YMTHE, Volume *32*

Supplemental Information

Safety concern of recombination between self-amplifying mRNA vaccines and viruses is mitigated *in vivo*

Tessy A.H. Hick, Corinne Geertsema, Wilson Nguyen, Cameron R. Bishop, Linda van Oosten, Sandra R. Abbo, Troy Dumenil, Frank J.M. van Kuppeveld, Martijn A. Langereis, Daniel J. Rawle, Bing Tang, Kexin Yan, Monique M. van Oers, Andreas Suhrbier, and Gorben P. Pijlman

Supplemental Tables and Figures

Table S1. RNA-Seq read pairs derived from feet. Polyadenylated positive strand RNA sequences of GETV, SAM vaccine and mouse present in RNA isolated from feet identified by RNA-seq. Chimeric read-pairs identified from all mouse foot RNA-Seq data. Chimeric cDNA species likely arise as artefacts due to random template switching during reverse-transcription (RT)[1] (see Supplemental Fig. 6 and 7).

Strain	Time (dpi)	Sample	Proper read pairs				Chimeric read pairs	
			Total	GETV	SAM	Mouse	GETV/ SAM	Mouse
C57BL/6J	2	CCg24	16,496,555	357	550	16,094,607	0	661
		Cg25	17,444,881	274	516	17,010,685	0	542
		Cg26	15,482,350	272	674	14,938,719	0	535
	3	Cg21	16,509,615	438	78	16,195,705	0	498
		Cg22	15,675,118	2,993	4	14,895,533	0	303
		Cg23	16,223,836	155	25	15,884,778	0	605
Ifnar$^{-/-}$	2	Cg51	16,047,840	421,495	5,116	15,071,584	1	438
		Cg52	17,425,852	462,229	10,783	16,204,692	2	240
		Cg61	15,998,134	347,707	5,836	15,130,673	0	435
	3	Cg21	18,994,287	1,655,929	5,293	16,886,302	0	516
		Cg41	17,238,761	2,142,702	5,556	14,694,258	1	467
		Cg42	15,706,105	882,896	2,818	14,422,143	0	498
	5	Cg11	21,969,798	3,583,158	3,925	17,884,208	0	1110
		Cg31	16,841,127	2,722,173	3,959	13,673,336	0	944
Rag1$^{-/-}$	2	Cg83	18,491,093	94	45	18,183,205	0	391
		Cg84	18,007,132	141	133	16,885,390	0	48
		Cg85	19,604,811	431	133	18,696,171	0	114
	3	Cg81	18,108,538	340	21	17,150,222	0	244
		Cg82	14,059,628	708	5	13,581,597	0	469
		Cg86	20,155,651	210	73	19,724,814	0	698
	14	Cg74	16,923,637	2,710	7	16,520,045	0	556
		Cg75	16,244,992	2,907,751	3,866	12,818,883	1	844
		Cg76	17,387,562	421	6	16,930,047	0	483
Total			**397,037,303**	**15,135,584**	**49,422**	**369,477,597**	**5**	**11,639**

