CellPress
OPEN ACCESS

Supplemental figures

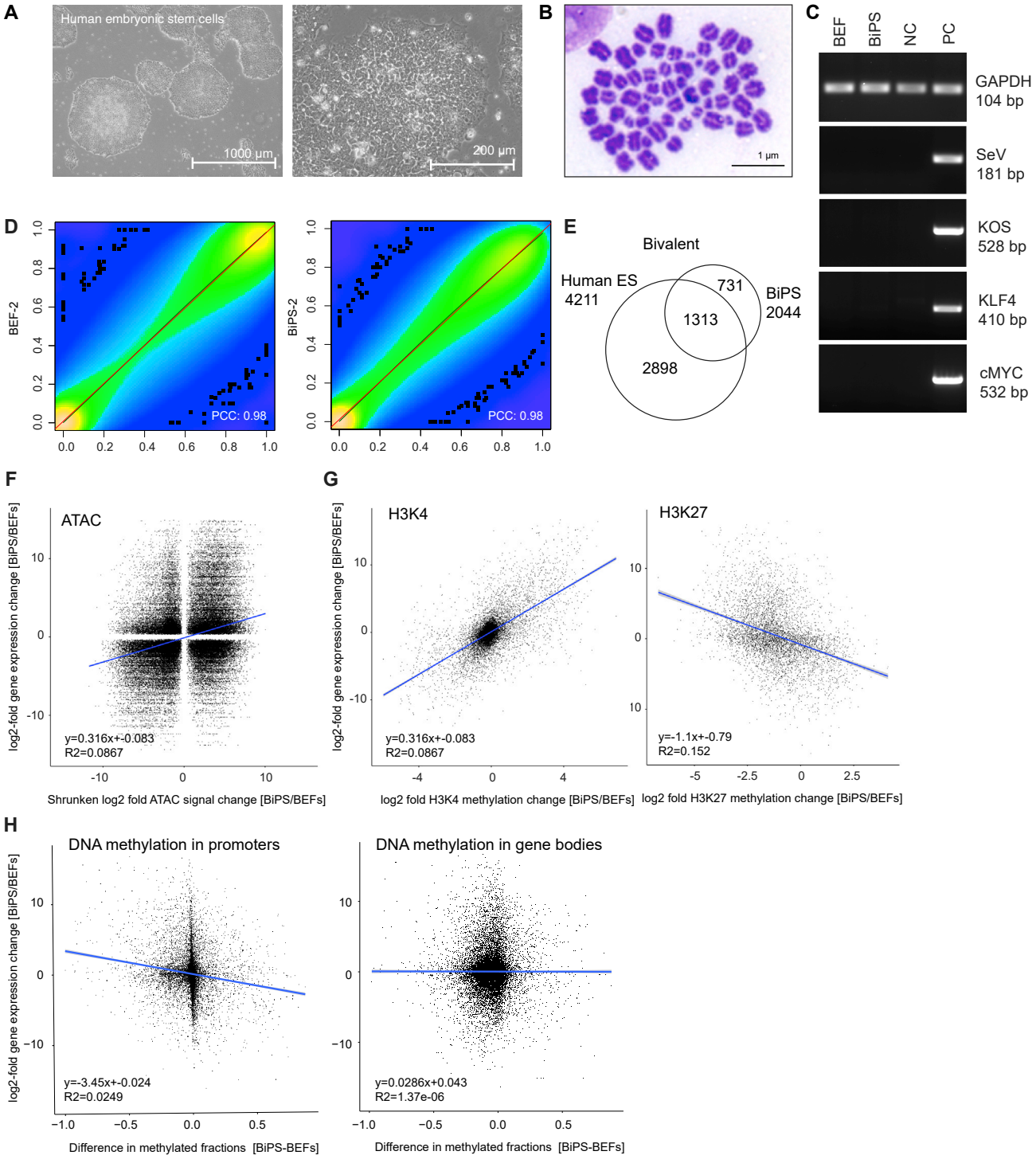


Figure S1. Characteristics of pluripotent stem cells generated from *Rhinolophus ferrumequinum* fibroblasts, related to Figure 1 and Table S1

(A) Microscopic images of human embryonic stem cells (H9) and bat pluripotent stem cells at different magnifications.

(B) Karyotype analysis of BiPS cells at passage 17. Shown is a representative image after Giemsa staining of metaphase spreads with 58 chromosomes.

(legend continued on next page)

(C) PCR verification of reprogramming-associated virus clearing. Bat iPS cells (BiPS) at passage 92 were tested for Sendai virus clearance in comparison to the embryonic fibroblasts used as starting material (BEF), adult fibroblasts as negative control (NC), and freshly transduced cells at passage 3 as a positive control (PC). bp, base pairs; SeV, Sendai virus; KOS, KLF4-OCT4-SOX2.

