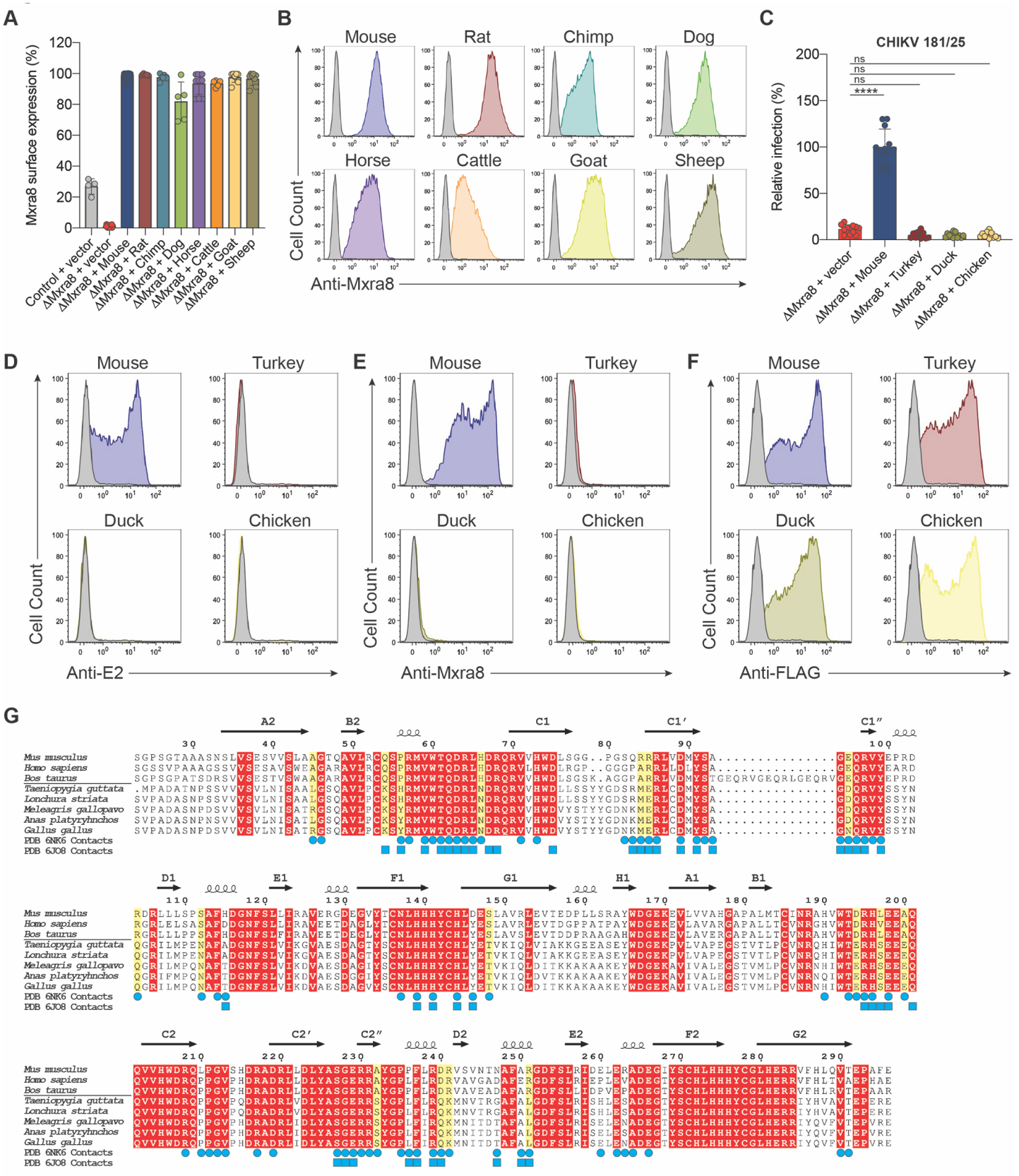

Figure S1. Expression and infection of mammalian and avian Mxra8 orthologs in Δ*Mxra8* 3T3 cells, Related to Figure 1. Data are pooled from two to ten experiments (n = 4 to 18 replicates) (**A**) and representative flow cytometry plots (**B**) showing cell surface expression of mouse, rat, chimpanzee, dog, horse, cattle, goat, and sheep Mxra8 after lentivirus trans-complementation of Δ*Mxra8* 3T3 cells and staining with species cross-reactive anti-Mxra8 mAbs. **C-D**. Lentivirus complementation of Δ*Mxra8* 3T3 with Mxra8 cDNA from mouse, turkey, duck, or chicken. Cells were inoculated with CHIKV (181/25) and analyzed by staining with anti-E2 mAbs. Data

are from four experiments (n = 12 replicates; one-way ANOVA with Dunnett's post-test: ****, $P < 0.0001$). **E-F**. Representative flow cytometry plots showing Mxra8 surface expression of mouse, turkey, duck, and chicken Mxra8 with (**E**) a pool of anti-Mxra8 mAbs or (**F**) anti-FLAG mAb. **G.** Structure-based alignment of mouse (*Mus musculus*), human, (*Homo sapiens*), cattle (*Bos taurus*), zebra finch (*Taeniopygia guttata*), striated finch (*Lonchura striata*), turkey (*Meleagris gallopavo*), duck (*Anas platyrhynchos*), and chicken (*Gallus gallus*) using ALINE; finches can act as amplifying hosts for some encephalitic (*e.g.*, WEEV) but not arthritogenic alphaviruses. The black line in the species names indicates grouping of mammals (*top*) and birds (*bottom*). Red boxes indicate conserved residues, white boxes indicate non-conserved residues, and yellow boxes indicate CHIKV contact residues that are conserved differently in mammals or birds. Secondary structure was assigned using DSSP and indicated above the sequence. Blue circles and blue squares represent mouse Mxra8 (PDB 6NK6) and human MXRA8 (PDB 6JO8) contact residues (>50% buried surface area) with the CHIKV E2-E1 heterodimer, respectively (Basore et al., 2019; Song et al., 2019).

