Naturally occurring T-cell-mediated antitumor immune response to LINE-1 antigens in mice that can be enhanced by vaccination and TLR5 agonist

Katerina I. Leonova, Anatoli S. Gleiberman, Victor Jaravine, Nicholas H. Trageser,

Ksenia A. Novototskaya-Vlasova, Marina P. Antoch, Oleg V. Kurnasov, Andrei L. Osterman,

Ekaterina L. Andrianova, and Andrei V. Gudkov

Supplementary Material

Supplementary Material and Methods

Putative T-cell epitopes in L1 proteins

The netMHC4.0 software is utilized to predict the binding affinities of peptide epitopes to MHC-I molecules. The predicted scores include various parameters such as MHC-I binding affinity (Aff), binding stability (Tstab), cleavage score (Acl), TAP transport score (ATAP), combined score ($Rank_{comb}$), and proteomics abundance (Abundance). The software sets default thresholds to designate the strength of epitopes. Epitopes are categorized as 'weak' if their IC50 affinity falls between 500nM and 50nM, while those with affinities below 50nM are categorized as 'strong'. These correspond to the top 2% by affinity rank-score and top 0.5% by rank-score, respectively. The prediction is performed for various MHC restrictions and three different peptide lengths (8, 9, and 10 amino acids). In the case of the mouse strain Balb/cJ, three MHC types are considered: H-2-Dd, H-2-Kd, and H-2-Ld. To convert the computed affinity scores to a normalized range between 0 and 1, a formula is used: A1 = 1 - log50k(Aff). This formula scales the affinity scores appropriately for further analysis.

The stability assessments for the identified epitopes were carried out using the netMHCstabpan1.0 software. In this context, the predicted stability of the binding between the epitopes and MHC-I molecules is measured in hours. The $Rank_{Comb}$ score is computed by the netCTL1.1 program, which combines several factors including MHC-I affinity (determined using netMHC3.0), TAP score, and cleavage rank (RankCL). This combined score offers a comprehensive ranking of the epitopes. The Abundance score is represented in parts per million (ppm) and reflects the summation of proteomics abundances associated with the corresponding proteins across all tissues. This data is sourced from the PaxDB database, specifically the file named "10090-WHOLE_ORGANISM-integrated.txt." Notably, the Abundance scores for all the predicted epitopes in this study were quite close to zero. This indicates an extremely low potential for cross-reactivity with the organism's self-peptidome. The collection of computed predicted peptides constitutes a dataset classified as a "short peptide" dataset[1,2].

A developed in-house algorithm for determination of “long-peptides” for each protein sequence performed the following: multiple alignments of the “short-peptide” sequences with the respective full protein sequences, and selection of continuous ‘long’ sequence stretches that contained maximum number of overlapping “short-peptides”. The algorithm has the following parameters: minimal long-peptide length (9 amino acids), a minimum of 2 ‘short’ peptides contained in ‘long’ segments, the thresholds for stability 0.2 hours and for converted affinity A1 scores of 0.25 for “short-peptides”. All “short-peptides” matched the resulting “long-segments” (or in other words, peptides “contained” in the selected longer segments).

Supplementary Table S1 and Supplementary Fig. S1. show the results of computational mapping of putative T-cell epitopes within L1-encoded ORF1p and ORF2p proteins, where T_{stab} values for the long-segments are calculated as averages of the values for constituting short-epitopes. This approach, compared to conventional approach, has circa 3 times more of potential epitopes in the final peptide vaccine construct, and, consequently, produces more immunogenic vaccines, while keeping the total peptide construct size within reasonable for practical applications limits.

Optimized string-of-beads sequence containing ORF2-derived T cell epitopes

Long predicted T cell epitope sequences were joined into string-of-beads sequence using previously described[3] OptiVac tool (https://github.com/FRED-2/OptiVac; www.epitoolkit.de). The resulting sequence is shown in Supplementary Fig. S1.

Engineering of constructs for expression of L1 antigens in mammalian cells

Three constructs were engineered in pLVX-EF1α-IRES-Puro vector (Clontech) and verified by sequencing as described below.

pLVX-EGFP-FLAG encoding a translational fusion of eGFP with C-terminal FLAG peptide in pLVX-EF1α-IRES-Puro vector (Clontech). The DNA coding sequence of eGFP (720 bp) was PCR-amplified from vector pINDUCER21 (https://www.addgene.org/browse/sequence/66698/) using the primers shown below: Forward primer introducing EcoRI cloning site and Kozak sequence (34 nt; Tm~66oC): 5’-gacgtcgaatt CCACCATGGTGAGCAAGGGCGAG -3’; Reverse primer introducing XbaI cloning site and FLAG-tag (70 nt; Tm

~ 65oC): 5’-ctagactctagattacttgtcgtcgtccttgtagtcCTTGTACAGCTCGTCCATGCCGAGAGTGATC-3’. The resulting DNA sequence after cloning into pLVX-EF1α-IRES-Puro vector was verified by Sanger sequencing:

ctcaagcctcagacagtggttcaaagttttttttcttccatttcaggtgtcgtgaggatctatttccggtgaattccaccATGGTGAGCAAGGGCGAGG
AGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCG
TGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCA
AGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACC
CCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAA
CCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAA
CTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCC
GCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACG
GCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGA
AGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGGACTACAAGGACGACGACGACAAGTAAtctagagcggccgcggatcccgcccctctccctccccccccccctaacgttactggc
cgaagccg

Encoded amino acid sequence of EGFP-FLAG protein (247 aa; 28 kDa):