(D) Correlation scatter plot of methylation level at common CpG sites in BEF or BiPS cells. Shown is the representative result of one of two replicates per cell type. BEF, bat embryonic fibroblast cells; BiPS, bat pluripotent stem cells; PCC, Pearson correlation coefficient.

(E) Venn diagram illustrating the overlap of bivalent genes in bat iPS cells and human ES cells.

(F) Correlation plot of shrunken log_2-fold changes in ATAC-seq signal with log_2-fold expression changes. Shown are the average values with $p < 0.05$ of two replicates per cell type.

(G) Correlation of log_2-fold changes in H3K4 trimethylation (H3K4me3, left) or H3K27 trimethylation (H3K27me3, right) with log_2-fold changes in gene expression. Shown are the results of one sample for each chromatin mark.

(H) Correlation of log_2-fold gene expression changes with the difference in the methylated fraction of promoters (left) or gene bodies (right) fractions. Shown are the average expression changes from three replicates per condition and methylated fractions from two replicates.

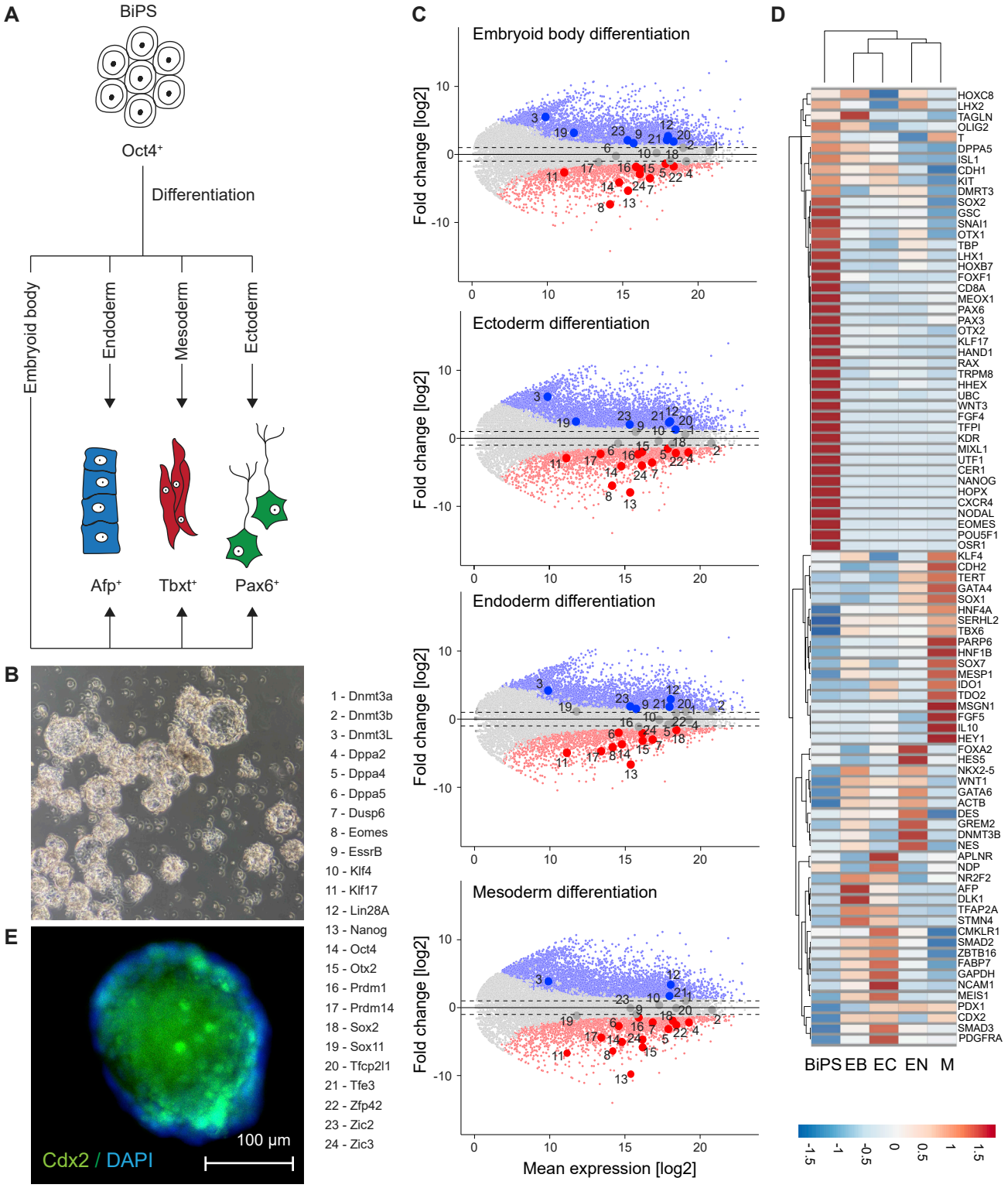


Figure S2. Differentiation potential of bat pluripotent stem cells, related to Figure 2 and Table S2

(A) Schematic of differentiation strategies.