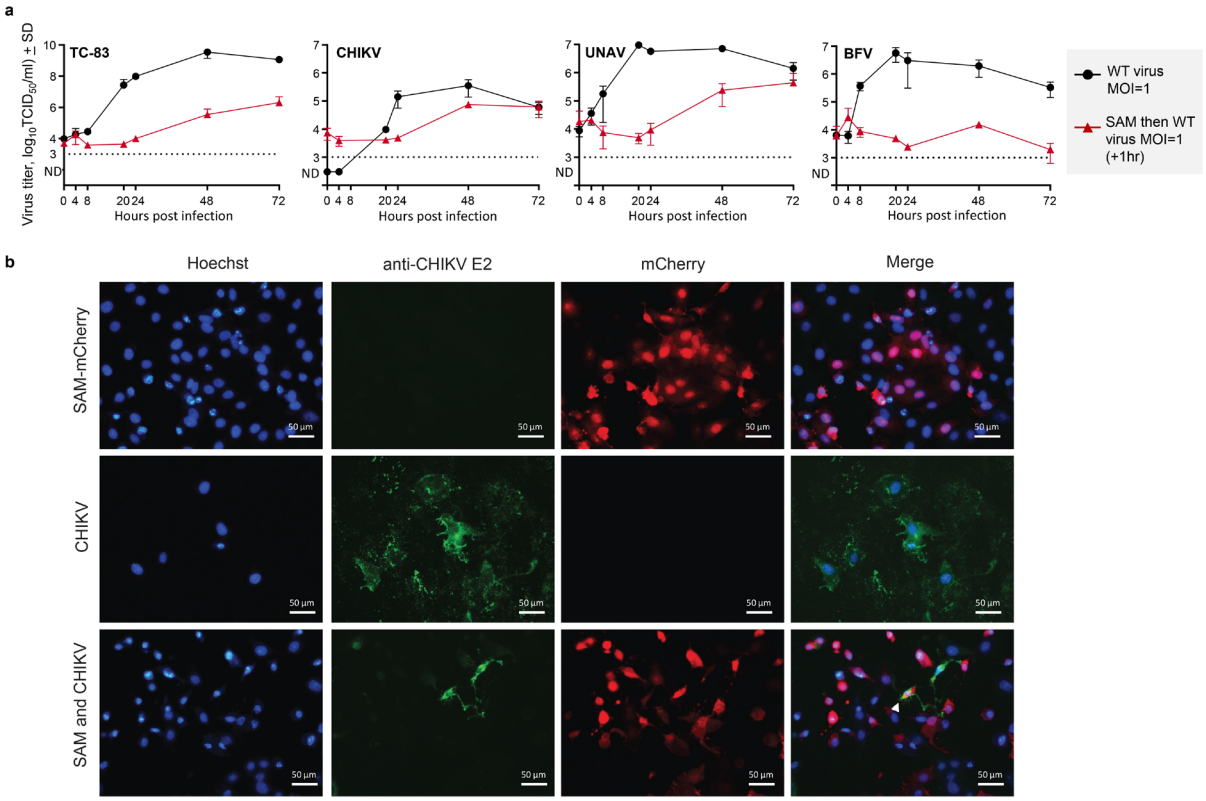

Figure S1. Superinfection exclusion and co-expression. Vero cells were transduced with SAM (10 VRPs/cell) and infected with WT alphaviruses (MOI=1 $TCID_{50}$/cell) with a 1 hour delay (n=2). Supernatant fractions were collected and titrated to determine WT alphavirus growth kinetics by serial dilutions on Vero cells (detection limit 3 $log_{10}TCID_{50}$/ml). **b**, Fluorescence microscopy of co-transduced/infected Vero cell; cell nuclei stained with Hoechst, WT alphavirus (CHIKV) stained with rabbit anti-CHIKV E2 (1:5000), and SAM vaccine mCherry expression.

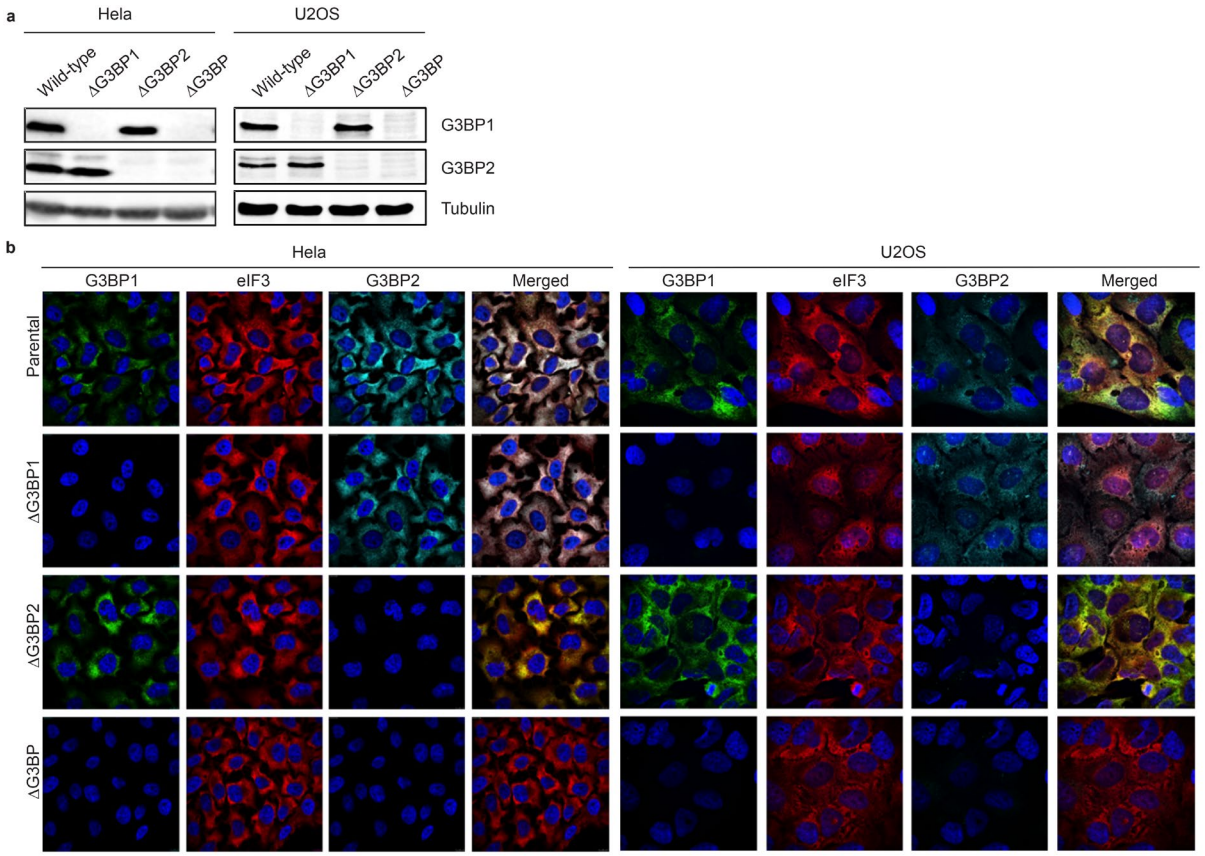

Figure S2. Hela and U2OS ΔG3BP cell lines. Hela and U2OS G3BP deficient cells (ΔG3BP) were generated using CRISPR-Cas9 technology to delete expression of G3BP1 (ΔG3BP1), G3BP2 (ΔG3BP2), or both (ΔG3BP) as described[2]. **a** Western blot showing loss of G3BP in Hela and U2OS cells. **b** Immunofluorescence assay using mouse anti-G3BP (BD biosciences, 1:2000), rabbit anti-G3BP2 (1:2000, Bethyl), anti-tubulin (1:5000, Sigma-Aldrich), and anti-eIF2 (stress granular identification factor, 1:200, SantaCruz) antibodies.

