

COMRADES determines in vivo RNA structures and interactions

Omer Ziv^{1,2,12*}, Marta M. Gabryelska^{3,12}, Aaron T. L. Lun⁴, Luca F. R. Gebert⁵, Jessica Sheu-Gruttadauria⁵, Luke W. Meredith⁶, Zhong-Yu Liu^{7,11}, Chun Kit Kwok⁸, Cheng-Feng Qin⁷, Ian J. MacRae⁵, Ian Goodfellow⁶, John C. Marioni^{4,9,10}, Grzegorz Kudla^{10,3*} and Eric A. Miska^{1,2,10*}

The structural flexibility of RNA underlies fundamental biological processes, but there are no methods for exploring the multiple conformations adopted by RNAs in vivo. We developed cross-linking of matched RNAs and deep sequencing (COMRADES) for in-depth RNA conformation capture, and a pipeline for the retrieval of RNA structural ensembles. Using COMRADES, we determined the architecture of the Zika virus RNA genome inside cells, and identified multiple site-specific interactions with human noncoding RNAs.

The conformational flexibility of RNA is essential for splicing, translation, and post-transcriptional regulation. Recently developed methods have used proximity ligation to reveal RNA base-pairing within cells^{1–6}, but because of insufficient probing depths and a lack of appropriate computational algorithms, it has been difficult to assess the in vivo structural dynamics of RNAs. We developed COMRADES, a method that couples in vivo probing of RNA base-pairing with selective RNA capture. We additionally established an algorithm to assess the structural complexity of RNA inside cells (Fig. 1a).

COMRADES uses a cell-permeable azide-modified psoralen derivative (psoralen-triethylene glycol azide) to facilitate the coupling of two effective affinity-capture steps while overcoming the limited cell permeability of biotin-labeled psoralen (Supplementary Fig. 1a). The azide group does not affect psoralen's cross-linking properties (Supplementary Fig. 1a). After in vivo cross-linking, an RNA of interest is selectively captured, allowing nearly 1,000-fold enrichment (Supplementary Fig. 1b,c). The RNA is then fragmented and a copper-free click-chemistry reaction is carried out to link a biotin moiety to in vivo cross-linked regions, which allows a second streptavidin-based affinity selection of cross-linked regions (Supplementary Fig. 1d). Half of the resulting RNA is proximity-ligated to create RNA chimeras, and then the cross-link is reversed, to enable high-throughput sequencing and assessment of the base-pairing (Supplementary Fig. 1e). The other half is used as a control in which reversal of the cross-link precedes the proximity ligation. COMRADES and control samples have essentially identical RNA compositions, which ensures accurate assessment of artificial chimeric reads originating from random ligation or reverse-transcription errors. COMRADES's dual enrichment substantially increases the structure-probing depth of the selected RNA, thus enabling an

unbiased and global view of coexisting conformations. In our study, COMRADES yielded levels of ligated chimeric reads that were fourfold higher than those in the control or in non-cross-linked samples (Fig. 1b). We successfully reported on the known secondary structure of human 18S ribosomal RNA with high sensitivity (Supplementary Fig. 2), and spurious interactions between cytoplasmic and mitochondrial ribosomal RNA subunits occurred at a very low level (Fig. 1c). The robustness of COMRADES is further demonstrated by its high reproducibility (Fig. 1d,e).

RNA viruses use RNA base-pairing to regulate various aspects of their life cycle^{7–12}. However, the full-length architecture of RNA genomes inside the host cell and their interactions with the host transcriptome are largely unknown. We used COMRADES to determine RNA base-pairing along the 10.8-kb-long single-stranded RNA genome of Zika virus (ZIKV) from the *Flavivirus* genus inside human cells. We identified 1.7 million nonredundant chimeric reads corresponding to the structure of the ZIKV genome (Fig. 1d,e, Supplementary Fig. 3). This high probing coverage is valuable for analysis of multiple coexisting conformations. Previous work mainly identified RNA structures in the untranslated regions (UTRs) of flaviviruses, leaving 95% of the genome unexplored^{7–10}. COMRADES identified base-pairing along the entire genome and between the open reading frame (ORF) and the UTRs (Supplementary Fig. 4a). Nearly 80% of the identified interactions spanned a distance of less than 1,000 nt, implying local structure with a certain degree of three-dimensional compaction (Supplementary Fig. 4b). Both short- and long-range interactions were supported by reproducible, well-defined clusters of chimeric reads, ligated in 5'–3' and 3'–5' orientations (Supplementary Data 1), and showed strong evidence of base-pairing when analyzed with the hybrid-min RNA-folding algorithm as compared to a shuffled-chimeras control (Wilcoxon test P value < 0.0001). COMRADES therefore allows deep and comprehensive analysis of RNA base-pairing inside cells.