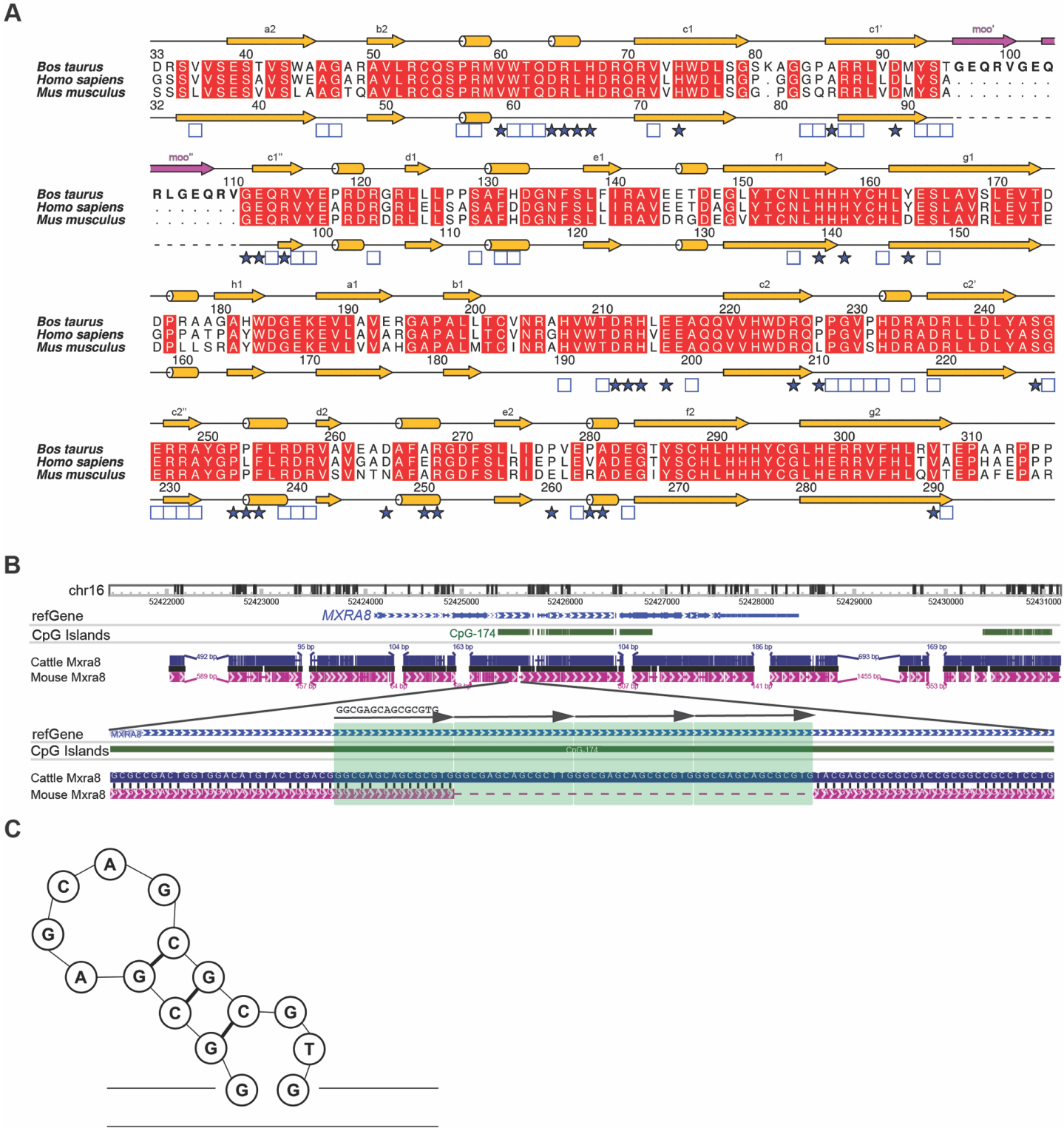

Figure S2. Sequence alignments and analysis of Mxra8, Related to Figure 2. A. Structure-based alignment of mouse, human, and cattle Mxra8 using ALINE. Red boxes indicate conserved residues and white boxes indicated non-conserved residues. Secondary structure was assigned using the DSSP algorithm and is shown in yellow for cattle (top) and mouse (bottom). The β-strands are labeled above the mouse secondary structure according to standard convention. The 15-amino acid cattle Mxra8 insertion is indicated in magenta. The symbols below the alignment indicate mouse Mxra8 contact residues with the CHIKV E2-E1 heterodimer (PDB 6NK6) as calculated by PDBePISA. Open boxes represent 10-40% buried surface area and stars represent 50-90% buried surface area as defined previously (Basore et al., 2019). **B**. Genome comparison of the cattle (blue) and mouse (magenta) Mxra8 gene using Washington University Epigenome Browser (https://epigenomegateway.wustl.edu). Predicted CpG islands are indicated in dark green. The primary 15-nucleotide sequence and the three copies of

the 15-nucleotide tandem repeat insertion are highlighted in green in the bottom panel. **C**. Predicted DNA secondary structure of cattle Mxra8 amino acids ^{111}GEQRV^{115} shows the formation of a DNA loop in the sense strand of the double stranded DNA. DNA secondary structure was generated using mfold (Zuker, 2003). During cell replication, polymerase slippage and subsequent reattachment may cause a bubble, or a single-stranded DNA, to form in the newly synthesized strand. Slippage is thought to occur in sections of DNA with tandem repeat patterns, such as those in the *Mxra8* gene. The single stranded DNA repeat is predicted to form a stem-loop structure, which potentially increases the likelihood of formation and stabilization of a bubble. DNA repair mechanisms subsequently realign the template with the new strand resulting in the straightening and removal of the bubble. Thus, DNA polymerase slippage can cause the newly created DNA strand to contain an expanded section, such as the 45-nucleotide insertion in cattle *Mxra8*.

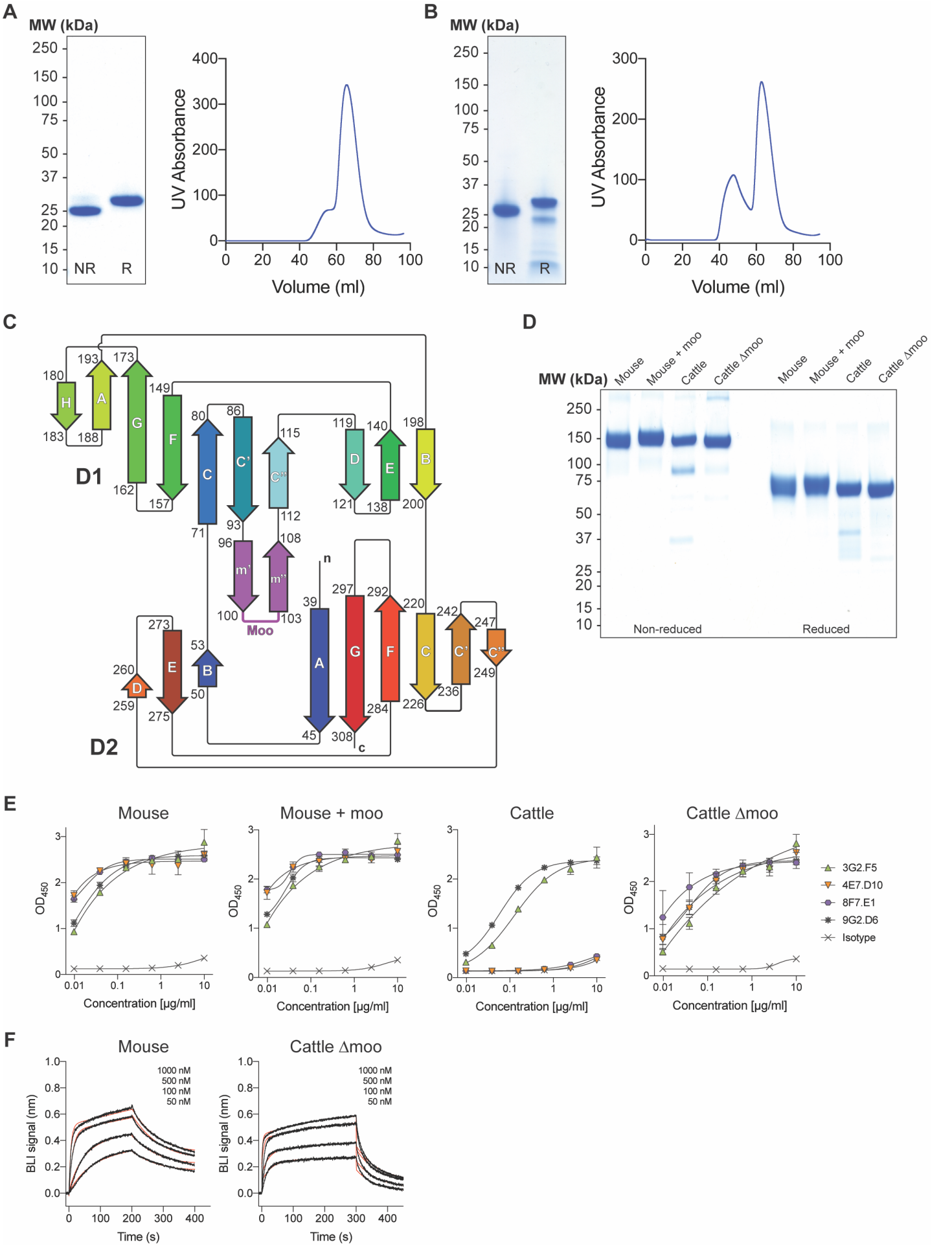
A
MW (kDa)
250
150
100
75
50
37
25
20
15
10
NR
R
UV Absorbance
400
300
200
100
0
0 20 40 60 80 100
Volume (ml)
B
MW (kDa)
250
150
100
75
50
37
25
20
15
10
NR
R
UV Absorbance
300
200
100
0
0 20 40 60 80 100
Volume (ml)
C
D1
D2
H
A
G
F
C
C'
C''
D
E
B
m'
m''
Moo
n
c
180
183
193
188
173
162
149
157
80
71
86
93
115
112
96
100
108
103
119
121
140
138
198
200
39
45
297
308
292
284
220
226
242
236
247
249
273
275
53
50
260
259
D
E
B
A
G
F
C
C'
C''
D
E
B
A
G
F
C
C'
C''
D
E
B
m'
m''
Moo
n
c
H
A
G
F
C
C'
C''
D
E
B
D1
D2
D
MW (kDa)
Mouse
Mouse + moo
Cattle
Cattle Δmoo
Mouse
Mouse + moo
Cattle
Cattle Δmoo
250
150
100
75
50
37
25
20
15
10
Non-reduced
Reduced
E
Mouse
Mouse + moo
Cattle
Cattle Δmoo
OD450
3
2
1
0
0.01 0.1 1 10
Concentration [µg/ml]
3G2.F5
4E7.D10
8F7.E1
9G2.D6
Isotype
F
Mouse
Cattle Δmoo
BLI signal (nm)
1.0
0.8
0.6
0.4
0.2
0.0
1000 nM
500 nM
100 nM
50 nM
0 100 200 300 400
Time (s)