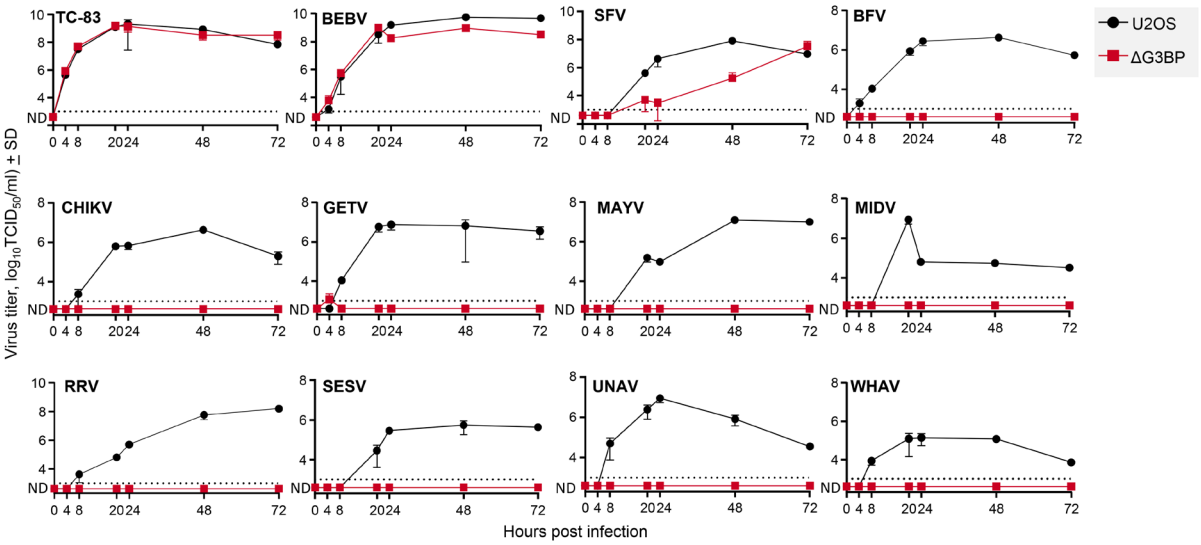

Figure S3. Alphavirus replication kinetics on U2OS ΔG3BP cells. U2OS and U2OS ΔG3BP cells were infected with various alphaviruses (0.1 $TCID_{50}$/cell, n=2). Supernatant fractions were collected and titrated to determine alphavirus growth kinetics (detection limit 3 $log_{10}TCID_{50}$/ml).Venezuelan equine encephalitis virus TC-83 (TC-83), Bebaru virus (BEBV), and Semliki Forest virus (SFV) demonstrated to replicate independently of G3BP, whereas Barmah Forest virus (BFV), chikungunya virus (CHIKV), getah virus (GETV), Mayaro virus (MAYV), Middelburg virus (MIDV), Ross River virus (RRV), southern elephant seal virus (SESV), Una virus (UNAV), and Whataroa virus (WHAV) required the presence of G3BP for efficient virus replication.