During replication, the flavivirus genome undergoes a global conformational change mediated by long-distance base-pairing between the 5' and 3' cyclization sequences (CSs)^{7,13}. Additional elements that contribute to genome cyclization are the 'upstream/downstream of AUG' regions (UAR and DAR, respectively)^{8,9}. COMRADES detected extensive and highly specific base-pairing between the known cyclization elements, thus demonstrating

¹Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, UK. ²Department of Genetics, University of Cambridge, Cambridge, UK. ³MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scotland, UK. ⁴Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK. ⁵Department of Integrative Structural and Computational Biology, the Scripps Research Institute, La Jolla, CA, USA. ⁶Department of Pathology, University of Cambridge, Cambridge, UK. ⁷State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing, China. ⁸Department of Chemistry, City University of Hong Kong, Hong Kong, China. ⁹European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge, UK. ¹⁰Wellcome Sanger Institute, Cambridge, UK. ¹¹Present address: School of Medicine (Shenzhen), Sun Yat-sen University, Guangzhou, China. ¹²These authors contributed equally: Omer Ziv and Marta M. Gabryelska. *e-mail: omer.ziv@gurdon.cam.ac.uk; gkudla@gmail.com; eric.miska@gurdon.cam.ac.uk

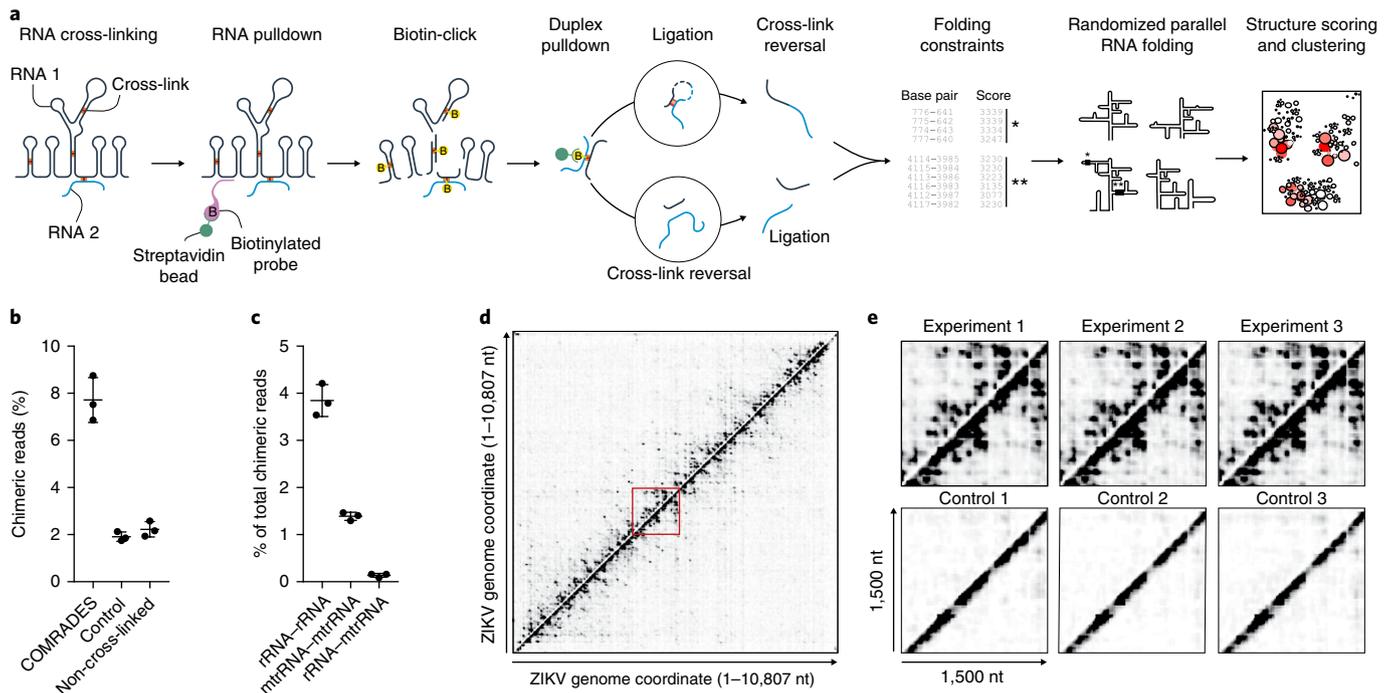


Fig. 1 | COMRADES methodology. **a**, Outline of the COMRADES experimental workflow and associated computational pipeline. **B**, biotin. **b**, The percentage of chimeric reads in COMRADES and control datasets. **c**, Probed interactions among cytoplasmic (rRNA) and mitochondrial (mtrRNA) ribosomal RNA subunits. Mean \pm s.d. of 3 independent experiments shown in **b,c**. **d**, Heat map of ZIKV RNA-RNA interactions. Each dot represents an interaction between the genomic coordinates on the x and y axes. Chimeras ligated in 5'-3' and 3'-5' orientations are plotted above and below the diagonal, respectively. **e**, Zoomed-in views of the 1,500-nt region outlined in red in **d**.

genome cyclization inside cells (Fig. 2a–c, Supplementary Fig. 5a). COMRADES further clarified the nature of the base-pairing associated with genome cyclization by identifying contact regions upstream of the 5' UAR and downstream of the 3' UAR (Fig. 2a,c). We confirmed the existence of previously defined functional RNA pseudoknots including the dumbbell pseudoknot¹⁴, the 'downstream of 5' cyclization sequence' pseudoknot¹⁵, and the stem-loop 1 pseudoknot¹⁶ (Supplementary Fig. 5b). We additionally detected an alternative 