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSR
YPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNS
HNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLL
EFVTAAGITLGMDELYKDYKDDDDK

pLVX-EGFP-smORF1 encoding a translational fusion of EGFP with mouse L1 ORF1 protein and C-terminal FLAG peptide in pLVX-EF1α-IRES-Puro vector. The DNA coding sequence of eGFP (720 bp) was PCR-amplified from vector pINDUCER21 (https://www.addgene.org/browse/sequence/66698/) using the primers shown below: Forward primer introducing EcoRI cloning site and Kozak sequence (34 nt; Tm~66oC): 5’-gacgtcgaattCCACCATGGTGAGCAAGGGCGAG -3’ Reverse primer introducing SpeI cloning site and GGGS linker (58 nt; Tm ~ 65oC): 5’-agatctactagtgctgcctccacctccCTTGTACAGCTCGTCCATGCCGAGAGTGATC-3’. The amplified fragment was digested with EcoRI and SpeI and ligated into pLVX-EF1α-IRES-Puro vector (Clontech) digested by EcoRI and SpeI to obtain the intermediate construct pLVX-EGFP. The DNA coding sequence of mouse ORF1 (1116 bp, as in published ORFeus sequence) was synthesized by GenScript and PCR amplified using the following primers: Forward primer introducing SpeI cloning site (30 nt; Tm ~ 66oC): 5’-gacgtcactagtGCCAAGGGCAAGCGCCGC-3’ Reverse primer introducing BamHI cloning site and FLAG-tag (60 nt; Tm ~ 67oC) 5’-agatctggatccttacttgtcgtcgtccttgtagtcGCGGCGGGTCTTCTCCAGGGC-3’. The amplified fragment was digested by SpeI and BamHI, and ligated into the intermediate construct pLVX-EGFP (described above) digested by SpeI and BamHI.

The resulting DNA sequence in pLVX-EF1α-IRES-Puro vector verified by Sanger sequencing:

ctcaagcctcagacagtggttcaaagttttttttcttccatttcaggtgtcgtgaggatctatttccggtgaattccaccATGGTGAGCAAGGGCGAGG
AGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCG
TGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCA
AGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACC
CCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAA
CCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAA
CTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCC
GCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACG
GCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGA
AGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGGGAGGTGGAGGCAGCACTAGTGCCAAGGGCAAGCGCCGCAACCTGACCAACCGCAACCAGGAC
CACAGCCCCAGCCCCGAGCCCAGCACCCCCACCAGCCCCAGCCCCGGCAACCCCAACACCCCCGAGAA
CCTGGACCTGGACCTGAAGGCCTACCTGATGATGATGGTGGAGGACATCAAGAAGGACTTCAACAAGAGC
CTGAAGGAGATCCAGGAGAACACCGCCAAGGAGCTGCAGGTGCTGAAGGAGAAGCAGGAGAACACCATC
AAGCAGGTGGAGGTGCTGACCGAGAAGGAGGAGAAGACCTACAAGCAGGTGATGGAGATGAACAAGACC

ATCCTGGACCTGAAGCGCGAGGTCGACACCATCAAGAAGACCCAGAGCGAGGCCACCCTGGAGATCGAG
ACCCTGGGCAAGAAGAGCGGCACCATCGACCTGAGCATCAGCAACCGCATCCAGGAGATGGAGGAGCGC
ATCAGCGGCGCCGAGGACAGCATCGAGAACATCGGCACCACCATCAAGGAGAACGGCAAGTGCAAGAAG
ATCCTGACCCAGAACATCCAGGAGATCCAGGACACCATCCGCCGCCCCAACGTGCGCATCATCGGCGTG
GACGAGAACGAGGACTTCCAGCTGAAGGGCCCCGCCAACATCTTCAACAAGATCATCGAGGAGAACTTCC
CCAACCTGAAGAACGAGATGCACATGAACATCCAGGAGGCCTACCGCACCCCCAACCGCCTGGACCAGA
AGCGCAACAGCTCTAGACACATCATCATCCGCACCAGCAACGCCCTGAACAAGGACCGCATCCTGAAGGC
CGTGCGCGAGAAGGGCCAGGTGACCTACAAGGGCAAGCCCATCCGCATCACCCCCGACTTCAGCCCCGA
GACCATGAAGGCCCGCCGCGCCTGGACCGACGTGATCCAGACCCTGCGCGAGCACAAGCTGCAGCCCC
GCCTGCTGTACCCCGCCAAGCTGAGCATCATCATCGAGGGCGAGACCAAGGTGTTCCACGACAAGACCA
AGTTCACCCACTACCTGAGCACCAACCCCGCCCTGCAGCGCATCATCACCGAGAAGAACCAGTACAAGAA
CGGCAACAACGCCCTGGAGAAGACCCGCCGCGACTACAAGGACGACGACGACAAGTAAggatccgcggccgcg
gatcccgcccctctccctcccccccccctaacgttactggccgaagccgcttggaataaggccggtgtgcgtttgtctatatgttattttccaccatattgccgtctt
tttggc

Encoded amino acid sequence of eGFP-mORF1-FLAG protein (624 aa; 71.3 kDa):

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSR
YPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNS
HNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLL
EFVTAAGITLGMDELYKGGGGSTSAKGKRRNLTNRNQDHSPSPEPSTPTSPSPGNPNTPENLDLDLKAYLMM
MVEDIKKDFNKSLKEIQENTAKELQVLKEKQENTIKQVEVLTEKEEKTYKQVMEMNKTILDLKREVDTIKKTQSE
ATLEIETLGKKSGTIDLSISNRIQEMEERISGAEDSIENIGTTIKENGKCKKILTQNIQEIQDTIRRPNVRIIGVDENE
DFQLKGPANIFNKIIEENFPNLKNEMHMNIQEAYRTPNRLDQKRNSSRHIIIRTSNALNKDRILKAVREKGQVTYK
GKPIRITPDFSPETMKARRAWTDVIQTLREHKLQPRLLYPAKLSIIIEGETKVFHDKTKFTHYLSTNPALQRIITEKN
QYKNGNNALEKTRRDYKDDDDK

pLVX-SBV15-EGFP encoding a translational fusion of eGFP with 15 mouse L1 ORF2-derived immunogenic peptides in pLVX-EF1α-IRES-Puro vector (Clontech)

