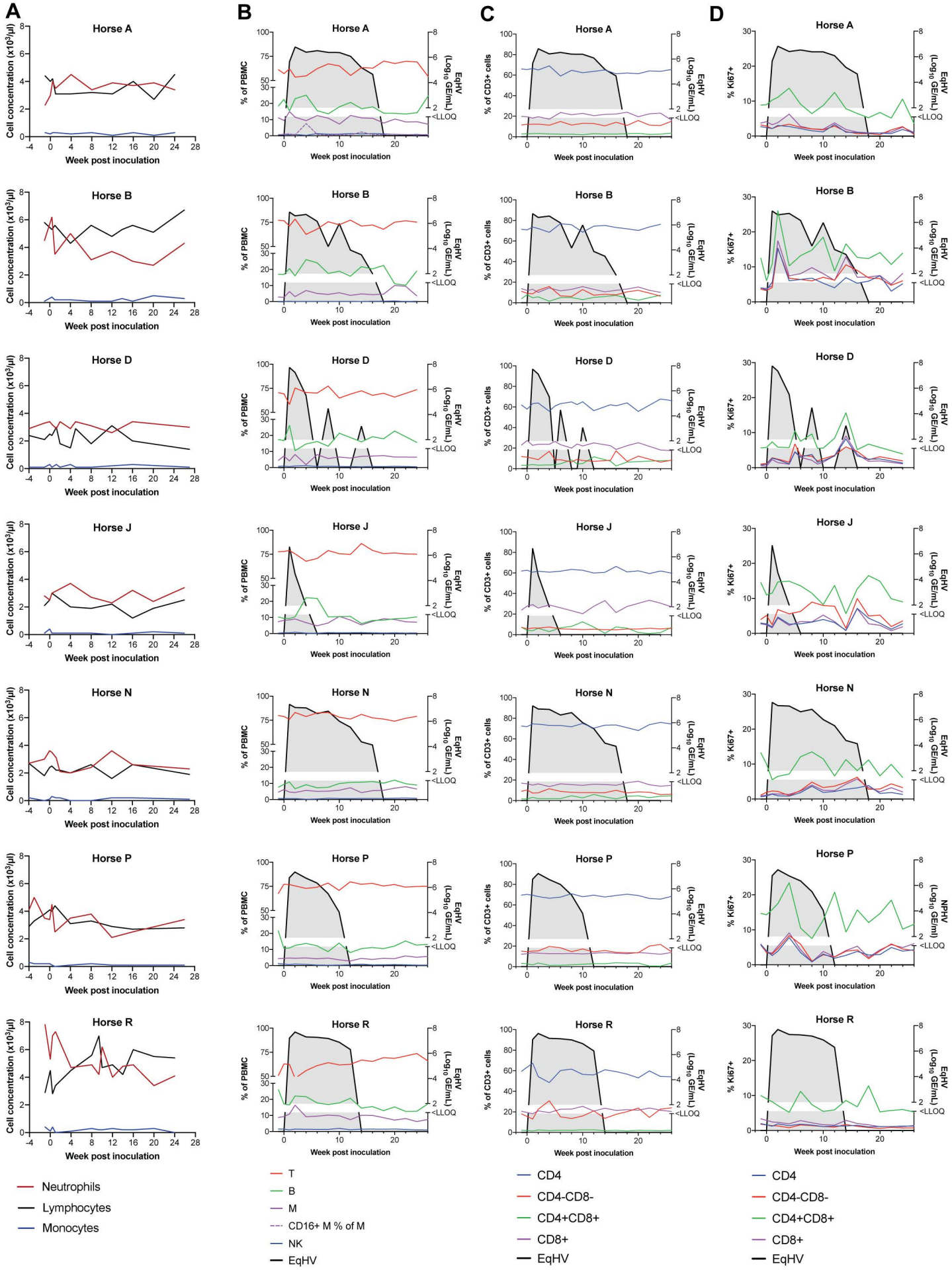

Horse A
RLU
10^7
10^6
10^5
10000
1000
virus cleared
challenge 1
productive
challenge 2
0
10
20
30
40
Week