5'-UTR conformation in which stem-loops A and B are not formed but rather engage in long-distance base-pairing with the downstream envelope coding sequence (Supplementary Fig. 5c–e). The essential role of stem-loop A during replication suggests that this structure is more likely to be involved in virus translation or packaging. Overall, COMRADES identified nearly all previously known flavivirus RNA structures and further defined critical base-pairing involving the UTRs.

Our intraviral RNA-RNA interaction map revealed the presence of multiple mutually exclusive RNA structures in which one region alternately base-paired with several other regions. The averaged Shannon entropy per nucleotide was 5.9 bits, which implies high folding plasticity (Supplementary Fig. 6a). We found a strong inverse correlation between the degree of experimental support for base-paired regions and their entropy (i.e., strong base-pairing correlated with low entropy; Supplementary Figs. 6a–d and 7). To explore the ensemble of alternative structures, we developed an algorithm to computationally fold \sim 1,000-nt-long regions using randomly selected subsets of high-confidence mutually compatible folding constraints derived from COMRADES data. For each region, a set of 1,000 structures was generated (Supplementary Data 2–11). The validity of this approach is demonstrated by the clear correlation between the thermodynamic stability and the number of reads supporting each structure (Fig. 2d, Supplementary Fig. 8).

Nevertheless, the most thermodynamically favored structures gained only moderate experimental support, which implies additional effects of the cellular environment on RNA folding^{17–19}.

We further computed the degree of similarity between all pairs of structures and applied multidimensional scaling to cluster structures on the basis of their similarity (Supplementary Data 12). The presence of separated well-defined clusters reflected the occurrence of alternative conformations (Fig. 2e, Supplementary Fig. 9). As a control, we randomly shuffled the interacting RNA partners between the chimeric reads; the resulting shuffled structures clustered separately from the structures recovered by COMRADES (Supplementary Fig. 10a,b). A single ZIKV structure typically accounted for \sim 30% of the in vivo observed interactions, whereas a reduced set of five structures was sufficient to capture 80–90% of the in vivo data (Supplementary Fig. 10c). Our analysis suggests that the intracellular folding complexity of the ZIKV genome might be explained by the coexistence of a small set of alternative conformations.