Figure S3. Purification and structural topology of mouse and cattle Mxra8 protein variants, Related to Figures 2 and 3. A-B. Mxra8 ectodomain of (**A**) mouse or (**B**) cattle was expressed in bacteria, oxidatively refolded, and purified by size exclusion chromatography. (*Left*) Coomassie-stained SDS-PAGE of refolded mouse Mxra8 under non-reducing and reducing conditions. (*Right*) Size exclusion chromatography profile of Mxra8 proteins. **C**. Topology diagram of cattle Mxra8. The β-strands of each Ig domain are labeled according to standard convention. The 15-residue 'moo' insertion is colored purple and forms β-strands (moo' [m'] and moo'' [m'']). The two Ig domains are labeled D1 and D2. The N- and C-termini of the Mxra8 protein are labeled in lowercase. **D**. Coomassie-stained SDS-PAGE of mouse, mouse + moo, cattle, and cattle ∆moo Mxra8-Fc proteins under non-reducing and reducing conditions. **E.** Binding of increasing concentrations of anti-Mxra8 mAbs 3G2.F5, 4E7.D10, 8F7.E1, 9G2.D6, and isotype control mAb to adsorbed mouse, mouse + moo, cattle, and cattle ∆moo Mxra8-Fc fusion proteins by ELISA. Data are pooled from four experiments performed in duplicate. **F.** Kinetic sensograms of mouse (left) and cattle ∆moo (right) Mxra8 binding to CHIKV VLPs fit to a 1:1 binding model. Raw experimental traces (1000, 500, 100, and 50 nM) are shown in black, and fit traces are shown in red. Data are representative of three (cattle ∆moo) to six (mouse) experiments.

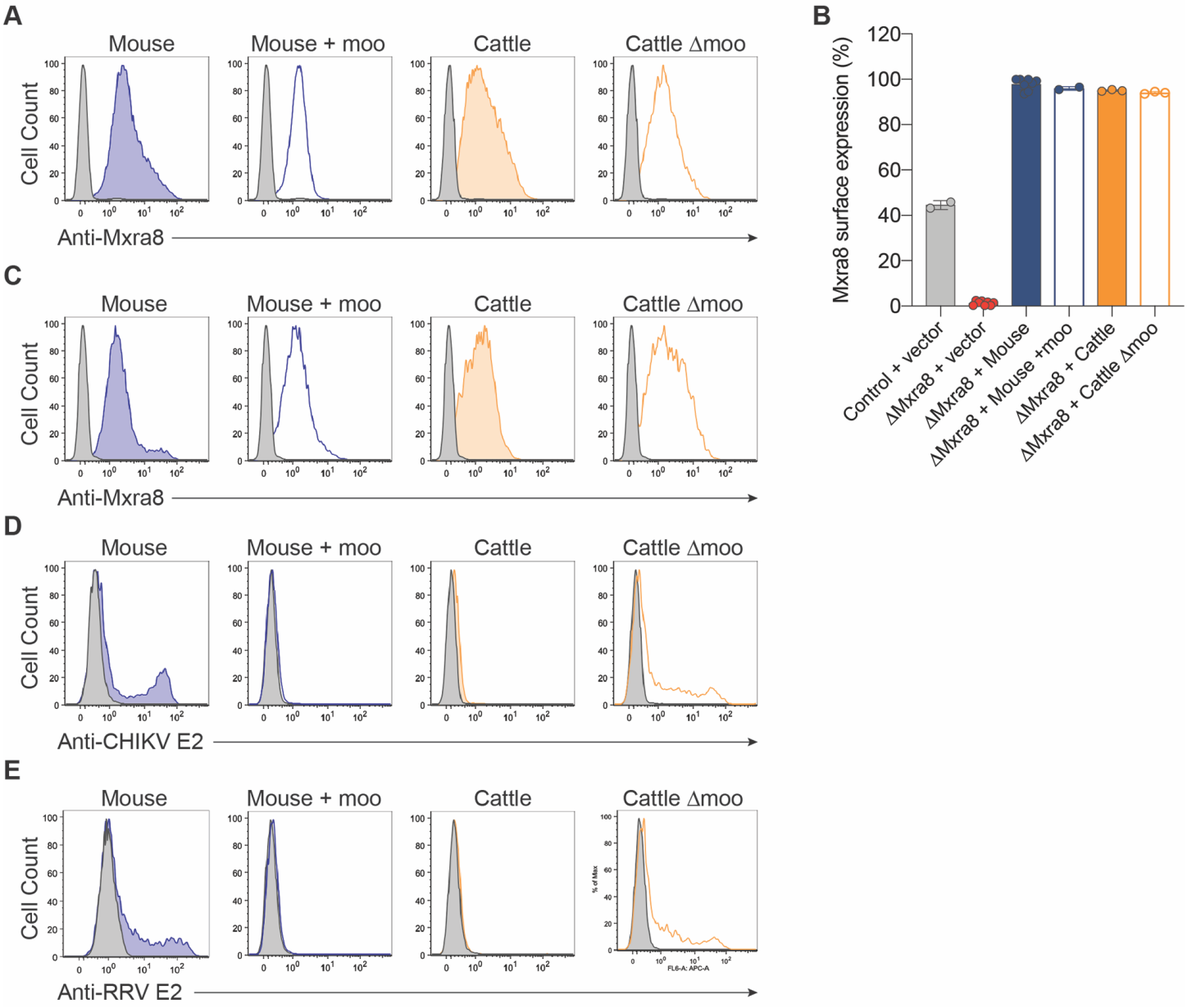

Figure S4. Expression of mouse and cattle Mxra8 variants on the surface of Δ*Mxra8* 3T3 cells and bovine cells, Related to Figure 3. A. 3T3 ΔMxra8 cells were complemented with empty vector, mouse, mouse + moo, cattle or cattle Δmoo Mxra8 and stained for Mxra8 surface expression using a pool of anti-Mxra8 mAbs. **B**. Representative flow cytometry histograms showing cell surface expression of mouse, mouse + moo, cattle, and cattle Δmoo Mxra8 variants. Data are representative of three experiments. **C.** Flow cytometry histograms showing cell surface expression of mouse, mouse + moo, cattle, and cattle Δmoo Mxra8 variants after lentivirus complementation of bovine corneal cells using a pool of anti-Mxra8 mAbs. Data are representative of three experiments. Histograms for wild-type bovine corneal cells are shown in gray. **D-E**. Flow cytometry histograms showing (**D**) CHIKV or (**E**) RRV infection of cow cells complemented with mouse, mouse + moo, cattle, and cattle Δmoo Mxra8 gene variants. Histograms for wild-type bovine corneal cells are shown in gray. Data are representative of three experiments.