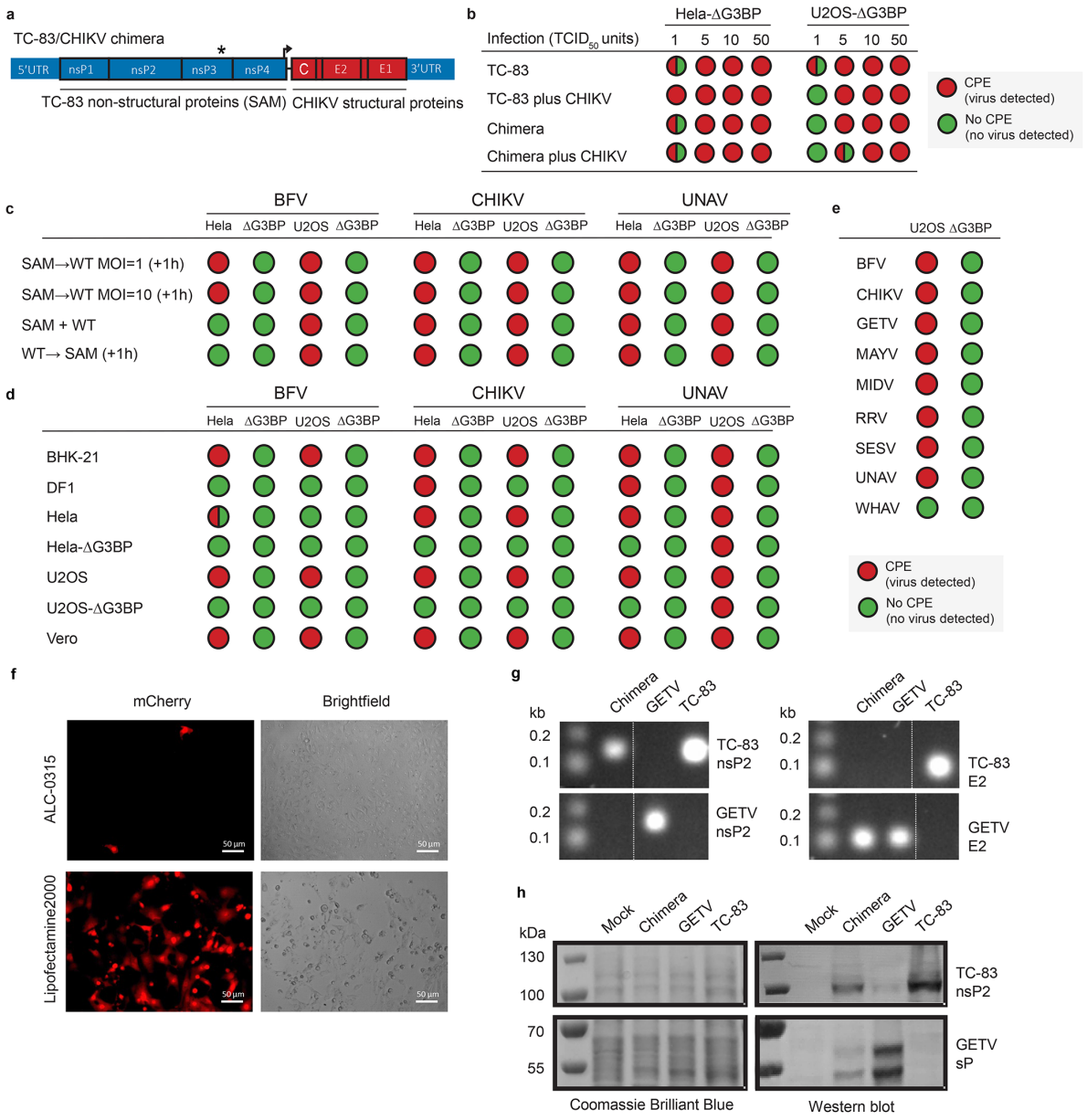

Figure S4. Alphavirus recombination detection assay using ΔG3BP cells. a, An alphavirus chimera was constructed, that encoded the nsPs of TC-83 (SAM) and the structural proteins of CHIKV S27, by cloning the latter into the multiple cloning site of the SAM plasmid. The hypervariable domain of nsP3 (indicated by asterisk) confers G3BP independent replication capacity. This infectious clone was recovered after lipofectamine-mediated transfection of Vero cells. **b**, The TC-83/CHIKV chimera and TC-83 were able to replicate and induce CPE in 10^5 ΔG3BP cells within four serial passages even at low MOI and in the presence of an excess of G3BP-dependent CHIKV ($5x10^5$ $TCID_{50}$), illustrating the assay's sensitivity and suitability. **c**, Supernatants from Vero cells transduced with a SAM vaccine encoding mCherry (10 VRPs/cell) and infected with the indicated WT alphaviruses (as in Fig. 2a) were evaluated for

the presence of chimeric alphaviruses using both Hela ΔG3BP and U2OS ΔG3BP cells. Vero cells were transduced with SAM vaccine and 1 hour later infected with the WT alphavirus at the indicated MOIs (SAM→WT MOI=1/10 (+1h)). SAM vaccine and WT virus (10 $TCID_{50}$/ml) were added to the Vero cells at the same time (SAM + WT), or Vero cells were infected with the WT alphavirus (10 $TCID_{50}$/ml) and 1 hour later transduced with SAM vaccine (WT→ SAM (+1h)). After 72 hours, supernatants were four times serial passaged on Hela, Hela ΔG3BP, U2OS, and U2OS ΔG3BP cells. No CPE (indicating absence of a chimeric alphavirus) was observed in ΔG3BP cells. **d**, Indicated cell lines were infected with the WT alphaviruses (10 $TCID_{50}$/ml) followed by transduction with SAM vaccine (10 VRPs/cell, 1 hour later). After 72 hours, supernatants were four times passaged on Hela, Hela ΔG3BP, U2OS, and U2OS ΔG3BP cells. No CPE (indicating absence of a chimeric alphavirus) was seen in ΔG3BP cells. **e**, Supernatants from Vero cells infected with the indicated WT alphaviruses followed by transduction with SAM vaccine (10 VRPs/cell, 1 hour later) were four times passaged on U2OS, and U2OS ΔG3BP cells. No CPE (indicating absence of a chimeric alphavirus) was seen in ΔG3BP cells. **f**, Delivery to Vero cells of SAM vaccine (SAM-mCherry) formulated as LNPs (ALC-0315 lipid mix, Echelon Biosciences) or using Lipofectamine-2000 (Thermo). **g,** One step RT-PCR analysis of the SAM-mCherry-GETV chimera (chimera), GETV, and TC-83 using RNA extracted from supernatants. Primers amplify fragments of the TC-83 non-structural protein 2 sequence (nsP2, 150 bp), GETV nsP2 sequence (150 bp), TC-83 structural envelope glycoprotein E2 sequence (100 bp), and GETV E2 sequence (GETV sP, 100 bp). **h**, SDS-PAGE and western blot analysis of SAM-mCherry-GETV chimera (chimera), GETV, and TC-83 infections. Infected cell lysate was heated in SDS sample buffer and loaded onto SDS-PAGE gel. The SDS-PAGE gel was stained by using Coomassie Brilliant Blue (protein stain) and transferred to PVDF membranes followed by western blot detection using goat anti-TC-83 nsP2 (1:1000, AlphaVax), rabbit anti-goat alkaline phosphatase (1:2500, Sigma), alkaline phosphatase, and NBT/BCIP (1:50, Roche).