Viral RNAs have an inherent capacity to form specific interactions through base-pairing with host RNAs²⁰, but little is known about the prevalence of such interactions. COMRADES revealed multiple interactions between the ZIKV genome and human small regulatory RNAs (Fig. 3a). We found site-specific interactions between the ZIKV ORF and the U1 small nuclear RNA (Supplementary Fig. 11a) that could plausibly affect host splicing. We also detected site-specific interactions with certain human tRNAs (Supplementary Fig. 11b). We identified several interactions between the ZIKV genome and human microRNAs such as miR-21, miR-19, miR-512, miR-515, and miR-1323, whereas we did not detect any interactions with microRNAs in the control datasets (Fig. 3b, Supplementary Fig. 11c). COMRADES indicated noncanonical base-pairing between the 5' CS of ZIKV and the seed region of miR-21 (Fig. 3c; Benjamini-Hochberg adjusted P value = 1.0×10^{-13}). We further verified the

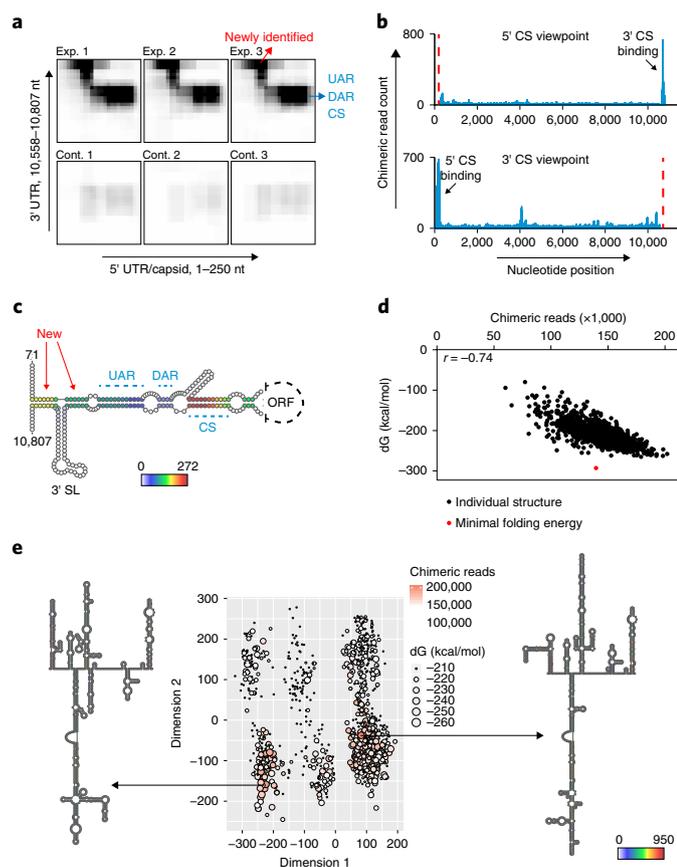


Fig. 2 | The genomic structure of ZIKV inside human cells. **a**, Heat maps of RNA-RNA interactions between cyclization elements. Exp, experiment; cont, control. **b**, Viewpoint histograms showing long-distance (>150 nt) binding positions of the CSs along the ZIKV genome. Viewpoint regions are marked by dashed red lines. **c**, Probed interactions along the circular genome conformation. Color-coding indicates the number of nonredundant chimeric reads supporting each base pair. New, newly identified base-pairing; SL, stem-loop. **d**, Folding energy (dG) and supporting experimental evidence (chimeric reads) for each of the 1,000 computationally predicted structures corresponding to ZIKV genome nucleotides 2,288–3,323. r , Pearson correlation coefficient. **e**, Clustering and prediction of alternative conformations for the region shown in **d**. Color-coding as described in **c**.

significance of the miR-21 interaction with the ZIKV genome by using an independent analysis pipeline (false discovery rate: 3.0×10^{-25} by quasi-likelihood moderated F -test). In vitro-synthesized miR-21 did not bind the ZIKV 5' CS on its own, whereas pre-loading of miR-21 onto purified Argonaute 2 (AGO2) facilitated a strong and sequence-specific interaction (Supplementary Fig. 12), supporting the involvement of AGO2 in this base-pairing. CRISPR-Cas9 deletion of *MIR21* or antisense inhibition of mature miR-21 in human cells decreased the intracellular level of ZIKV genome (Supplementary Fig. 13a–d). Abrogation of the miR-21-binding ability of a ZIKV replicon through point mutations rendered it insensitive to miR-21 antisense inhibition (Supplementary Fig. 13e), which indicates that miR-21 acts through direct interaction with the 5' CS. The ZIKV envelope protein was similarly affected in the *MIR21*-null cells (Supplementary Fig. 14). Although the effect size was relatively small, the strong evidence for the miR-21–5' CS interaction presented here suggests that miR-21 might assume a greater proviral role in the physiological context.