The DNA sequence encoding 15 immunogenic peptides derived from L1 ORF2 connected by linkers optimized for proteasome cleavage was synthesized by Life Technologies and PCR-amplified using the primers shown below: Forward primer introducing EcoRI cloning site and Kozak sequence (36 nt; Tm~66oC): 5'-gacgtcgaattcGCCACCATGCAGCCTCTGTGGAAG-3' Reverse primer introducing SpeI cloning and GGGS linker (53 nt; Tm ~ 66oC): 5'-ctagacactagtgctgcctccacctccCATTCGCACAGGTGTCAGGTGGAACC-3'. The amplified fragment was digested with EcoRI and SpeI and ligated into pLVX-EF1α-IRES-Puro vector (Clontech) digested by EcoRI and SpeI to obtain the intermediate construct pLVX-SBV15. The DNA coding sequence of eGFP (720 bp) was PCR-amplified from vector pINDUCER21 (https://www.addgene.org/browse/sequence/66698/) using the primers shown below: Forward primer introducing SpeI cloning site (38 nt; Tm~66oC): 5'-gacgtcgaattCCACCATGGTGAGCAAGGGCGAG -3' Reverse primer introducing XbaI cloning site (43 nt; Tm ~ 65oC): 5'-agatctactagtgctgcctccacctccCTTGTACAGCTCGTCCATGCCGAGAGTGATC-3'. The amplified fragment was digested by SpeI and XbaI and ligated into the intermediate construct pLVX-SBV15 digested by SpeI and XbaI.

The resulting DNA sequence in pLVX-EF1α-IRES-Puro vector verified by Sanger sequencing:

ctcaagcctcagacagtggttcaaagttttttttcttccatttcaggtgtcgtgaggatctatttccggtgaattcgccaccATGCAGCCTCTGTGGAAG
TCTGTGTGGCGGTTCTGTCACCACCACCACGCCAAGAGAAGCAGAAGGCAAGAGATCATCCACCACCATC
ACATGAAGTTTCTGGCCAAGTGGATGGACCTGCACCACCATGCCGGCAGCAACAACTACTTCAGCCTGAT
CCATCACCACGCTAAGTCCATCAACGTGATCCACTACATCCACCATCATGCCACGCCTTTCTCTATCGCCAC
CAACAACATCCATCACGCCCACTGCAGCCCCTACCTGTTCAACATCGTGCTGCACCACTGTGCCCCTCAC
GGCACCTTCAGCAAGATCCACCATGCTCTGCCCAAGGCCATCTACAGATTCAACGCTATCCCCATCAAGAT
CCCCACGCAGTTCTTCAACGAGCTGCATCACGCCGCTCCTAACACCAGAGCCGCTACCTTTCACCATCAC
GCTTACCCCAAGACCAAGGGCTACACCTTCTTTCATCACCACTACAAGCTGTACTACCGGGCCATCGTGAT
TCACCACCATTGCAGCTTCTACGAGGCCACAATCACCCTGCACCATCACTGTCAGAAGATGTGGTACATCT
ACACCATGCATCACCATTGCCGGTTCCACCTGACACCTGTGCGAATGGGAGGTGGAGGCAGCACTAGTGT
GAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACG

GCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCA
TCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGT
GCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGT
CCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGG
CGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCA
CAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGG
TGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACA
CCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCA
AAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCG
GCATGGACGAGCTGTACAAGTAAtctagagcggccgcggatcccgcccctctccctcccccccccctaacgttactggccgaagccgcttgg
aataaggccggtgtgcgtttgtctatatgttatttccaccatattgccgtcttttggc

Encoded amino acid sequence of the protein containing 15 immunogenic peptides derived from L1 ORF2 fused with eGFP (455 aa; 52.7 kDa):

MQPLWKSVWRFCHHHHAKRSRRQEIIHHHHMKFLAKWMDLHHHAGSNNYFSLIHHHAKSINVIHYIHHHATPF
SIATNNIHHAHCSPYLFNIVLHHCAPHGTFSKIHHALPKAIYRFNAIPIKIPTQFFNELHHAAPNTRAATFHHHAYP
KTKGYTFFHHHYKLYYRAIVIHHHCSFYEATITLHHHCQKMWYIYTMHHHCRFHLTPVRMGGGGSTSVSKGEE
LFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMA
DKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAG
ITLGMDELYK

Lentiviral transduction and stable cell line generation

EMT6 and 4T1 cell lines were purchased from American Type Culture Collection (ATCC). All cell lines were analyzed for mycoplasma contamination using MycoAlert Mycoplasma Detection Kit (Lonza). Cells were maintained in RPMI supplemented with 10% fetal bovine serum, and antibiotic/antimycotic 10,000 units/ml. These cell lines were infected with EGFP (empty vector), mORF1p-eGFP or mORF2$^{15TCE\text{-}eGFP}$ lentivirus, and stable cell lines were generated by puromycin selection.

Colony formation and cell proliferation assays

Parental and modified 4T1 cells (1000/plate) were seeded in triplicate 10cm plates. Cells were maintained in culture for 10 days and were fixed and stained in methylene blue. The number of colonies (>50 cells) were counted manually.

For cell proliferation, parental and modified 4T1 cells (1000/well) were seeded in triplicate 6-wells. Cells were manually counted on day 1, 2, and 6. Statistical difference between groups was analyzed using one-way ANOVA; $p<0.05$ was considered significant.

Western immunoblotting

Proteins were extracted from cells using RIPA lysis buffer (Sigma-Aldrich). A total of 30 mg of total protein was loaded for electrophoresis into 4–12% Bis-Tris protein gels (Invitrogen). Proteins were transferred to a polyvinylidene difluoride membrane (Bio-Rad). Membranes were blocked with 5% milk in tris-buffered saline (TBS) with 0.1% Tween 20. Protein levels were detected using antibodies against GFP (C2) (Santa Cruz, sc-390394), and actin (BD Biosciences), followed by secondary horseradish-peroxidase conjugated antibodies. Bands were visualized using enhanced chemiluminescence (ECL) Plus western blotting detection reagents (Thermo Scientific).