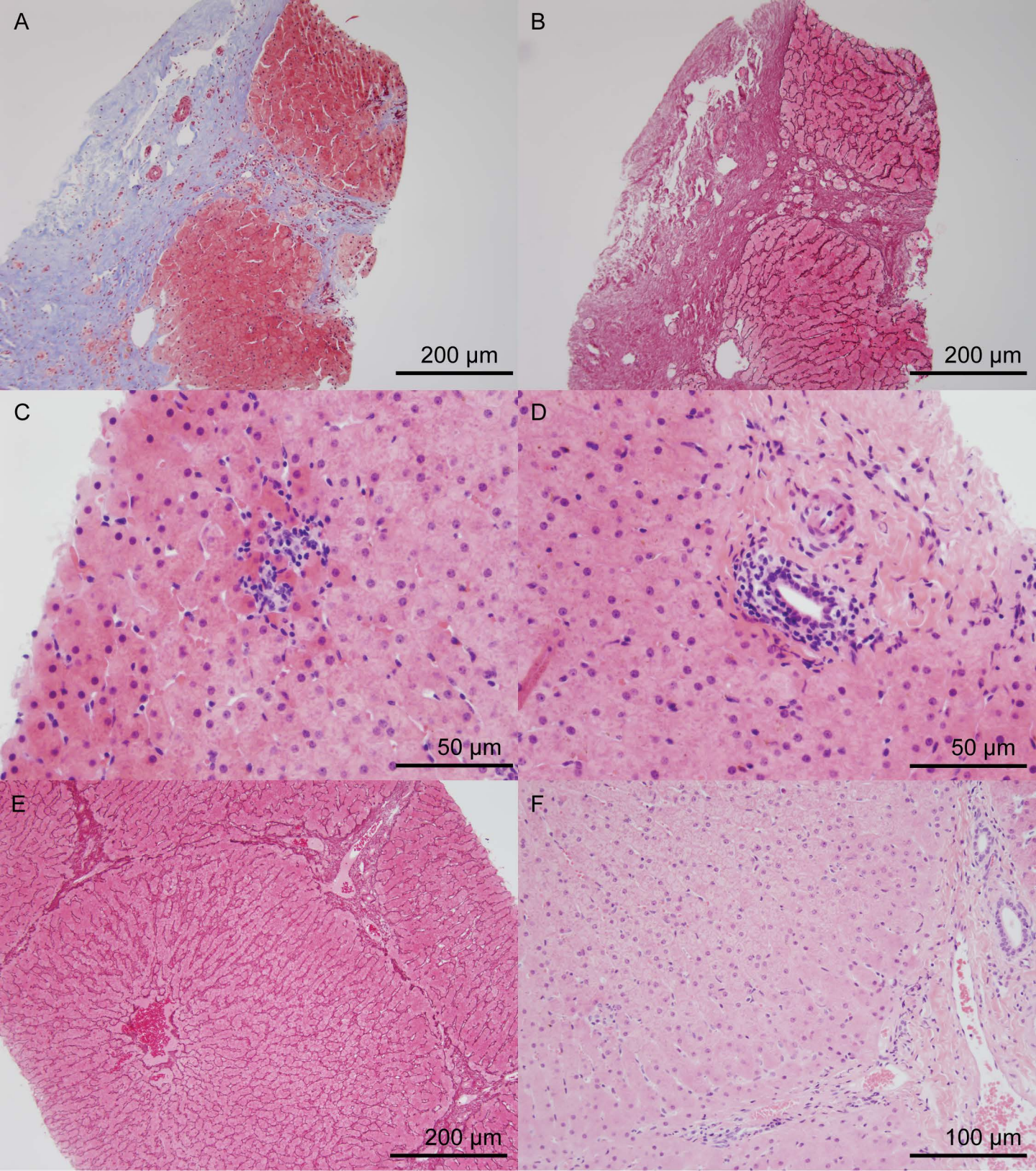

Horse R
RLU
10^7
10^6
10000
1000
virus cleared
challenge 1
challenge 2
productive
0
10
20
30
40
Week