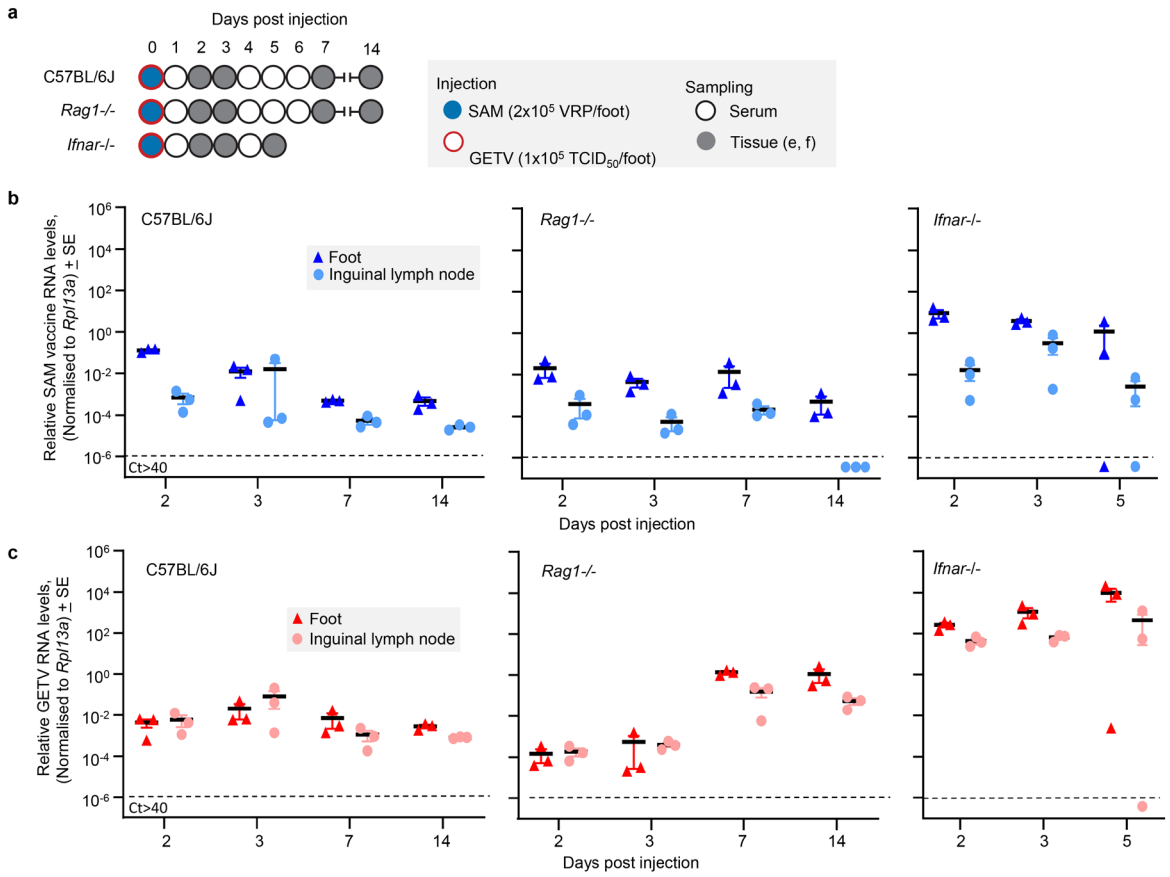

Figure S5. SAM vaccine and GETV co-injection qRT-PCR analyses by tissue and time. **a,** C57BL/6J mice were co-injected subcutaneously into the hind feet with SAM vaccine ($2x10^5$ VRP/foot) and with GETV (10^5 $TCID_{50}$/foot) into hind feet of indicated mouse strains. **b,** SAM vaccine RNA in feet and inguinal lymph nodes detected by qRT-PCR. **c,** GETV RNA in feet and inguinal lymph nodes detected by nsP2 sequence qRT-PCR. C57BL/6J feet on days 2 and 3, *Rag1*$^{-/-}$ feet on days 2, 3 and 14, and *Ifnar*$^{-/-}$ feet on days 2, 3 and 5 were chosen for RNA-Seq as they had the highest consistent levels of RNA for both GETV and SAM vaccine. The persistence of alphavirus RNA in feet of C57BL/6J mice has been reported previously[3,4].