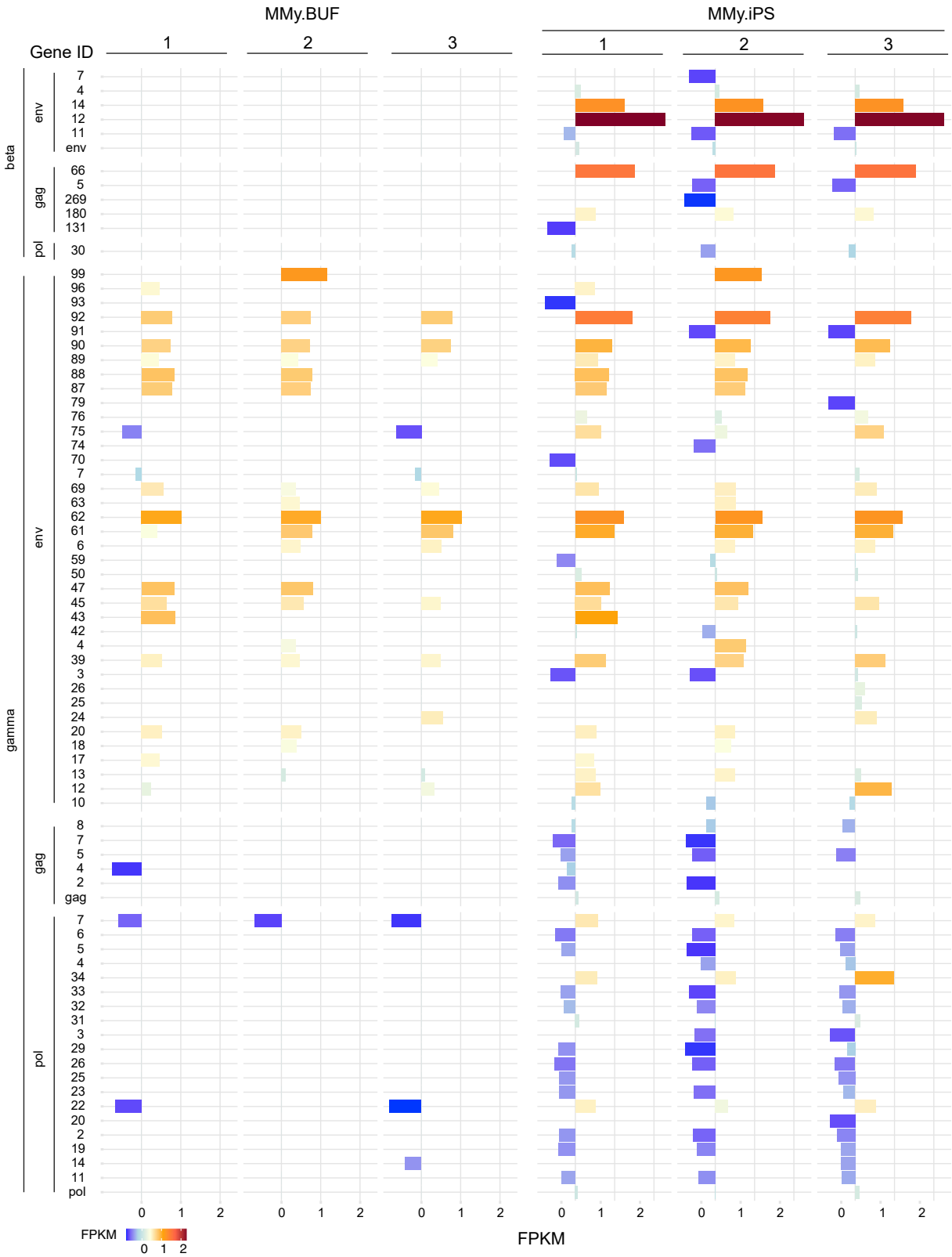
(B) Phase contrast image of BiPS2-derived embryoid bodies after 7 days of differentiation.

(legend continued on next page)

(C) MA plots depicting the log_2 mean expression and log_2-fold expression changes of all genes in bat induced pluripotent stem cells (BiPS) after exposure to the noted differentiation conditions illustrated in (A). Shown is the mean average of each gene from three replicates per differentiation condition (n = 3). EB, Embryoid body differentiation; EC, human ectoderm differentiation conditions; EN, human endoderm differentiation conditions; M, human mesoderm differentiation conditions.