COMRADES revealed the highly dynamic nature of an RNA genome inside cells, as well as its ability to engage in base-pairing with

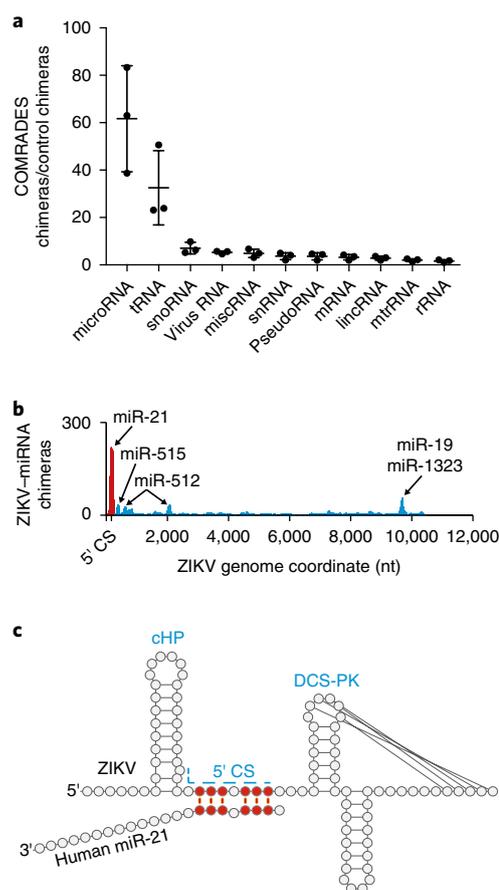


Fig. 3 | Host-virus RNA-RNA interactions. **a**, Human RNA types that interact with the ZIKV genome. snoRNA, small nucleolar RNA; miscRNA, miscellaneous small RNA; snRNA, small nuclear RNA; lincRNA, long intergenic noncoding RNA; mtrRNA, mitochondrial ribosomal RNA; rRNA, cytoplasmic ribosomal RNA. Data are shown as the mean \pm s.d. of 3 independent experiments. **b**, Probed interactions between the ZIKV genome and specific miRNAs in COMRADES samples. **c**, COMRADES determined base-pairing between ZIKV and miR-21 (red). CHP, capsid-coding-region hairpin element; DCS-PK, 'downstream of 5' cyclization sequence' pseudoknot.

multiple host regulatory RNAs. The involvement of the conserved 5' CS of ZIKV in genome cyclization, capsid translation, and miR-21 binding further demonstrates the intracellular structural complexity of viral RNA genomes. The general applicability of COMRADES provides an opportunity to undertake an unbiased analysis of the dynamic nature of RNA inside cells and can be used to investigate the structure and interaction partners of any cellular or foreign RNA in any species.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41592-018-0121-0>.

Received: 22 May 2018; Accepted: 20 July 2018;
Published online: 10 September 2018

References

- Lu, Z. et al. *Cell* **165**, 1267–1279 (2016).
- Aw, J. G. A. et al. *Mol. Cell* **62**, 603–617 (2016).