Immunofluorescent staining

4T1 and EMT6 cells grown on Easy-Grip Falcon tissue culture 35mm dishes were fixed 15 min in 4% paraformaldehyde in PBS, washed with PBS, permeabilized and blocked with blocking solution (5% normal donkey serum, 0.2% triton x-100, PBS or TBS, 1% glycine) 15min. Cells were stained 30 min with cocktail of mouse monoclonal antibodies against LINE-1 ORF1 (1:400, MABS1152, Millipore) and chicken anti-GFP (Aves Labs, GFP-1020, 1:1000) in blocking solution, washed with PBS and stained 30 min with secondary donkey anti-

mouse antibody against whole molecule IgG conjugated with Cy3 and donkey anti-chicken antibody conjugated with Alexa488 (both from JacksonImmunoResearch, 2mg/ml concentration in blocking solution), washed with PBS and mounted with ProLong Gold antifade reagent with DAPI.

Tumor samples were fixed in 4%formaldehyde/PBS overnight at +4°C, washed with PBS, cryopreserved by incubation with 15-30% sucrose in PBS overnight at +4°C, embedded in freezing medium Neg-50. 12mm cryosections were stained 30 min at room temperature with cocktail of rabbit monoclonal antibodies against CD3[SP7], 1:200 dilution (Abcam, cat.No.16669) and rat monoclonal anti-mouse antibodies B220, 1:100 dilution (Biolegend, cat.No.103225) in blocking solution. After washing with PBS, cells were stained with cocktail of secondary donkey anti-rat and anti-rabbit IgG antibodies conjugated with Cy3 and Alexa488 correspondingly (both from JacksonImmunoResearch, 2mg/ml concentration in blocking solution), washed with PBS and mounted with ProLong Gold antifade reagent with DAPI.

Images were collected with Zeiss AxioImager 2 fluorescent microscope equipped with AxioCam 702 digital camera using ZEN software 2.6 version.

Subcutaneous tumor growth

Female Balb/cJ, SCID or RAG2 mice (8-10 weeks old; 104-week-old) were randomly assigned into groups. Mice were inoculated with 1×10^6 cells per mouse with either modified 4T1 or EMT6 (EGFP, ORF2^{15TCE}-EGFP, ORF1p-EGFP) subcutaneously into the vicinity of the mammary fat pad in suspension of 100µl. The growth of subcutaneous tumors was measured using a caliper, and the tumor volume was calculated using the following formula: tumor volume = length × (width^2)/2. Mice were euthanized for the following ethical reasons: tumor ulceration, tumor size reaching 2000 mm^3, or weight loss of greater than 20%. Complete tumor regression was defined as the absence of a measurable tumor on three consecutive measurements.

Experimental metastasis model

In naïve mice, for metastasis model, 4T1 cells with EGFP (empty vector), ORF1p-EGFP, ORF2^{15TCE}-EGFP were injected in 8-week-old female Balb/c mice via tail vein ($3x10^4$ per mouse). Mice were sacrificed two weeks later and metastases were enumerated by India ink staining. Briefly, ribs of the mice were cut to expose the lungs and trachea. Next 15% of India ink was slowly injected into the lungs (intratracheal injection) until the lungs inflated. The lungs were then excised, washed in PBS and fixes in Fekete's solution overnight. The tumor nodules do not absorb India ink and appear white in color were counted.

In the second set of experiments, we used C57BL/6 mice which received three doses of vaccine with two weeks in between. Three weeks after the last booster, these mice were challenged with B16-F10 cells ($1x10^5$ per mouse) injected via tail vein. Four weeks later, all mice were sacrificed and nodules formed in the lungs were counted.

Antigen specific IFN-ɣ production in naïve splenocytes

Spleens were aseptically harvested from naïve mice and transferred to six-well culture plates containing RPMI 1640 medium supplemented with 10% fetal bovine serum and penicillin (10 U/ml)-streptomycin (10 µg/ml). The spleens were minced and passed twice over a sterile 60µm mesh, and the erythrocytes were lysed. The remaining cell pellet was resuspended in medium and represented the total splenic mononuclear cell population. Total cell count was performed using hematocytometer and viability was determined by trypan blue exclusion. Freshly isolated splenocytes were plated at $1x10^6$ cells per well in 48-well format. Splenocytes were stimulated with 1µg/ml of synthetic peptide or PMA/Ionomycin (20ng/ml and 1ug/ml, respectively) for 72 hours.

IFN-ɣ levels in conditioned media were determined by enzyme-linked immunosorbent assay (ELISA) in accordance with the protocol of the manufacturer (R&D Systems, DY485).

Vaccination of MMTV-neu mice

FVB/N-TgN (MMTV-neu) 202Mul female mice[4] were vaccinated at 8-weeks of age (tumor free) or when spontaneously developed mammary tumors reached 50-75mm^3. Vaccination groups were as followed: PBS (vehicle), Alum-bound Entolimod (1µg per mouse), 0.25mg/ml of each of the 15 peptides were pooled together (TCV^{ORF2}) and 15 peptides pooled together with entolimod (TCV^{ORF2}-Entolimod). Tumor incidence in MMTV-

HER2/Neu mice was assessed via inspection of mammary pads for palpable masses. Mice were monitored every other day for new tumor formation. Masses were measured using calipers along the two perpendicular diameters. Tumor volume was calculated by: V(volume) = L(length) × W(width)2/2. Data are expressed as ± SEM and statistical analysis was performed using Mantel-Cox test.