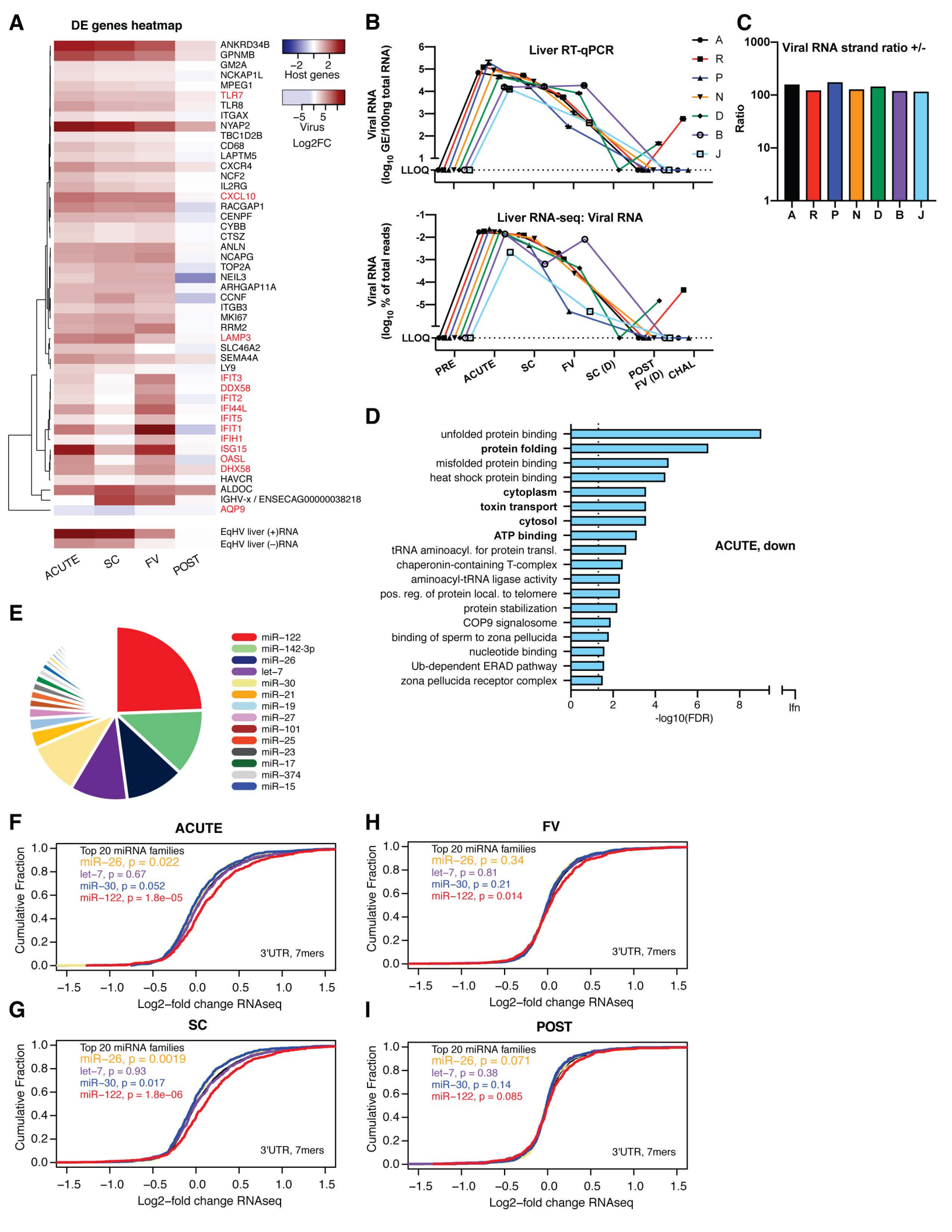

Horse B
RLU
10^7
10^6
10000
1000
virus cleared
challenge 1
productive
challenge 2
0
10
20
30
40
Week


Horse J
RLU
10^7
10^6
10^5
10000
1000
virus cleared
challenge 1
challenge 2
0
10
20
30
40
Week


A
NZP1 Inoculum
Horse B, ACUTE (day 7)
Horse N, ACUTE (day 9)
Horse A, ACUTE (day 11)
Horse B, FV (day 77)
Horse N, FV (day 84)
Horse A, FV (day 91)
Horse D, ACUTE (day 9)
Horse R, ACUTE (day 11)
Horse A, Challenge (hom.)
Horse D, FV (day 28)
Horse R, FV (day 84)
Horse D, FV (day 183)
100%
10%
1%
0.1%
0
2000
4000
6000
8000
B
CU heterologous Inoculum
Horse O, ACUTE (Naïve)
Horse P, 1y Re-challenge (het.)
Horse R, Challenge (het.)
Horse N, 1y Re-challenge (het.)
C
E1
E2
NS2
NS3
NS4B
NS5A
NS5B
p7
NS4A
C
Sequence Logo
Bits
NZP1
380
390
394
403
413
423
TTNNYFKLCLLLLAYSGLVSC---D----NYLNVSLSCNFTVKEMWGWTFFPKWALLNGQRLNC
E1
E2
CU sub-I
CU sub-II
CU sub-III
AFJ20703
AJF44502
AFJ20705
AFJ20704
AFJ20709
AFY03631
AGT17571
AFJ20707
AFJ20706
AFJ20708