3. Sharma, E., Sterne-Weiler, T., O'Hanlon, D. & Blencowe, B. J. *Mol. Cell* **62**, 618–626 (2016).
4. Helwak, A., Kudla, G., Dudnakova, T. & Tollervey, D. *Cell* **153**, 654–665 (2013).
5. Ramani, V., Qiu, R. & Shendure, J. *Nat. Biotechnol.* **33**, 980–984 (2015).
6. Sugimoto, Y. et al. *Nature* **519**, 491–494 (2015).
7. Liu, Z.-Y. et al. *eLife* **5**, e17636 (2016).
8. Alvarez, D. E., Lodeiro, M. F., Ludueña, S. J., Pietrasanta, L. I. & Gamarnik, A. V. *J. Virol.* **79**, 6631–6643 (2005).
9. Friebe, P. & Harris, E. *J. Virol.* **84**, 6103–6118 (2010).
10. Filomatori, C. V. et al. *Genes Dev.* **20**, 2238–2249 (2006).
11. Pirakitikulr, N., Kohlway, A., Lindenbach, B. D. & Pyle, A. M. *Mol. Cell* **62**, 111–120 (2016).
12. Watts, J. M. et al. *Nature* **460**, 711–716 (2009).
13. Hahn, C. S. et al. *J. Mol. Biol.* **198**, 33–41 (1987).
14. Manzano, M. et al. *J. Biol. Chem.* **286**, 22521–22534 (2011).
15. Liu, Z.-Y. et al. *J. Virol.* **87**, 6804–6818 (2013).
16. Akiyama, B. M. et al. *Science* **354**, 1148–1152 (2016).
17. Rouskin, S., Zubradt, M., Washietl, S., Kellis, M. & Weissman, J. S. *Nature* **505**, 701–705 (2014).
18. Ding, Y. et al. *Nature* **505**, 696–700 (2014).
19. Spitale, R. C. et al. *Nature* **519**, 486–490 (2015).
20. Guo, Y. E. & Steitz, J. A. *Mol. Cell. Biol.* **34**, 3780–3787 (2014).

Acknowledgements

The authors thank A. Kohl (Centre for Virus Research, University of Glasgow, Glasgow, UK) and L.J. Pena and R.F. França (Fiocruz Recife, Pernambuco, Brazil) for providing the PE243 ZIKV RNA used to generate the virus stock. We thank Y. Galanty and F.M. Martínez for assisting with CRISPR–Cas9 knockout; T.D. Domenico and W. Matsushima for collapsing UMIs; G. Sanguinetti for suggesting the RNA-folding strategy; A.M. Sharkey and L. Gardner (University of Cambridge, Cambridge, UK) for providing the JEG-3 cells; S. Moss for assisting with risk assessment; M.S. Diamond, T. Sweeney, A. Firth, D. Jordan, A. Zeisel, and members of the Miska group for their comments; and C. Flandoli for illustrations. This work was supported by Cancer Research UK (C13474/A18583 and C6946/A14492 to E.A.M.; award no. A17197 to J.C.M.), the Wellcome Trust

(104640/Z/14/Z and 092096/Z/10/Z to E.A.M.; 207507 to G.K. and M.M.G.; Senior Fellowship in Basic Biomedical Science 207498/Z/17/Z to I.G.), the Human Frontier Science Program (HFSP; LT000558/2015 to O.Z.), the European Molecular Biology Organization (EMBO; ALTF1622–2014 to O.Z.), the Blavatnik Family Foundation (postdoctoral fellowship to O.Z.), the UK Medical Research Council (G.K. and M.M.G.), EMBL (core funding to J.C.M.), NIGMS (grants R01GM104475 and R01GM115649 to I.J.M., L.F.R.G., and J.S.-G.), City University of Hong Kong (projects 9610363 and 7200520 to C.K.K.), the Croucher Foundation (project 9500030 to C.K.K.), the Hong Kong RGC (projects 9048103 and 9054020 to C.K.K.), the NSFC (Excellent Young Scientist Fund 81522025 to C.-F.Q.), and the Academy of Medical Sciences, UK (Newton Advanced Fellowship to C.-F.Q.).

Author contributions

O.Z. designed, developed, and performed COMRADES; E.A.M. supervised the study; M.M.G. and G.K. developed the associated analysis pipeline and analyzed coexisting conformations and interactions; A.T.L.L. and J.C.M. developed an independent analysis pipeline and discovered the ZIKV–miR-21 interaction; O.Z. performed the in vivo miR-21 experiments with assistance from L.W.M., I.G., and C.K.K.; L.F.R.G., J.S.-G., and I.J.M. performed the in vitro miR-21-binding experiments; Z.-Y.L. and C.-F.Q. provided the ZIKV replicons under an MTA agreement; and O.Z., G.K., and E.A.M. wrote the paper with input from all other authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41592-018-0121-0>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to O.Z. or G.K. or E.A.M.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

COMRADES. Each independent experiment was carried out on a different day and included three sequencing libraries: COMRADES, control, and a non-cross-linked sample.