Physiological Frailty Index (PFI)

To determine PFI (see references 32 & 33), male Balb/c mice were immunized with vehicle (n=18), entolimod (n=10) and TCV^{ORF2}-entolimod (n=14) at 12, 16 and 20 weeks. Mice were bled and CBCs were measured at 63 weeks of age. Physiological Frailty Index (PFI) was calculated based on the 11 CBCs parameters.

Blood samples were collected from a single submandibular vein bleed from an individual mouse at different ages. 100ml of blood was collected into EDTA-treated Vacutainer tubes (BD) and used for whole blood cell counts. Another 100µl of blood was collected into LiCl-treated plasma separator tubes; plasma was purified by centrifugation at 5000g for 5 min and used for blood biochemistry analysis. Whole blood cell analysis was performed in 20 µl of blood using Hemavet 950 Analyzer (Drew Scientific). The following parameters were measured: while blood cell counts (WBC), neutrophil (NE), lymphocyte (LY), monocyte (MO) and eosinophil (EO) counts and percentage, red blood cell (RBC) counts, hemoglobin (Hb), hematocrit (HCT), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), red cell distribution width (RDW), platelet (PLT) counts and mean platelet volume (MPV).

To quantitate the accumulation of health deficits PFI was created for each individual mouse as previously described. Total blood cell composition (white and red blood cell counts and differentials, nine total) were used for analysis. Data are expressed as mean ± SEM. Statistical analysis was performed using one or two-way ANOVA with Turkey post hoc test or t-test where appropriate. P-values <0.05 were considered significant.

References

1 Jaravine V, Raffegerst S, Schendel DJ, Frishman D. Assessment of cancer and virus antigens for cross-reactivity in human tissues. *Bioinformatics* 2017; **33**: 104–111.

2 Jaravine V, Mösch A, Raffegerst S, Schendel DJ, Frishman D. Expitope 2.0: A tool to assess immunotherapeutic antigens for their potential cross-reactivity against naturally expressed proteins in human tissues. *BMC Cancer* 2017; **17**. doi:10.1186/s12885-017-3854-8.

3 Schubert B, Kohlbacher O. Designing string-of-beads vaccines with optimal spacers. *Genome Med* 2016; **8**. doi:10.1186/s13073-016-0263-6.

4 Usary J, Darr DB, Pfefferle AD, Perou CM. Overview of genetically engineered mouse models of distinct breast cancer subtypes. *Curr Protoc Pharmacol* 2016; **2016**: 14.38.1-14.38.11.

Supplementary Figures

Mouse L1 ORF1 sequence | 371 aa

MAKGKRRNLTNRNQDHSPSPEPSTPTSPSPGNPNTPENLDLDLKAYLMMMVEGIKKDFNKSLKEI
QENTAKELQVLKEKQENTIKQVEVLTEKEEKTYKQVMEMNKTILDLKREVDTIKKTQSEATLEIETL
GKKSGTIDLSISNRIQEMEERISGAEHSIENIGTTIKENGKCKKILTQNIQEIQDTIRRPNVRIIGVDEN
EDFQLKGPANIFNKIIEENFPNLKNEMHMNIQEAYRTPNRLDQKRNSSRHIIIRTSNALNKDRILKAV
REKGQVTYKGKPIRITPDFSPETMKARRAWTDVIQTLREHKLQPRLLYPAKLSIIIEGETKVFHDKTK
FTHYLSTNPALQRIITEKNQYKNGNNALEKTRR

Mouse L1 ORF2 sequence | 1281aa

MPPLTTKITGSNNYFSLISLNINGLNSPIKRHRLTNWLHKQDPTFCCLQETHLREKDRHYLRMKG
WKTIFQANGMKKQAGVAILISDKIDFQPKVIKKDKEGHFILIKGKILQEELSILNIYAPNTRAATFTKE
TLVKLKAHIAPHTIIVGDFNTPLSPMDRSWKQKLNRDTLKLTEVMKQMDLTDIYRTFYPKTKGYTF
FSAPHGTFSKIDHIIGHKSGLNRLKNIEIVPCILSDHHALRLIFNNKINNRKPTFTWKLNNTLLNDTLV
KEGIKKEIKDFLEFNENEATTYPNLWDTMKAFLRGKLIAMSAFKKKRERAHTSSLTTHLKALEKKE
ANSPKRSRRQEIIKLRGEINQVETRRTIQRINQTRSWFFEKINKIDKPLARLTKGHRDKILINKIRNE
KGDITTDPEEIQNTIRSFYKRLYSTKLENLDEMDKFLDRYQVPKLNQDQVDLLNSPISPKEIEAVIN
SLPAKKSPGPDGFSAEFYQTFKEDLTPVLHKLFHKIEVEGILPNSFYEATITLIPKPQKDPTKIENFR
PISLMNIDAKILNKILANRIQEHIKAIIHPDQVGFIPGMQGWFNIRKSINVIHYINKLKDKNHMIISLDA
EKAFDKIQHPFMIKVLERSGIQGQYLNMIKAIYSKPVANIKVNGEKLEAIPLKSGTRQGCPLSPYLF
NIVLEVLARAIRQQKEIKGIQIGKEEVKISLFADDMIVYISDPKNSNRELLNLINSFGEVAGYKINSNK
SMAFLYTKNKQAEKEIRETTPFSIATNNIKYLGVTLTKEVKDLYDKNFKSLKKEIKEDLRRWKDLPC
SWIGRTNIVKMAILPKAIYRFNAIPIKIPTQFFNELEGAICKFIWNNKKPRIAKTLLKDKRTSGGITMP
DLKLYYRAIVIKTAWYWYRDRQVDQWNRIEDPEMNPHTYGHLIFDKGAKTIQWKKDSIFNNWCW
HNWLLSCRRMRIDPYLSPCTKVKSKWIKELHIKPETLKLIEEKVGKSLEDMGTGEKFLNRTAMAC
AVRSRIDKWDLMKLQSFCKAKDTVYKTKRPPTDWERIFTYPKSDRGLISNIYKELKKVDLRKSNN
PLKKWGSELNKEFSPEEYRMAEKHLKKCSTSLIIREMQIKTTLRFHLTPVRMAKIKNSGDSRCWR
GCGERGTLLHCWWDCRLVQPLWKSVWRFLRKLDIVLPEDPAIPLLGIYPEEAPTGKKDTCSTMFI
AALFIIARNWKEPRCPSTEEWIQKMWYIYTMEYYSAIKKNEFMKFLAKWMDLESIILSEVTQSQRN
SHNMYSLISGY

Figure S1. Computational mapping of putative T-cell epitopes within L1-encoded ORF1p and ORF2p proteins. Mouse L1 ORF1 and ORF2 sequences, with highlighted locations of predicted T-cell epitopes.