A
B
C
D
Horse A
Horse B
Horse D
Horse J
Horse N
Horse P
Horse R
Cell concentration (x10³/µl)
Week post inoculation
% of PBMC
% of CD3+ cells
% Ki67+
EqHV (Log10 GE/mL)
NPHV (Log10 GE/ml)
<LLOQ
Neutrophils
Lymphocytes
Monocytes
T
B
M
CD16+ M % of M
NK
EqHV
CD4
CD4-CD8-
CD4+CD8+
CD8+
EqHV
CD4
CD4-CD8-
CD4+CD8+
CD8+
EqHV


A
200 µm
B
200 µm
C
50 µm
D
50 µm
E
200 µm
F
100 µm


A
DE genes heatmap
ANKRD34B
GPNMB
GM2A
NCKAP1L
MPEG1
TLR7
TLR8
ITGAX
NYAP2
TBC1D2B
CD68
LAPTM5
CXCR4
NCF2
IL2RG
CXCL10
RACGAP1
CENPF
CYBB
CTSZ
ANLN
NCAPG
TOP2A
NEIL3
ARHGAP11A
CCNF
ITGB3
MKI67
RRM2
LAMP3
SLC46A2
SEMA4A
LY9
IFIT3
DDX58
IFIT2
IFI44L
IFIT5
IFIT1
IFIH1
ISG15
OASL
DHX58
HAVCR
ALDOC
IGHV-x / ENSECAG00000038218
AQP9
−2
2
Host genes
−5
5
Virus
Log2FC
EqHV liver (+)RNA
EqHV liver (−)RNA
ACUTE
SC
FV
POST
B
Liver RT-qPCR
Viral RNA (log10 GE/100ng total RNA)
LLOQ
A
R
P
N
D
B
J
Liver RNA-seq: Viral RNA
Viral RNA (log10 % of total reads)
LLOQ
PRE
ACUTE
SC
FV
SC (D)
POST FV (D)
CHAL
C
Viral RNA strand ratio +/-
Ratio
A R P N D B J
D
unfolded protein binding
protein folding
misfolded protein binding
heat shock protein binding
cytoplasm
toxin transport
cytosol
ATP binding
tRNA aminoacyl. for protein transl.
chaperonin-containing T-complex
aminoacyl-tRNA ligase activity
pos. reg. of protein local. to telomere
protein stabilization
COP9 signalosome
binding of sperm to zona pellucida
nucleotide binding
Ub-dependent ERAD pathway
zona pellucida receptor complex
ACUTE, down
-log10(FDR)
Ifn
E
miR−122
miR−142-3p
miR−26
let−7
miR−30
miR−21
miR−19
miR−27
miR−101
miR−25
miR−23
miR−17
miR−374
miR−15
F
ACUTE
Top 20 miRNA families
miR−26, p = 0.022
let−7, p = 0.67
miR−30, p = 0.052
miR−122, p = 1.8e−05
3'UTR, 7mers
Cumulative Fraction
Log2−fold change RNAseq
G
SC
Top 20 miRNA families
miR−26, p = 0.0019
let−7, p = 0.93
miR−30, p = 0.017
miR−122, p = 1.8e−06
3'UTR, 7mers
Cumulative Fraction
Log2−fold change RNAseq
H
FV
Top 20 miRNA families
miR−26, p = 0.34
let−7, p = 0.81
miR−30, p = 0.21
miR−122, p = 0.014
3'UTR, 7mers
Cumulative Fraction
Log2−fold change RNAseq
I
POST
Top 20 miRNA families
miR−26, p = 0.071
let−7, p = 0.38
miR−30, p = 0.14
miR−122, p = 0.085
3'UTR, 7mers
Cumulative Fraction
Log2−fold change RNAseq