(D) Heatmap depicting expression changes of genes that are known as markers for human ectoderm, mesoderm, or endoderm during the differentiation of BiPS under the conditions described in (A). Shown is the average expression of 3 replicates per condition.

(E) Representative microscopic image of a bat blastoid after immunostaining using a Cdx2-specific antibody.



(legend on next page)

Figure S3. Reactivation of endogenized viral elements in *Myotis myotis* induced pluripotent stem cells, related to Figure 6 and Table S4

Expression of indicated ERV elements in *Myotis myotis* bat uropatagium fibroblasts (MMy.BUF) and induced pluripotent stem cells (MMy.iPS) as determined by extracting the overlap between RNA-seq reads mapped to the *Myotis myotis* genome and known mapped ERV elements. Shown are the elements with the most evident differences from three replicates.

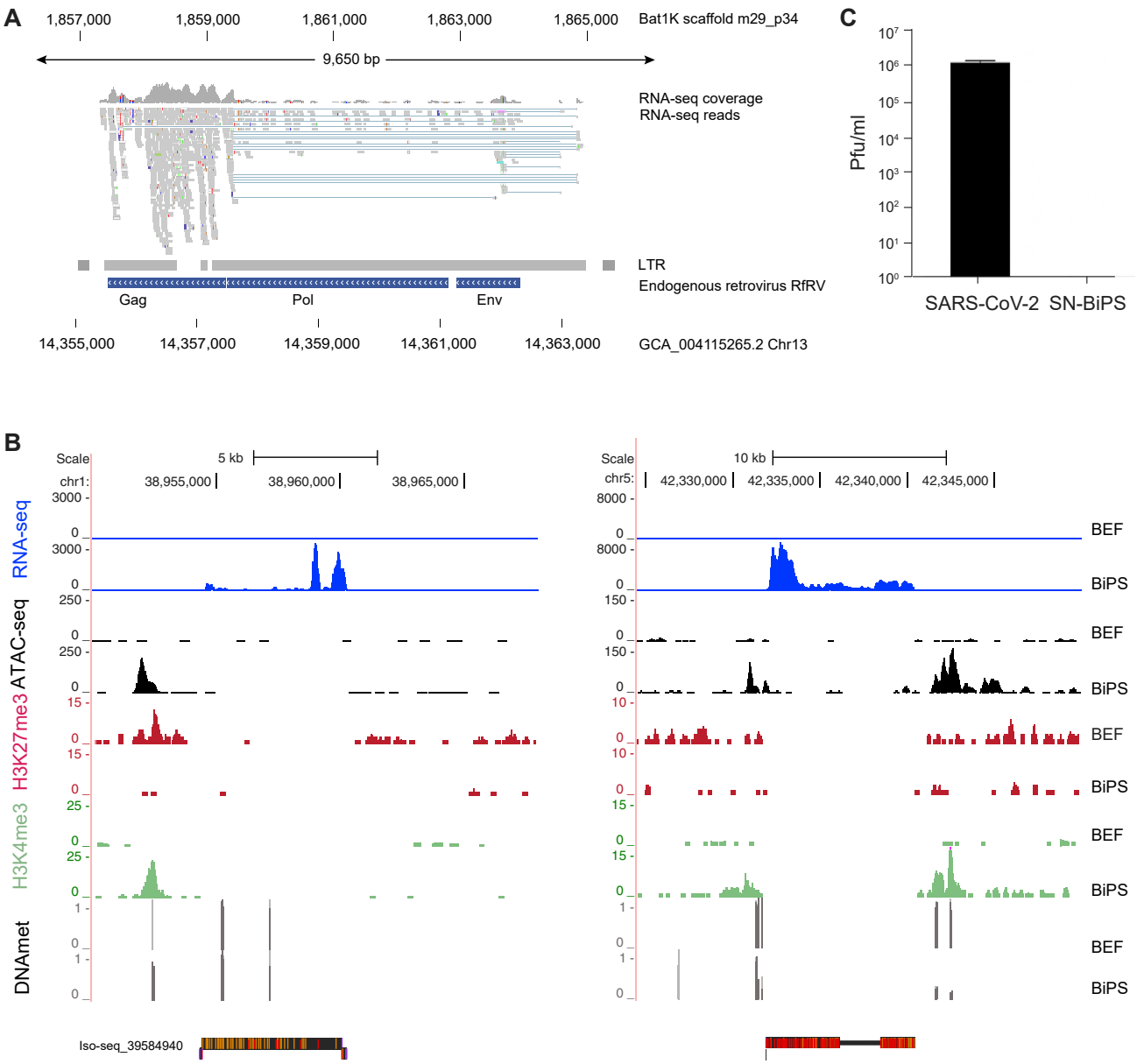


Figure S4. Viral activity in BiPS cells, related to Figure 6 and Table S4

(A) RNA-seq tracks aligning with the genomic *R. ferrumequinum* region (Bat1K annotated scaffold, *top*, and GCA assembly, *bottom*) containing the previously identified endogenous retrovirus RfRV. LTR, long terminal repeat.