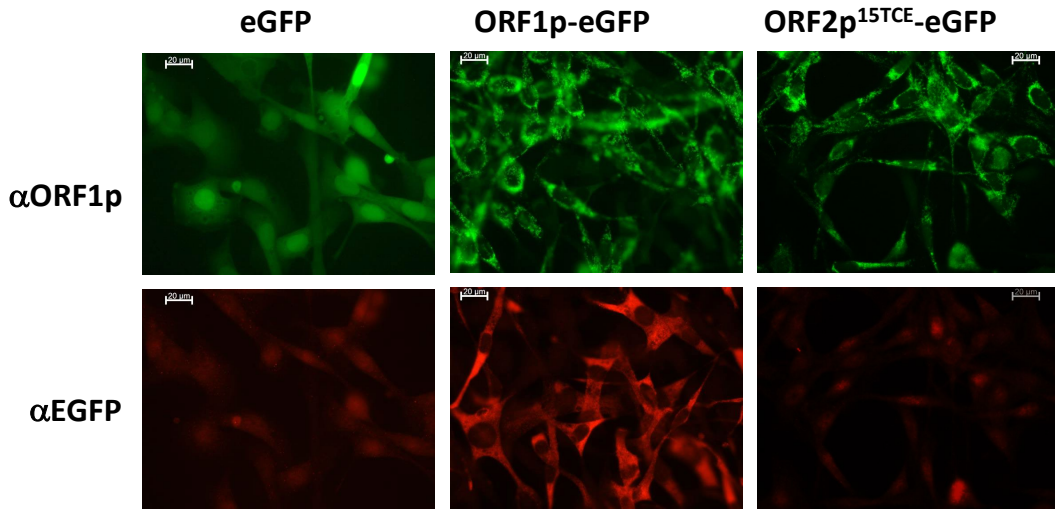

Figure S2. Stable overexpression of ORF1p-EGFP, ORF2p15TCE-EGFP or EGFP in EMT6 cells. Parental EMT6 cells were infected with lentivirus containing L1 proteins. After puromycin selection protein expression was confirmed by immunofluorescent staining.

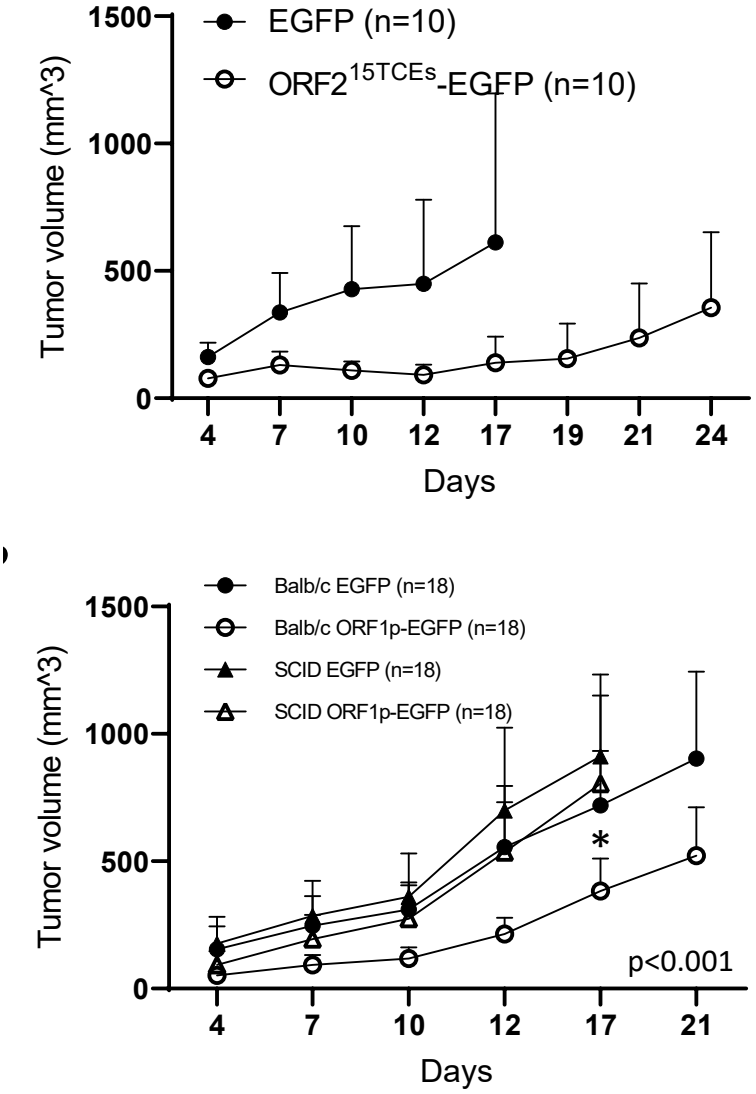

Figure S3. Comparison of subcutaneous tumor growth of 4T1 cells. (a) 4T1-eGFP or 4T1-ORF2^{15TCE}-eGFP were grown subcutaneously in naïve Balb/c mice. (b) 4T1-eGFP or 4T1-ORF1p-eGFP were grown subcutaneously in immunodeficient (SCID) or immunocompetent (Balb/c) mice.