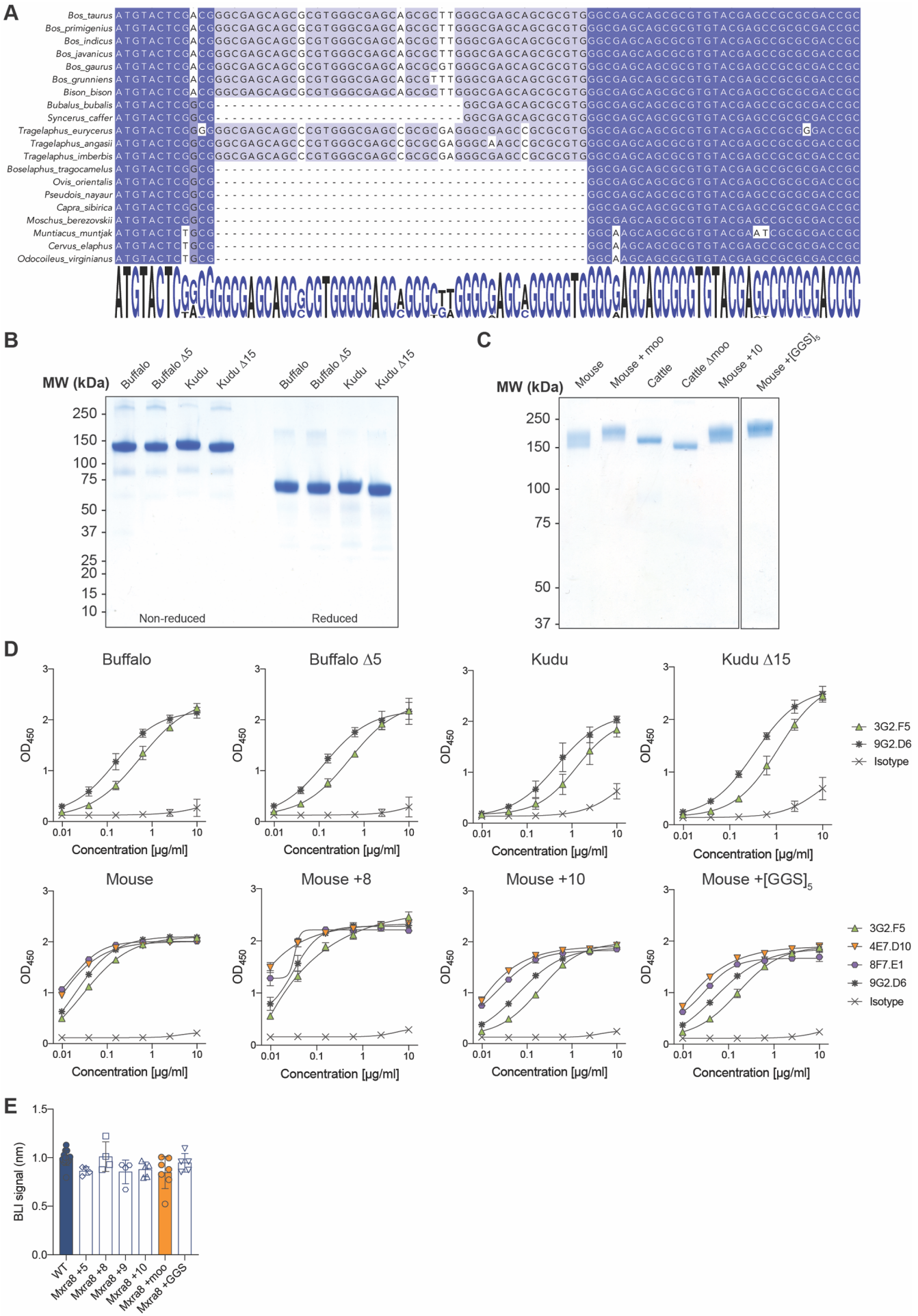
A
Bos_taurus
Bos_primigenius
Bos_indicus
Bos_javanicus
Bos_gaurus
Bos_grunniens
Bison_bison
Bubalus_bubalis
Syncerus_caffer
Tragelaphus_eurycerus
Tragelaphus_angasii
Tragelaphus_imberbis
Boselaphus_tragocamelus
Ovis_orientalis
Pseudois_nayaur
Capra_sibirica
Moschus_berezovskii
Muntiacus_muntjak
Cervus_elaphus
Odocoileus_virginianus
B
MW (kDa)
Buffalo
Buffalo Δ5
Kudu
Kudu Δ15
250
150
100
75
50
37
25
20
15
10
Non-reduced
Reduced
C
MW (kDa)
Mouse
Mouse + moo
Cattle
Cattle Δmoo
Mouse +10
Mouse +[GGS]5
250
150
100
75
50
37
D
Buffalo
Buffalo Δ5
Kudu
Kudu Δ15
Mouse
Mouse +8
Mouse +10
Mouse +[GGS]5
OD450
Concentration [μg/ml]
0.01
0.1
1
10
3G2.F5
9G2.D6
Isotype
4E7.D10
8F7.E1
E
BLI signal (nm)
1.5
1.0
0.5
0.0
WT
Mxra8 +5
Mxra8 +8
Mxra8 +9
Mxra8 +10
Mxra8 +moo
Mxra8 +GGS

Figure S5. Expression and antigenic characterization of water buffalo, kudu, and mouse Mxra8-Fc protein variants, Related to Figures 5 and 6. A. Nucleotide sequence alignments of *Mxra8* sequences from Bovidae family members in the region of the insertion site in D1. The sequences were obtained after assembly of deposited sequences (**Table S4**) or extraction of mRNA and primary sequencing (**Table S5** and **STAR Methods**), aligned using MUSCLE, and visualized using Jalview. Nucleotide consensus plot was generated using WebLogo3. For species with both types of data, we used sequencing from tissue samples as primary data and WGS data as confirmation. Five versions of the Mxra8 alignment were generated: (1) the complete alignment contains all sites for all species; (2) the trimmed alignment is identical to the complete alignment, but removes all sites following the first stop codon (*Bos taurus* genomic coordinates 16:51,173,039 – 51,176,528); (3) the no-insertion alignment contains all sites from the trimmed alignment except for the 45 nucleotide insertion for a total of 1344 nucleotides [*Bos taurus* 45-nucleotide insertion genomic coordinates 16:51,173,324 – 51,173,368]; (4) the insertion alignment includes only the 45-nucleotide Bovinae insertion plus the GEQRV repeat unit conserved across all mammals for a total of 60 nucleotides (*Bos taurus* insertion + GEQRV genomic coordinates 16:51,173,324 – 51,173,383); and (5) the insertion + flank alignment includes the same sites as the insertion alignment plus two 30-nucleotide flanking regions for a total of 120 nucleotides (*Bos taurus* insertion + flank genomic coordinates 16:51,173,294 – 51,173,413). **B**. Coomassie-stained SDS-PAGE of water buffalo, water buffalo ∆5, kudu, and kudu ∆15 Mxra8-Fc proteins under non-reducing and reducing conditions. **C**. Coomassie-stained SDS-PAGE of mouse, mouse + moo, cattle, cattle ∆moo, mouse +10, and mouse+$[GGS]_5$ Mxra8-Fc proteins under non-reducing conditions. **D.** Binding of increasing concentrations of anti-Mxra8 mAbs 3G2.F5, 9G2.D6, and isotype control to adsorbed water buffalo, water buffalo ∆5, kudu, kudu ∆15, mouse, mouse +8, mouse +10, and mouse +$[GGS]_5$ Mxra8-Fc fusion proteins by ELISA. Data are pooled from two to three experiments performed in duplicate. **E**. Binding of anti-Mxra8 mAb 9G2.D6 to bacterially-derived mouse Mxra8 and insertion variants Mxra8+5, Mxra8+8, and Mxra8+9, Mxra8+10, Mxra8 +moo, and Mxra8 +$[GGS]_5$ by BLI. Data are the mean and standard deviation of four to eight experiments.