(B) Sequencing tracks showing expression, ATAC-seq signal, histone H3K27 trimethylation (H3K27me3), histone H3K4 trimethylation (H3K4me3) as well as DNA methylation (DNAmet) status in the genomic region on chromosome 1 surrounding RFe-V-MD1 shown in Figure 6 (left) and a genomic region on chromosome 5 that contains a sequence highly similar to RFe-V-MD1 (right). Note that the 6088 bp-long Iso-seq fragment aligns with the (-) strand on chromosome 1 and the (+) strand on chromosome 5.

(C) Plaque assay with supernatant of *R. ferrumequinum* induced pluripotent stem cells compared to active SARS-CoV-2 viral particles in Vero cells. BEF, bat embryonic fibroblasts; BiPS, bat induced pluripotent stem cells; SN, supernatant.

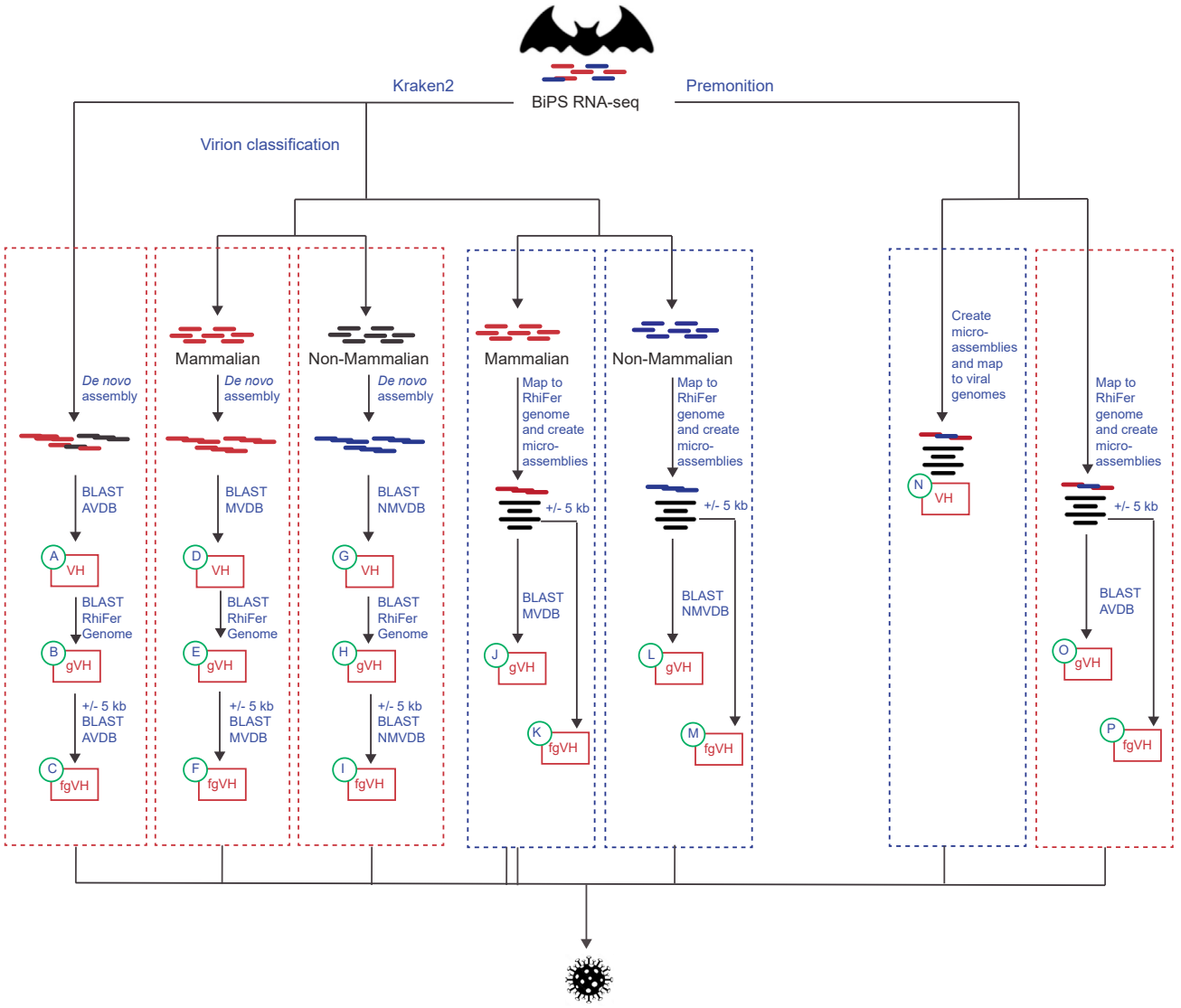
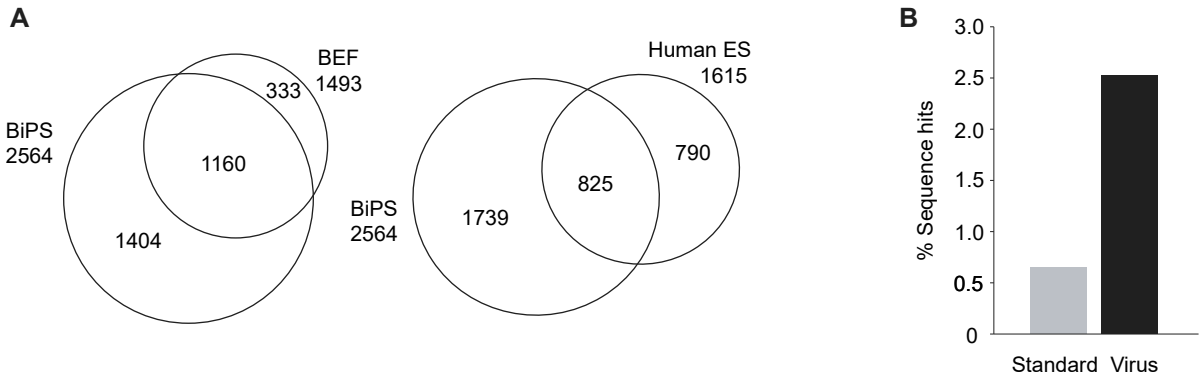
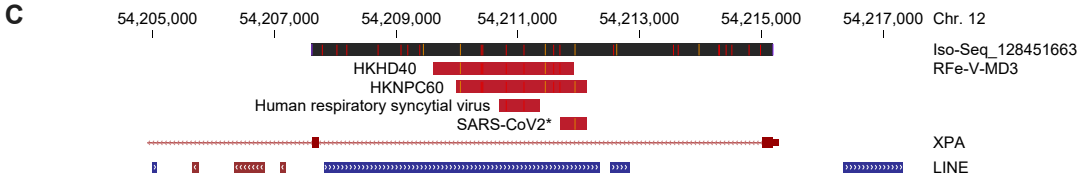
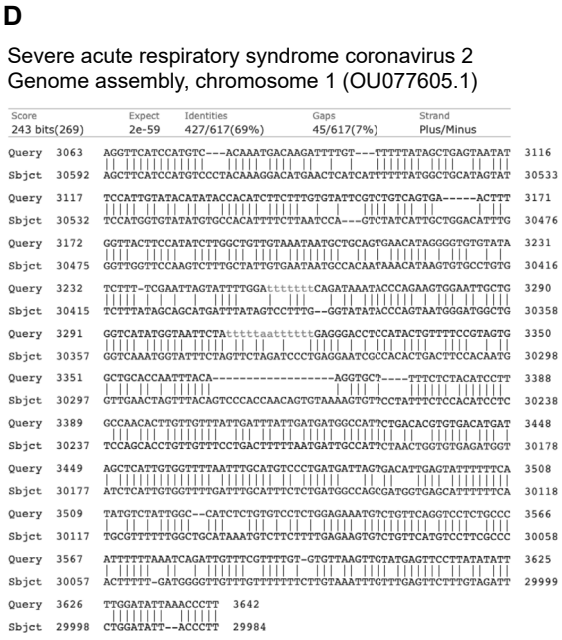