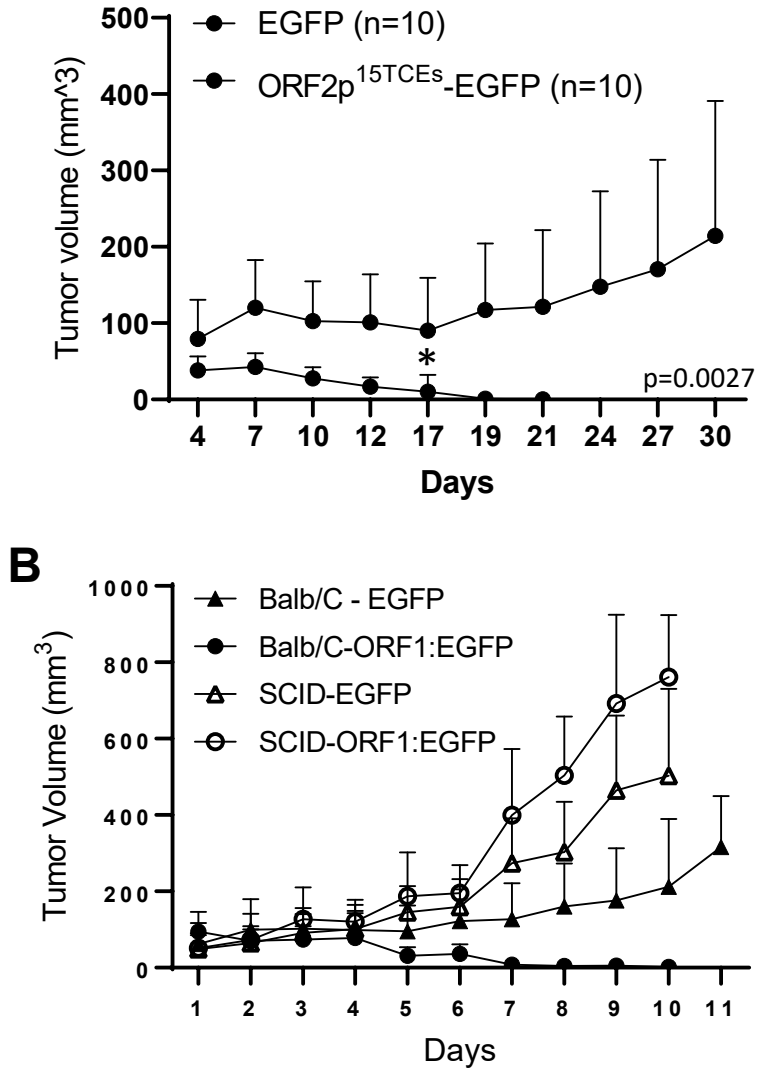

Figure S4. Comparison of subcutaneous tumor growth of EMT6 cells. (a) EMT6-eGFP or EMT6-ORF2^{15TCE}-eGFP were grown subcutaneously in naïve Balb/c mice. (b) EMT6-eGFP or EMT6-ORF1p-eGFP were grown subcutaneously in immunodeficient (SCID) or immunocompetent (Balb/c) mice. .

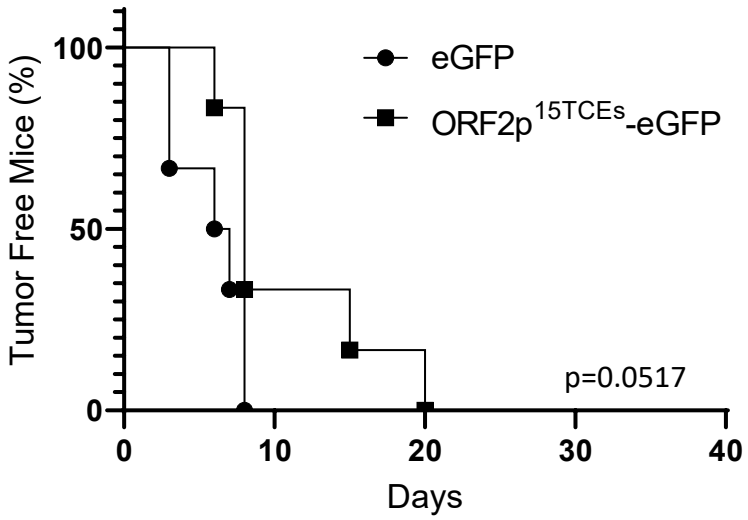

Figure S5. Subcutaneous tumor growth of modified 4T1 cells in old (104-week) Balb/c mice. 4T1-EGFP (n=4) or 4T1-ORF2p^{15TCE}-EGFP (n=5) cells (1x10^6 cells per mouse) were injected subcutaneously in the vicinity of mammary fat pad. Tumor growth was monitored every other day and tumor free survival was calculated.