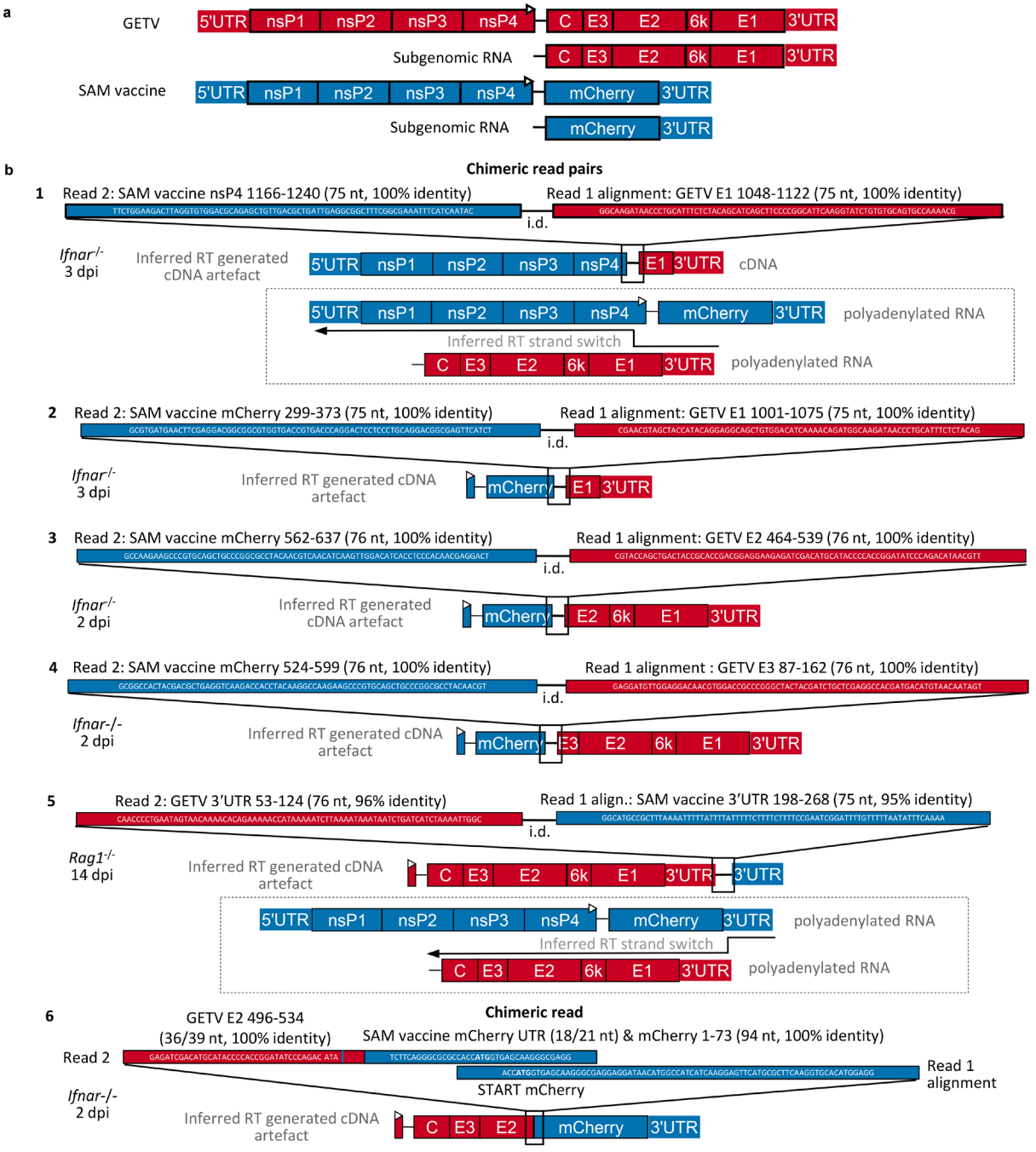

Figure S6. Characterization of alphavirus chimeric read pairs and read. a, Polyadenylated positive strand RNA sequences of GETV and SAM vaccine present in RNA isolated from feet identified by RNA-seq. **b,** Six chimeric read-pairs/read identified from all mouse foot RNA-Seq data (3 mice per group, 3 mouse strains, all times post injection, see Fig. 3a, c). i.d. – inter-read distance. Chimeric cDNA species likely arise as artefacts due to random template switching during reverse-transcription[1] (see Fig. S8). The artefactual chimeric cDNA species and the strand switching events (dashed boxes) can be inferred. To detect chimeric read pairs, alignments (BAMs) were viewed using Samtools to identify paired alignments in which each read in the pair aligned to a different sequence. To detect chimeric reads, read files (Fastqs)

were searched to identify reads containing 22-mers relating to both the GETV and SAM vaccine reference sequences, using an original Python script in Python version 3.7 (original script https://github.com/CameronBishop/detect_chimeric_sequences). Putative chimeric read pairs and reads were subsequently interrogated using Integrative Genomics Viewer version 2.9.4, BioEdit version 7.2.5, Genome Analysis Toolkit version 4.2.4.1 and R version 4.1.0.

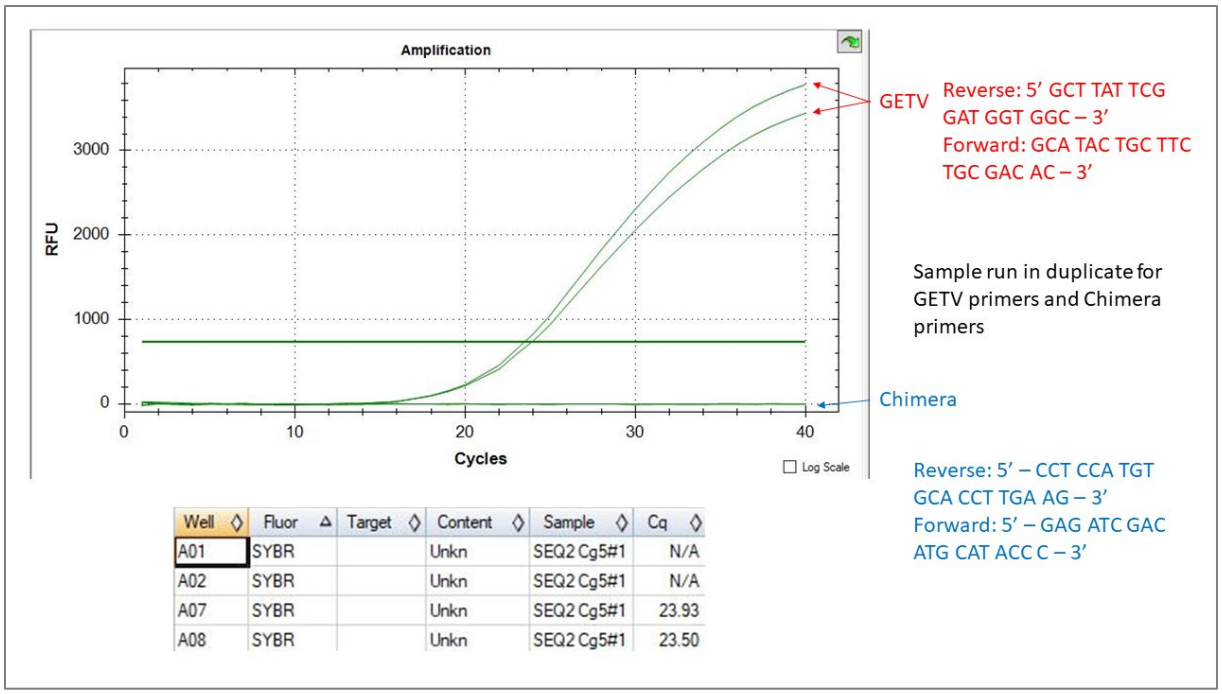

Figure S7. Quantitative RT-PCR detection of putative chimeric reads. Results are shown for the sample (IFNAR mouse 2 dpi) from which chimeric read 6 was obtained. This is the chimeric read with no inter read distance (for details see Fig. S6). Primers were designed to amplify this exact chimeric read. GETV primers served as positive control for qRT-PCR amplification of viral RNA. Chimeric read 6 cannot be amplified and most likely is an artefact.