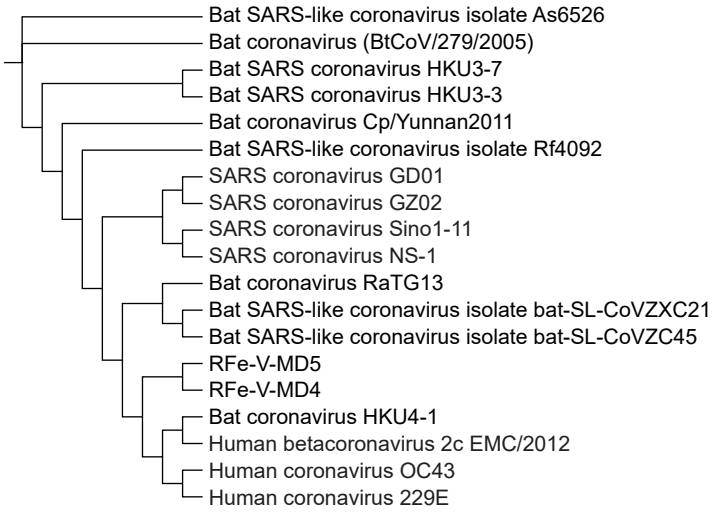
Figure S5. Virome mining approaches, related to Figure 7 and Tables S5 and S6