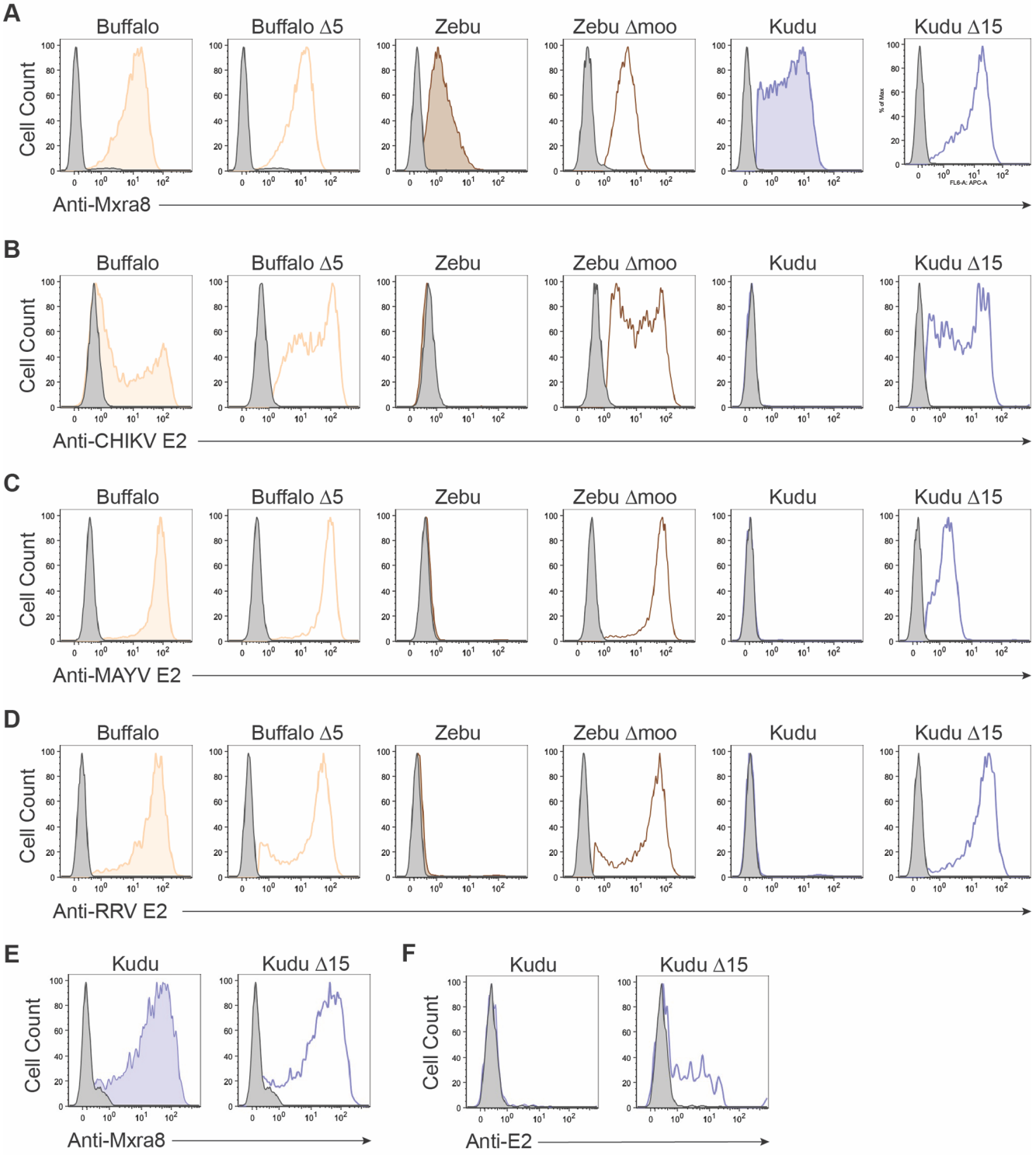

Figure S6. Cell surface expression and alphavirus infection of water buffalo, zebu, and kudu Mxra8 variants, Related to Figure 6. A. Flow cytometry histograms showing cell surface expression of water buffalo, water buffalo Δ5, zebu, zebu Δmoo, kudu, and kudu Δ15 Mxra8 after lentivirus complementation of Δ*Mxra8* 3T3 cells using a pool of anti-Mxra8 mAbs. Data are representative of three experiments. **B-D.** Flow cytometry histograms showing (**B**) CHIKV, (**C**) MAYV, or (**D**) RRV infection of 3T3 Δ*Mxra8* complemented with water buffalo, water buffalo Δ5, zebu, zebu Δmoo, kudu, and kudu Δ15 Mxra8 gene variants. Data are representative of three experiments. **E.** Flow cytometry histograms showing cell surface expression of kudu and kudu Δ15 Mxra8 after lentivirus transduction of primary kudu fibroblasts. Cell surface expression of Mxra8 was detected using a pool of anti-Mxra8 mAbs. Data are representative of two experiments. **F.** Flow cytometry histograms showing MAYV infection of primary kudu fibroblasts complemented with kudu and kudu Δ15 *Mxra8* gene variants. Data are representative of four experiments.