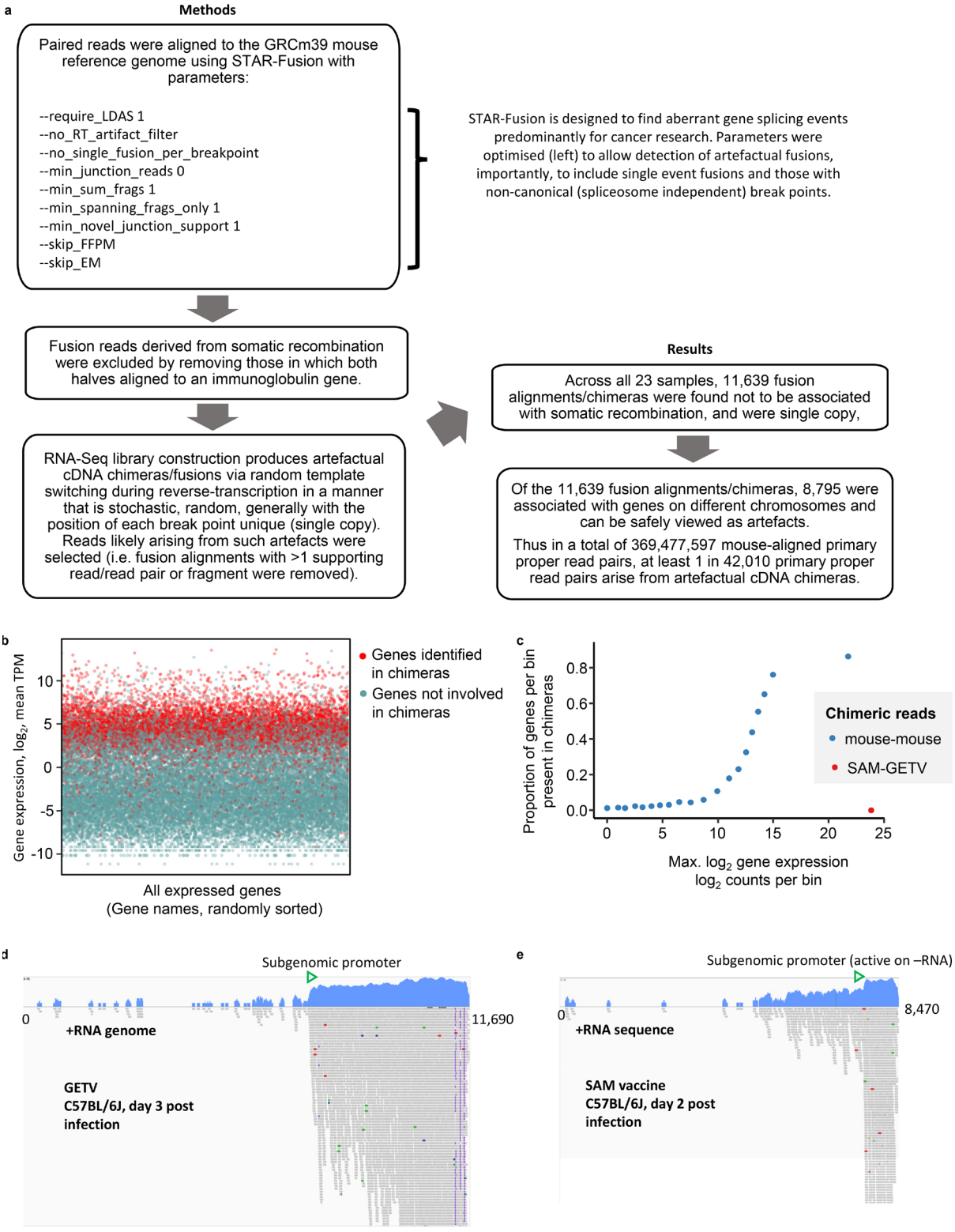

Figure S8. Artefactual mouse chimeric reads generated during cDNA synthesis. a, Process description of STAR-Fusion, illustrating that at least 1 in 42,010 of total mouse reads represent fusion/chimeric read pairs. These likely arise from generation of cDNA artefacts during reverse transcription (RT) [1]. **b**, The expression levels of mouse genes that are present in chimeric reads are higher than expression levels of mouse genes not present in in chimeric reads. **c**, Derived

from b by dividing y axes into 20 bins (and using raw counts rather than TPM). The greater the abundance of an mRNA species (high expression), the greater the chance that it will be incorporated into an artefactual cDNA chimera/fusion during RT. Curve is consistent with a stochastic process[1]. Importantly, the SAM-GETV chimeric read frequency is much lower than the artefactual mouse-mouse chimeric read frequencies. The SAM-GETV chimeric reads are thus highly likely to also be artefacts. **d,** Views of GETV reads aligned to the GETV genome from one mouse. Each grey cigar represents a read. The blue graphs show log read coverage at each nucleotide position (genome is 11, 690 bp). As expected, there are many more reads aligning to the subgenomic RNA encoding the structural genes. The subgenomic promoter operates during synthesis of +RNA from a -RNA template. The less abundant full length +RNA (also generated from the –RNA template) is initiated at the 3′UTR. **e**, As for (d) but for SAM vaccine alignments, where the subgenomic promoter drives production of +RNA carrying the GOI (mCherry).