RNA-seq reads were first classified using Kraken2 (left). The classified reads were then either analyzed directly or separated by their homology to a mammalian or a non-mammalian virus using virion. In a 'bottom-up' approach (*red boxes*), the reads that were classified to belong to the same virus were assembled *de novo* and blasted against databases containing all viruses (AVDB), mammalian viruses (MVDB) or non-mammalian viruses (NMVDB). The identified sequences with viral hits were then further mapped to the *R. ferrumequinum* (RhiFer) genome (gVH) and in a final step, reads/assemblies that did map to the genome were extended by 5 kb on both sides and the sequences again blasted against the databases as before to extend the search for viral hits in flanking genomic regions (fgVH). In a parallel 'top-down' approach (blue boxes), the Kraken2 reads were mapped directly to the *R. ferrumequinum* genome, micro-assemblies were generated based on the mapped reads, and blasted directly against the different databases as before or first extended by 5 kb on both sides. In an orthogonal approach, the reads were similarly assigned to viruses using the Microsoft Premonition metagenomics pipeline (right). First all reads were classified as aligning to a virus or as being unaligned, micro-assemblies were generated and mapped to viral genomes (*red box*). Consensus sequences were extracted and blasted against the *R. ferrumequinum* genome, extended by 5 kb and blasted against a database containing all viruses as before. In another 'top-down' approach (*blue box*), the Premonition classified reads were mapped to the *R. ferrumequinum* genome first, and micro-assemblies generated that were blasted against the 'All virus' database either directly or after being extended by 5 kb on both sides. AVDB, all viruses database; MVDB, mammalian virus database; NMVDB, non-mammalian virus database; VH, viral hits; gVH, viral hits with homology to the bat genome; fgVH, viral hits using flanked genomic regions; RhiFer, *Rhinolophus ferrumequinum*; kb, kilobase pairs.









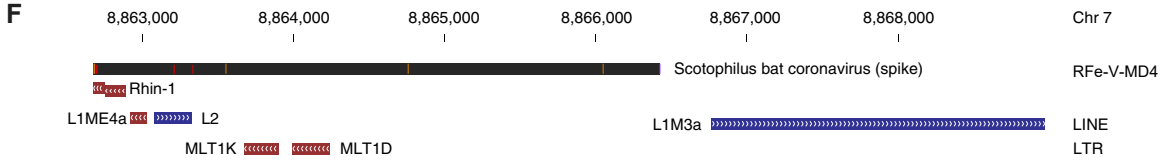


Figure S6. Virome mining in *Rhinolophus ferrumequinum* BiPS cells, related to Figure 7 and Tables S6 and S7

(A) Venn diagram showing the overlap of viral species as classified by Kraken2 in RNA-seq reads of BiPS and BEF cells (this study; *left*) and human ES cells (GSM4192140 and GSM4192141)[98] and *R. ferrumequinum* BiPS cells (right).

(B) Comparison of sequence hits using the Standard or a virus-specific database as reference in Kraken2 analysis with BiPS RNA-seq reads.

(legend continued on next page)

(C) Illustration of a short viral integration with homology to two *human herpesvirus 4* isolates (*HKD40* and *HKNPC60*), *human respiratory syncytial virus* (Kilifi isolate), and a ~500 bp fragment that was identified at the end of a *SARS-CoV2* isolate from an infected patient (NCBI ref. sequence OU077605.1). Shown is the alignment of the 7955 bp-long fragment to the *R. ferrumequinum* genomic sequence.
(D) Sequence alignment of the Iso-seq read from BiPS (RFe-V-MD3) and the ~500 bp SARS-CoV2 fragment described in (C).
(E) Phylogenetic relationship of the identified sequences RFe-V-MD4 and RFe-V-MD5 to a selection of corona viruses.
(F) Genome track of the 6404 bp-long sequence (RFe-V-MD4) with homology to the *Scotophilus bat coronavirus 512* containing the spike coding region.