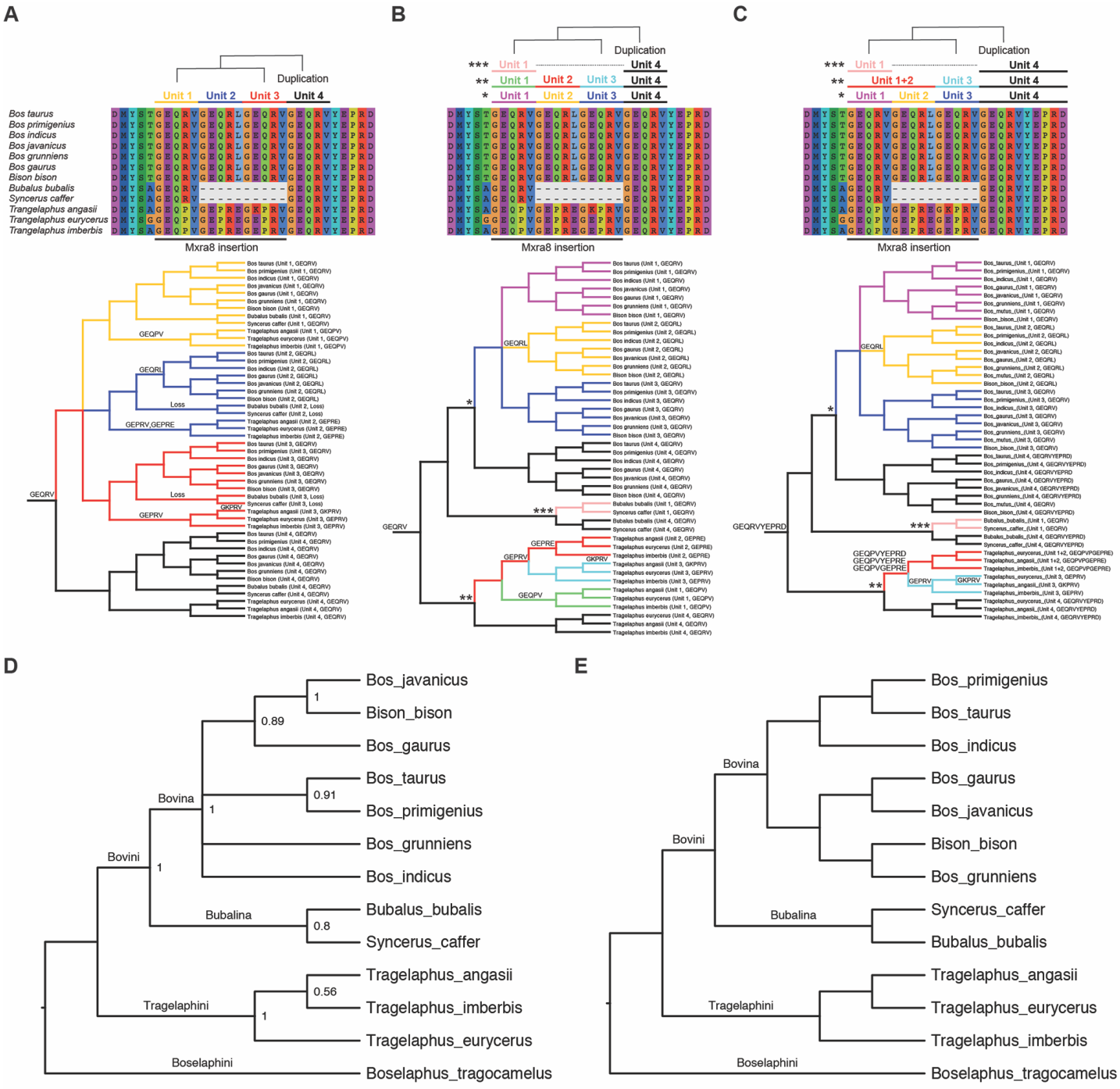

Figure S7. Evolutionary history of Mxra8 insertion, Related to Figure 6. A-C. Three possible evolutionary histories of unit duplication within the Mxra8 insertion, which were reconstructed under maximum parsimony. The sequence alignments (*top panels*) indicate the series of duplications of the insertion repeat units. The Mxra8 insertion is composed of one to three length-5-amino acid sequences (repeat units), where each unit resembles either the length-5 (GEQRV) or potentially the length-10 (GEQRVYEPRD) unit flanking the 3' end of the insertion that is conserved in all sampled Bovidae species. The Mxra8 insertion is likely derived from this conserved unit through a complex sequence of duplication, loss, and substitution events. To our knowledge, we are not aware of any software suited to reconstructing the complex substitution-duplication-loss history of the Mxra8 insertion due to the short length of the repeat unit (GEQRV). Hence, the substitution-duplication-loss history was manually reconstructed under maximum parsimony in Mesquite. The tips of any historical scenario correspond to repeat units that are homologous through duplication events (i.e., paralogous; colors change at bifurcation event) or speciation events (i.e., orthologous; colors constant at bifurcation event). The sequence alignments above each evolutionary history show how the repeat units relate to the Mxra8 insertion alignment. We considered three scenarios to assert the topology for the repeat unit paralogs. Under the L5 scenario (**A**), where duplication events

are rare and loss events are common, 3 duplication, 2 loss, and 6 nonsynonymous substitution events were reconstructed. Under the D5 scenario (**B**), where duplication events are common and loss events are rare, at least 7 duplication, 0 loss, and 5 nonsynonymous substitution events were reconstructed. Under the D10 scenario (**C**), where the Tragelaphini (nyala and bongo) insertion was duplicated from Unit 4, 6 duplication, 0 loss, and 6 nonsynonymous substitution events were reconstructed. All three phylogenetic histories require at least 5-amino acid substitutions, with the majority (>80%) occurring within older Tragelaphini lineages. Loss of insertion unit(s) or changes in the insertion sequence are indicated at branch points and branches (**A-C**, *bottom panels*). To reflect three duplication histories, the Mxra8 insertion and insertion + flank alignments were restructured further to correctly assign homology to sites belonging to the paralogous repeat units within the insertion. In the restructured format, each row of the insertion alignment corresponded to one repeat unit from one species. For the insertion + flank alignment was processed in the same way, except the flanking regions were concatenated to the left and right of the progenitor GEQRV repeat unit. **D-E**. Topologies for the Mxra8 gene (**D**) and species (**E**) trees are congruent for backbone relationships among Bovina, Bubalina, Tragelaphini, and Boselaphini. Gene tree topology shows clades with posterior support of $P > 0.5$ (node values). The unrooted Mxra8 gene tree topology from the trimmed alignment of the Bovidae gene sequences described in **Table S4** was estimated using RevBayes. Nucleotide evolution was modeled by an HKY substitution process with flat Dirichlet priors assigned to the exchangeability rates and base frequencies. Site-rate variation was modeled by a discrete +Γ4 model with shape and scale parameters following an exponential prior density with rate of 0.1. Relative prior branch lengths followed a flat Dirichlet distribution, which were multiplied by the tree length, L ~ Exponential(10), to model actual branch lengths. The gene was partitioned by codon site position, where the relative clock rate for each partition was modeled by a lognormal prior density with log-mean equal to 1 and log-sd equal to 0.5. The tree was rooted with *Boselaphus tragocamelus* as the outgroup (Zurano et al., 2019). Analysis scripts and the sequence alignment are available at: https://github.com/mlandis/mxra8_bovinae. The estimated gene tree topology was compared to a synthetic species tree topology that we constructed by grafting the phylogenomic relationships among *Bos taurus*, *Bos indicus*, *Bos grunniens*, *Bison bison*, and *Bison bonasus* inferred by others (Wang et al., 2018) into the broader Bovinae species relationships estimated by (Zurano et al., 2019). **Fig S7D** presents a majority rule consensus topology ($p > 0.5$) for the Mxra8 gene tree, which shows that the backbone relationships among Bos + Bison, Bubalus + Syncerus, and Tragelaphini have high posterior support ($p > 0.99$) and are congruent with accepted species tree relationships (**Fig S7E**).

Table S1. Amino acid and nucleotide identity and similarity of Mxra8 orthologs, Related to Figure 1.

	Mouse	Rat	Human	Chimp	Dog	Horse	Cattle	Goat	Sheep	Turkey	Duck	Chicken
Mouse		93.9	78.2	77.5	79.8	82.9	76.5	79.6	79.7	58.4	56.3	58.7
Rat	96.6		79.1	78.4	80.7	83.1	76.1	79.4	79.5	58.4	56.5	58.7
Human	84.5	84.5		98.6	82.7	84.2	80.0	82.9	83.3	59.6	57.8	59.6
Chimp	84.0	84.0	99.1		82.0	83.8	79.3	82.3	82.6	59.1	57.1	59.1
Dog	86.3	86.5	88.1	87.6		88.1	83.7	86.3	86.4	60.8	59.9	60.8
Horse	87.2	87.4	89.0	88.5	92.1		85.6	88.7	89.1	60.0	58.7	60.0
Cattle	81.9	81.5	83.4	83.0	88.0	88.4		93.3	93.1	57.9	55.7	57.9
Goat	85.1	84.7	86.3	85.8	90.5	91.6	94.4		98.7	59.7	57.7	59.7
Sheep	85.3	84.9	86.4	86.0	90.7	92.2	94.4	98.9		60.2	58.2	60.2
Turkey	71.5	71.9	71.5	71.2	72.0	72.8	69.7	71.2	71.6		92.7	99.3
Duck	68.9	69.4	69.2	69.0	70.0	70.8	67.5	69.0	69.3	94.3		92.0
Chicken	71.5	71.9	71.5	71.2	72.0	72.8	69.7	71.2	71.6	100.0	94.3	

	Mouse	Rat	Human	Chimp	Dog	Horse	Cattle	Goat	Sheep	Turkey	Duck	Chicken
Mouse	100.0	92.7	77.6	77.3	78.9	80.6	75.1	77.3	77.7	62.2	63.1	62.9
Rat		100.0	77.8	77.6	78.7	80.5	75.1	77.5	77.9	62.1	63.2	62.7
Human			100.0	99.2	84.5	86.0	80.7	82.9	83.2	62.3	64.9	62.6
Chimp				100.0	84.0	85.8	80.5	82.6	83.0	61.9	64.7	62.2
Dog					100.0	87.4	82.4	84.7	85.0	63.7	65.9	64.4
Horse						100.0	84.3	86.6	87.0	63.5	65.9	64.0
Cattle							100.0	92.9	93.4	60.3	62.4	61.2
Goat								100.0	98.7	61.9	63.9	62.5
Sheep									100.0	62.4	64.4	63.0
Turkey										100.0	88.8	97.3
Duck											100.0	89.2
Chicken												100.0

(*Top*) Amino acid sequences were aligned using MUSCLE and amino acid identity (red) and similarity (yellow) were determined using Ident and Sim. (*Bottom*) Nucleotide sequence identity (red) of Mxra8 orthologs was determined using MUSCLE and Ident and Sim.