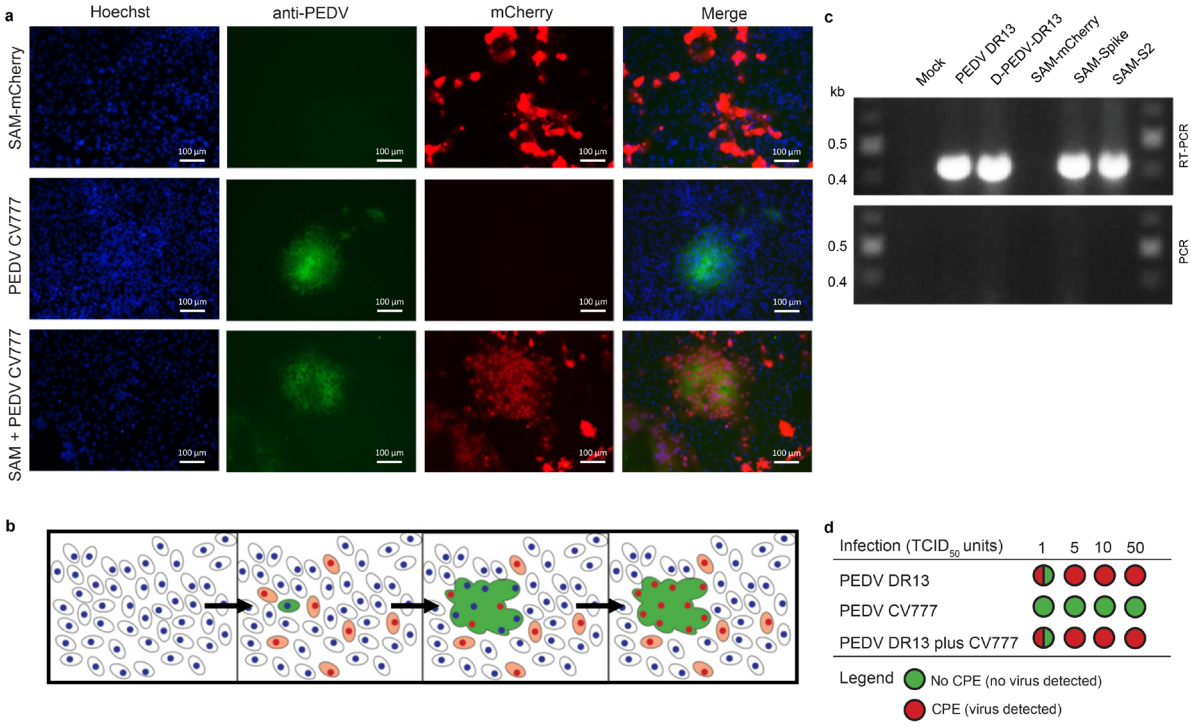

Figure S9. Evaluating recombination between PEDV CV777 and SAM vaccine. a, Fluorescence microscopy of co-transfected/infected Vero cell. Vero cells were transfected with SAM vaccine encoding mCherry and infected with PEDV CV777 (0.1 $TCID_{50}$/cell) 24 hours post transfection. Cell nuclei were stained with Hoechst, PEDV CV777 stained with mouse anti-PEDV spike, and SAM vaccine was visualized by expression of mCherry. **b**, Schematic representation of the development of PEDV CV777 syncytia expressing SAM encoded mCherry. PEDV CV777 infected cells formed large syncytia with SAM-transduced cells. Thereby, the SAM-induced mCherry expression can be detected in the syncytia, visualized by mCherry-positive nuclei in the syncytia. **c**, Detection of spike S2 domain RNA after transfection of SAM templates (SAM-mCherry, SAM-Spike and SAM-S2) and defective PEDV DR13 (D-PEDV-DR13) RNA template in Vero cells. RNA was extracted from cells, DNase treated, and used as input for RT-PCR and PCR analysis using specific spike S2 sequence primers (430 bp product). The presence of the spike RNA after transfection was confirmed by a RT-PCR (amplification of RNA and DNA) product in combination with the absence of a PCR (amplification DNA) product. **d**, Sensitivity of the cell-based coronavirus recombination detection assay. PEDV CV777 was unable to replicate and induce CPE in 10^5 Vero cells in the absence of trypsin, whereas PEDV DR13 was able to replicate and induce CPE within 4 serial passages even at low MOI and in the presence of an excess of PEDV CV777 (10^5 $TCID_{50}$/ml), illustrating the assay's sensitivity and suitability.

References

1 Yan, B. *et al.* Host-virus chimeric events in SARS-CoV-2-infected cells are infrequent and artifactual. *J Virol* **95**, e0029421 (2021). https://doi.org:10.1128/JVI.00294-21

2 Visser, L. J. *et al.* Essential role of enterovirus 2A protease in counteracting stress granule formation and the induction of type I interferon. *J Virol* **93** (2019). https://doi.org:10.1128/JVI.00222-19

3 Poo, Y. S. *et al.* Multiple immune factors are involved in controlling acute and chronic chikungunya virus infection. *PLoS Negl Trop Dis* **8**, e3354 (2014). https://doi.org:10.1371/journal.pntd.0003354

4 Wilson, J. A. *et al.* RNA-Seq analysis of chikungunya virus infection and identification of granzyme A as a major promoter of arthritic inflammation. *PLoS Pathog* **13**, e1006155 (2017). https://doi.org:10.1371/journal.ppat.1006155