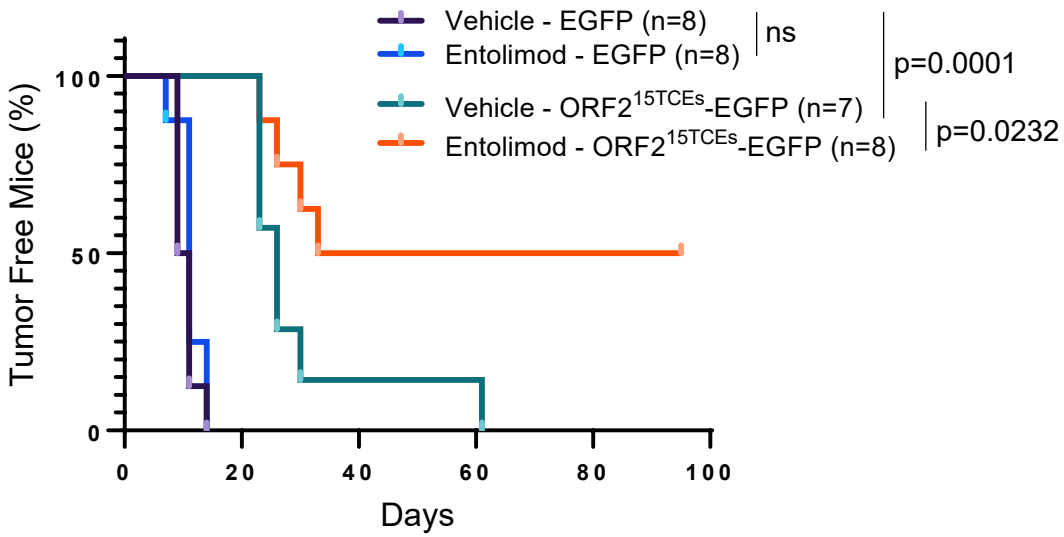

Figure S6. Systemic stimulation of TLR5 by entolimod has no effect on growth of transplanted control (EGFP) 4T1-EGFP cells but significantly reduces tumorigenicity of 4T1 cells expressing L1 antigen (4T1-ORF2p^{15TCE}-eGFP).

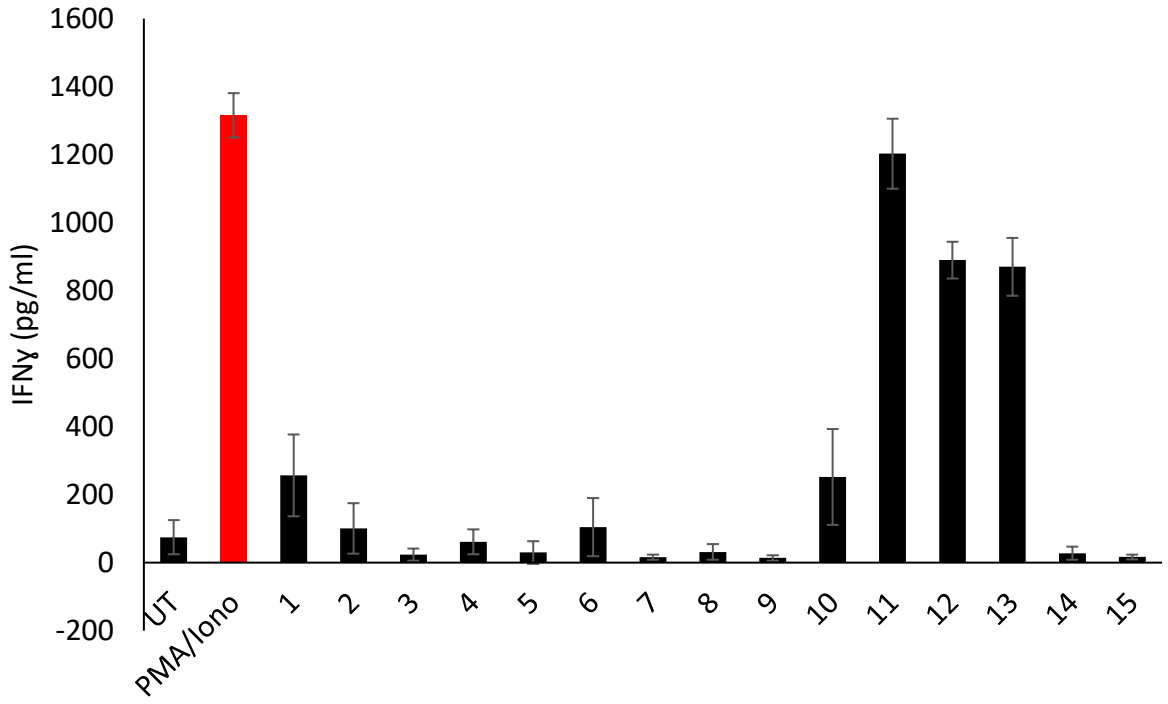

Figure S7. Splenocytes, isolated from naïve C57BL/6 mice, were stimulated with 1µg/ml of each individual peptide or PMA/ionomycin as positive control. Levels of secreted IFNɣ were measured by ELISA 72 hours later.

Supplementary Table

Table S1. Position, length, and predicted stability of MHC binding and immunogenicity index (IG) of putative T-cell epitopes in L1-encoded proteins.

	start	end	length	Tstab (h)	IG	sequence
ORF1p	316	326	11	1.3	2	LLYPAKLSIII
	341	350	10	0.8	2.1	HYLSTNPALQ
	280	289	10	0.6	1	GKPIRITPDF
	358	366	9	0.4	1.1	QYKNGNNAL
	208	218	11	0.4	0.8	KGPANIFNKII
	100	108	9	0.4	0.8	QVMEMNKTI
	3	11	9	0.4	0.7	KGKRRNLTN
	238	255	18	0.3	0.7	RTPNRLDQKRNSSRHIII
ORF2p	825	847	23	1.2	3.3	LPKAIYRFNAIPIKIPTQFFNEL
	671	679	9	0.8	2.7	SPYLFNIVL
	192	201	10	0.7	3.4	YPKTKGYTFF
	340	348	9	0.6	2.1	KRSRRQEII
	1248	1257	10	0.6	1.3	MKFLAKWMDL
	1118	1126	9	0.5	1.3	RFHLTPVRM
	585	593	9	0.4	0.9	KSINVIHYI
	883	891	9	0.4	0.7	KLYYRAIVI
	764	773	10	0.3	1.2	TPFSIATNNI
	1228	1236	9	0.3	0.3	QKMWYIYTM
	515	523	9	0.2	0.7	SFYEATITL
	203	211	9	0.2	1.1	APHGTFSKI
	123	131	9	0.2	0.9	APNTRAATF
	1158	1167	10	0.2	1.3	QPLWKSVWRF
	10	18	9	0.2	0.7	GSNNYFSLI