Table S2. Crystallographic data collection and refinement statistics for cattle Mxra8, Related to Figure 2.

PDB ID code	6ORT
[a]Resolution range	48.47 - 2.30 (2.38 - 2.30)
Space group	P 6(5) 2 2
Unit cell (Å) a, b, c	77.53, 77.53, 242.35
Total number of reflections	1,434,086 (146,561)
Unique reflections	20,172 (1,913)
Multiplicity	71.1 (76.6)
Completeness (%)	100.0 (100.0)
Mean I/σ(I)	49.2 (3.7)
Wilson B-factor	49.9
R_{merge}	0.146 (2.234)
$CC_{1/2}$	1.000 (0.93)
Reflections used in refinement	19,084 (1,842)
Reflections used for R_{free}	1,004 (97)
R_{work}	0.2180 (0.2683)
R_{free}	0.2416 (0.2899)
Number of non-hydrogen atoms	2,364
protein	2,229
solvent	117
Protein residues	277
RMS(bonds) (Å)	0.002
RMS(angles) (°)	0.44
Ramachandran favored (%)	96.73
Ramachandran allowed (%)	3.27
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
Clashscore	2.29
Average B-factor(Å^2)	51.7
protein	51.7
solvent	49.8

[a]Values in parentheses refer to the highest resolution shell. Data was collected at ALS Beamline 4.2.2 using an RDI CMOS_8M detector. Data processing was carried out in XDS and refinement in Phenix (Adams et al., 2010).

Table S3. Quantitative analysis of mouse and cattle ∆moo Mxra8 binding to CHIKV VLPs by biolayer interferometry, Related to Figure 3.

	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	$t_{1/2}$ (s)	$K_{D, kinetic}$ (nM)	$K_{D, equilibrium}$ (nM)
Mouse	$(1.4 \pm 0.6) \times 10^{5}$	$(8.3 \pm 2.0) \times 10^{-3}$	86.9 ± 17.3	61.5 ± 13.9	66.4 ± 13.3
Cow ∆moo	$(2.8 \pm 0.7) \times 10^{5}$	$(1.8 \pm 0.3) \times 10^{-2}$	40.2 ± 8.3	64.2 ± 15.5	72.3 ± 19.9

Data are the mean and standard deviations of three independent experiments.

Table S4. Source and *Mxra8* coverage statistics from whole genome and deposited sequences, Related to Figure 5.

Genus	Species	Common Name	Study Accession	Run Accession	Average alignment coverage per base of insert/insert junction	Minimum number of reads aligned at any base of insert/insert junction	Amino acid sequence of insertion
Bos	taurus	Cow	N/A	NM_001075830	N/A[#]	N/A[#]	GEQRVGEQRLGEQRV
Bos	primigenius	Auroch	PRJNA294709	SRR2465682	Modi et al., 2004	Modi et al., 2004	GEQRVGEQRLGEQRV
Bos	indicus	Zebu	PRJNA360096	XM_019976191	N/A[#]	N/A[#]	GEQRVGEQRLGEQRV
Bos	javanicus	Banteng	PRJNA325061	SRR4035276	3.33	3	GEQRVGEQRLGEQRV
Bos	grunniens	Domestic yak	PRJNA359997	SRR5140177	6.93	6	GEQRVGEQRLGEQRV
Bison	bison	Bison	PRJNA257088	SRR1659060	N/A[%]	N/A[%]	GEQRVGEQRLGEQRV
Bubalus	bubalis	Water buffalo	PRJNA207334	AWWX01000000	Contigs*	Contigs*	GEQRV
Syncerus	caffer	African cape buffalo	PRJNA341313	SRR4104498	7.53	7	GEQRV
Tragelaphus	angasii	Nyala	PRJNA388863	SRR5647659	18.8	16	GEQPVGEPREGKPRV
Ovis	orientalis	Mouflon	PRJEB5463	ERR454948	11.4	10	No insertion
Pseudois	nayaur	Himalayan blue sheep	PRJNA361448	SRR5439716	7.85	7	No insertion
Capra	sibirica	Siberian ibex	PRJNA361447	SRR5260693	6	6	No insertion
Moschus	berezovskii	Dwarf musk deer	PRJNA289641	SRR2098995	256.7	248	No insertion
Cervus	elaphus	Red deer	PRJNA324173	SRR4013902	50.5	48	No insertion

N/A indicates that coverage statistics are not available as sequences were obtained from deposited NCBI sequences[#], RNAseq data[%], or sequence contigs*.

Table S5. Primer sequences and annealing temperatures used to amplify *Mxra8* from primary tissue samples, Related to Fig 5.

Genus	Species	Common Name	Nested Primer	Primer sequence	Annealing Temp (ºC)
Bos	taurus	Cow	Outer	FOR: GCGCCTCCGGGCCAGGCGGGCGCCATGGAG REV: CAGAGCTGCTGGCCCAGCCAGGAGCCCAGAGTC	65
			Inner	FOR: CGGGCCTGGGTCCTGCTCTGGAGACTTGTG REV: GCAGTACTCCTTCCTGAACTCTTTGTCCAAGTC	60
Bos	javanicus	Banteng	Outer	FOR: GCGCCTCCGGGCCAGGCGGGCGCCATGGAG REV: CAGAGCTGCTGGCCCAGCCAGGAGCCCAGAGTC	65
			Inner	FOR: CGGGCCTGGGTCCTGCTCTGGAGACTTGTG REV: GCAGTACTCCTTCCTGAACTCTTTGTCCAAGTC	60
Bos	gaurus	Gaur	Outer	FOR: GCGCCTCCGGGCCAGGCGGGCGCCATGGAG REV: CAGAGCTGCTGGCCCAGCCAGGAGCCCAGAGTC	65
			Inner	FOR: CGGGCCTGGGTCCTGCTCTGGAGACTTGTG REV: GCAGTACTCCTTCCTGAACTCTTTGTCCAAGTC	60
Bos	grunniens	Domestic yak	Outer	FOR: GCGCCTCCGGGCCAGGCGGGCGCCATGGAG REV: CAGAGCTGCTGGCCCAGCCAGGAGCCCAGAGTC	65
			Inner	FOR: CGGGCCTGGGTCCTGCTCTGGAGACTTGTG REV: GCAGTACTCCTTCCTGAACTCTTTGTCCAAGTC	60
Bison	bison	Bison	Outer	FOR: GCGCCTCCGGGCCAGGCGGGCGCCATGGAG REV: CAGAGCTGCTGGCCCAGCCAGGAGCCCAGAGTC	65
			Inner	FOR: CGGGCCTGGGTCCTGCTCTGGAGACTTGTG REV: GCAGTACTCCTTCCTGAACTCTTTGTCCAAGTC	60
Bubalus	bubalis	Water buffalo	Outer	FOR: GCGCCTCCGGGCCAGGCGGGCGCCATGGAG REV: CAGAGCTGCTGGCCCAGCCAGGAGCCCAGAGTC	65
			Inner	FOR: CGGGCCTGGGTCCTGCTCTGGAGACTTGTG REV: GCAGTACTCCTTCCTGAACTCTTTGTCCAAGTC	60
Syncerus	caffer	African cape buffalo	Outer	FOR: GCGCCTCCGGGCCAGGCGGGCGCCATGGAG REV: CAGAGCTGCTGGCCCAGCCAGGAGCCCAGAGTC	65
			Inner	FOR: CGGGCCTGGGTCCTGCTCTGGAGACTTGTG REV: GCAGTACTCCTTCCTGAACTCTTTGTCCAAGTC	60
Tragelaphus	eurycerus	Bongo	Outer	FOR: CCATCAGGGCCCGCGACCTCCGAC REV: CCCAGCCAGGAGCCCAGAGTCGCC	67
Tragelaphus	angasii	Nyala	Outer	FOR: GCGCCTCCGGGCCAGGCGGGCGCCATGGAG REV: CAGAGCTGCTGGCCCAGCCAGGAGCCCAGAGTC	68
Tragelaphus	imberbis	Lesser kudu	Outer	FOR: GCGCCTCCGGGCCAGGCGGGCGCCATGGAG REV: CAGAGCTGCTGGCCCAGCCAGGAGCCCAGAGTC	68
Boselaphus	tragelaphus	Nilgai	Outer	FOR: CCATCAGGGCCCGCGACCTCCGAC REV: CCCAGCCAGGAGCCCAGAGTCGCC	67
Muntiacus	muntjak	Indian muntjac	Outer	FOR: GCCATGGAGCTGCGGGCCTGGGTCCTGCTC REV: CAGAGCTGCTGGCCCAGCCAGGAGCCTGGAG	65
			Inner	FOR: CTTGTGCTTCTGCAGAGTTCTGCCGTC REV: GCAGTACTCCTTCCTGAACTCTTTGTCCAAG	60
Odocoileus	virginianus	White-tailed deer	Outer	FOR: GCCATGGAGCTGCGGGCCTGGGTCCTGCTC REV: CAGAGCTGCTGGCCCAGCCAGGAGCCTGGAG	68
			Inner	FOR: CTTGTGCTTCTGCAGAGTTCTGCCGTC REV: GCAGTACTCCTTCCTGAACTCTTTGTCCAAG	60