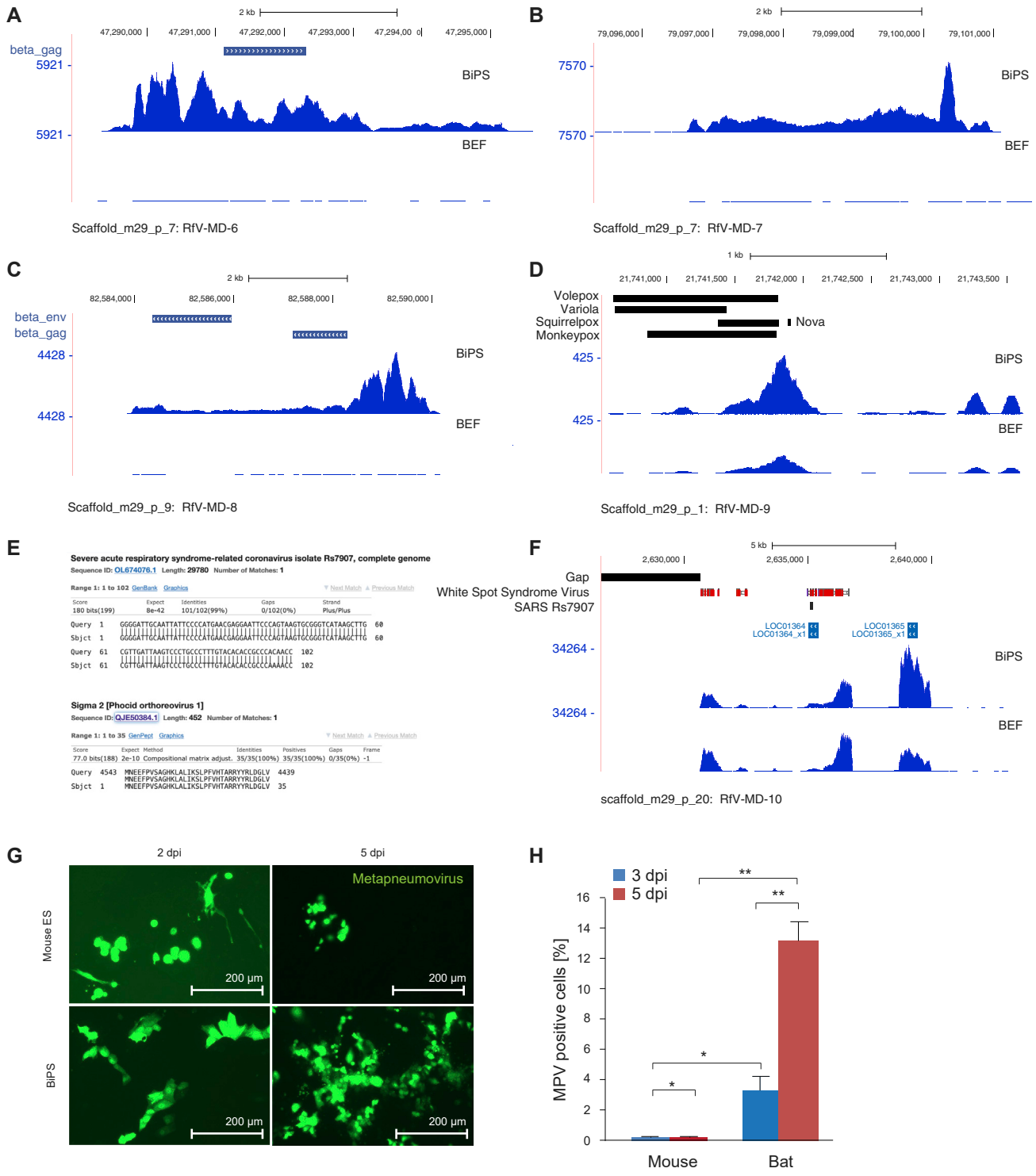


Figure S7. Virome mining in *Rhinolophus ferrumequinum* BiPS cells, related to Figure 7 and Tables S6 and S7

Chromosomal regions within indicated *R. ferrumequinum* scaffolds detected to contain homologies to the (A) *Mason-Pfizer monkey virus*, the (B) *Jaagsiekte sheep retrovirus*, or the (C) Simian endogenous retrovirus. Shown is the surrounding region covered by mapped RNA-seq reads that show further homologies to known viruses on DNA and protein levels as outlined in Table S7.

(D) Chromosomal region in the vicinity of DNA sequences with homology to several DNA viruses covered by RNA-seq reads that show homologies to *cowpox* and *monkeypox virus* proteins after translation using BLASTx.

(legend continued on next page)

(E) Short homology of expressed sequence reads with the *severe acute respiratory syndrome-related coronavirus isolate Rs7907* on DNA level (top) which translates into an N-terminal fragment of a dsRNA-binding protein (bottom).
(F) The chromosomal region surrounding the fragment shown in (E) that is covered by mapped sequence reads with further homologies to the White spot syndrome virus.
(G) Microscopic images of R1 mouse embryonic stem (ES) cells and *R. ferrumequinum* induced pluripotent stem cells (BiPS) at 2- and 5-days post-infection (dpi) with GFP-labeled Metapneumovirus (MPV) particles.
(H) Quantification of experiment shown in (A) at 3- and 5-days post-infection (dpi) by FACS. Data are shown as mean +/- SD; *$p < 0.05$,**$p < 0.001$.