Table S6. Summary of Bovinae Mxra8 insertion evolution, Related to Figure 6.

Evolutionary scenario	Duplication or loss events preferred?	Length of repeat unit	Minimum event counts			
			Duplication	Loss	Nonsynonymous substitution	Synonymous substitution
L5 insert only	Loss	5	3	2	6	1
D5 insert only	Duplication	5	7	0	5	1
D10 insert only	Duplication	5 or 10	6	0	6	2
L5 insert+flank	Loss	5	3	2	8	2
D5 insert+flank	Duplication	5	7	0	7	2
D10 insert+flank	Duplication	5 or 10	6	0	8	3

Evolutionary scenarios and minimum event counts correspond to **Fig S7A-C**. Three proposed evolutionary scenarios explain the distribution of Mxra8 insertions among Bovinae lineages: (1) adaptive inheritance; (2) neutral inheritance; or (3) adaptive introgression. Scenarios are distinguished by the order and geological timing of their evolutionary events and, in particular, whether the events are ancient (millions of years ago) or recent (within the past ten to hundred thousand years). Each evolutionary scenario involves the origin of the Mxra8 insertion one to three times before Bovini, Bubalina, and Tragelaphini originated, followed by an adaptive phase that imprinted detectable patterns of selection within the Mxra8 insertion. The “adaptive introgression” scenario also requires introgression of the insertion between Bovinae species. Under the “adaptive inheritance” scenario, the insertion originates and acquires an adaptive role before Bovinae (or, at latest, before Tragelaphini) first diversifies millions of years ago. The insertion, and any acquired amino acid substitutions, are then inherited vertically during speciation. Vertical transmission would implicate the species tree and Mxra8 gene tree to be largely congruent. Adaptive substitutions would fall primarily along deep branches of the phylogeny, with an imbalance of nonsynonymous or synonymous substitutions, depending on whether positive or purifying selection was in effect, respectively. Severe incongruence between the species tree and Mxra8 gene tree, failure to detect positive or purifying selection within the insertion, or disproportionate numbers of substitutions along terminal branches of the tree would render the “adaptive inheritance” scenario unlikely; these features were not observed in our reconstructions. Similar to the “adaptive inheritance” scenario, the “neutral inheritance” also involves the origin and inheritance of the insertion one to three times in Bovinae. In contrast with the “adaptive inheritance” scenario, the insertion would have faced neutral or nearly neutral selection pressures after it first originated. Like the “adaptive inheritance” scenario, the species tree and gene tree topologies would be congruent. If the insertion evolved under neutrality across all Bovinae lineages, both synonymous and nonsynonymous substitutions would be distributed randomly throughout the Bovinae phylogeny. Alternatively, the insertion might have evolved neutrally and only recently faced new selection pressures, in which case nonsynonymous substitutions should be concentrated in the terminal lineages of the phylogeny. However, we reconstructed at least five nonsynonymous and zero synonymous substitutions along internal branches, and only one nonsynonymous and one synonymous substitution along the terminal lineages. Therefore, the sequence data do not appear to support the “neutral inheritance” scenario. Finally, the “adaptive introgression” scenario is designed to allow the insertion to be far younger than the age of Bovinae and assumes that (1) the Mxra8 insertion was deleterious in Bovinae (i.e., insertions would have been rapidly lost) before acquiring a recent adaptive role; (2) as the insertion is found in most extant Bovinae, it must have occurred recently enough to have survived, but too recently to have been inherited vertically while Bovina, Bubalina, and Tragelaphini diversified; and (3) the insertion instead must have been inherited horizontally in the recent past, most likely through numerous introgression events between distantly related Bovinae lineages. If the introgression scenario was true, the Mxra8 gene tree topology should be incongruent with the species tree, and variation in chromosome count in Bovinae should be low. However, the gene tree and species tree estimates are topologically congruent and chromosome counts are highly variable in Bovinae. Thus, the “adaptive introgression” scenario appears unlikely. Of the three evolutionary scenarios considered, our reconstructions are most consistent with the “adaptive inheritance” scenario as described above. If the sequence identity of the Mxra8 insertion has indeed been shaped by selection, it remains to be determined as to what selective force or forces are responsible for those changes.

Table S7. Clade ages of Bovinae members, Related to Figure 6.

Clade	Minimum age (Ma)	Maximum age (Ma)	*Mxra8* insertion calibration scenario
Bovinae	14.0	17.6	-
Bovini + Tragelaphini	13.5	16.6	L5
Tragelaphini	6.5	8.1	D5, D10
Bovini	9.4	12.0	-
Bubalus + *Syncerus*	5.0	8.2	-
Bos	3.4	4.9	-
Bubalus	1.1	2.0	-

Minimum and maximum clade ages are equal to the 95% highest posterior densities estimated by Zurano et al., 2018. Calibration scenarios identify the minimum clade age that corresponds to each of the Mxra8 insertion histories shown in **Fig S7**.

Table S8. Chromosome counts and geographical ranges for selected Mxra8 species, Related to Figure 6.

Family	Subfamily	Tribe	Species name	Chromosome count (2N)	Geographical range
Bovidae	Bovinae	Tragelaphini	*Tragelaphus angasii*	55	South Africa
Bovidae	Bovinae	Tragelaphini	*Tragelaphus imberbis*	38	East Africa
Bovidae	Bovinae	Tragelaphini	*Tragelaphus eurycerus*	34F/33M	African rainforests
Bovidae	Bovinae	Tragelaphini	*Taurotragus oryx*	32F/31M	South and East Africa
Bovidae	Bovinae	Boselaphini	*Boselaphus tragocamelus*	46	South Asia
Bovidae	Bovinae	Bovini	*Bubalus bubalis*	50	South, East and Southeast Asia
Bovidae	Bovinae	Bovini	*Syncerus caffer*	54–56	Sub-Saharan Africa
Bovidae	Bovinae	Bovini	*Bos taurus*	60	Cosmopolitan
Bovidae	Bovinae	Bovini	*Bos javanicus*	60	Southeast Asia
Bovidae	Bovinae	Bovini	*Bos gaurus*	58	South and Southeast Asia
Bovidae	Bovinae	Bovini	*Bos grunniens*	60	Himalayas
Bovidae	Bovinae	Bovini	*Bison bonasus*	60	Europe
Bovidae	Caprinae	Caprini	*Pseudois nayaur*	56	Himalayas
Bovidae	Caprinae	Caprini	*Ovis orientalis*	54	Eurasia
Cervidae	Cervinae	Muntiacini	*Muntiacus muntjak*	6F/7M	Southeast Asia
Cervidae	Cervinae	Cervini	*Cervus elaphus*	68	Northern hemisphere

Chromosome data is sourced from the Atlas of Mammalian Chromosomes (O'Brien, 2006). Geographical ranges based on Global Biodiversity Information Facility records (https://gbif.org) that were filtered for quality.