Psoralen cross-link. JEG-3 cells (~50 million cells per experiment) were inoculated with ZIKV isolate PE243 at a multiplicity of infection of 2× the 50% tissue culture infective dose (TCID₅₀) per cell. 20 h after inoculation, cells were washed three times in PBS and then incubated for 20 min with 0.4 mg/ml psoralen-triethylene glycol azide (Berry & Associates) dissolved in PBS and diluted in OptiMEM 1 with no phenol red (Gibco). Cells were irradiated on ice with 365-nm UV radiation for 10 min in a CL-1000 cross-linker (UVP). Prolonged UVA irradiation should be avoided because it might decompose the azide moiety. Cells were lysed with RNeasy lysis buffer. Proteins were degraded by proteinase K (NEB), and RNA was purified with the RNeasy mini kit (Qiagen).

Viral RNA enrichment. Total RNA was mixed with an array of 50 biotinylated DNA oligos (IDT), each 20 nt long, designed to capture ZIKV genomic RNA and was maintained at 37 °C for 6 h with rotation in the following hybridization buffer: 500 mM NaCl, 0.7% SDS, 33 mM Tris-Cl, pH 7, 0.7 mM EDTA, 10% formamide. Hybridization and wash conditions were adapted from ref. ²¹. At the end of incubation, MyOne Streptavidin C1 Dynabeads (Invitrogen) were added and the RNA was incubated for an additional hour at 37 °C. Beads were captured on a magnet and were washed five times with 2× SSC buffer containing 0.5% SDS. RNA was released from beads by degradation of the DNA probes with 0.1 units/μl Turbo DNase (Invitrogen) at 37 °C for 30 min. RNA was cleaned with RNA Clean & Concentrator (Zymo Research).

Cross-link pulldown. RNA was fragmented to an average size of 100 nt with RNase III (Ambion) and was cleaned with RNA Clean & Concentrator (Zymo Research). A copper-free click reaction was carried out at 37 °C for 90 min in the presence of 150 μM Click-IT Biotin DIBO alkyne (Life Technologies) and 0.5 units/μl Suprase-In (Invitrogen). The reaction was terminated by RNA Clean & Concentrator (Zymo Research). Biotinylated RNA was pulled down with MyOne Streptavidin C1 Dynabeads (Invitrogen) under the following reaction conditions: 100 mM Tris-Cl, pH 7.5, 10 mM EDTA, 1 M NaCl, 0.1% Tween-20, 0.5 unit/μl Suprase-In. Beads were captured on a magnet and were washed five times with 100 mM Tris-HCl, pH 7.5, 10 mM EDTA, 3.5 M NaCl, 0.1% Tween-20. We eluted RNA by adding 95% formamide, 10 mM EDTA solution and incubating the mixture at 65 °C for 5 min. To avoid the enrichment of small RNA chimeric reads that could not be double-aligned to the reference ZIKV genome/human transcriptome, we size-fractionated RNA on 10% TBE-urea gel and eluted fragments corresponding to a size of 100–200 nt overnight at 4 °C in 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 250 mM NaCl, 0.1% SDS. RNA was concentrated with RNA Clean & Concentrator (Zymo Research).

Proximity ligation and cross-link reversal. At this stage, the RNA sample was divided into two equal samples. In one half we applied proximity ligation followed by cross-link reversal (i.e., COMRADES sample), whereas in the other half we carried out cross-link reversal before proximity ligation (i.e., control sample). We included an additional control containing an equimolar concentration (albeit a non-similar composition) of non-psoralen-treated, non-cross-linked enriched RNA (i.e., non-cross-linked sample). Before proximity ligation, the RNA was heated to 85 °C for 2 min and then cooled down rapidly on ice. Proximity ligation was done under the following conditions: 1 unit/μl RNA ligase 1 (New England Biolabs), 1× RNA ligase buffer, 50 μM ATP, 1 unit/μl Suprase-In (Invitrogen), in a final volume of 200 μl. The reaction was incubated for 16 h at 16 °C and was terminated by cleaning with RNA Clean & Concentrator (Zymo Research). For cross-link reversal we irradiated the RNA on ice with 2.5 kJ/m² UVC.

Sequencing-library preparation. Library preparation was done as described in ref. ²², with the following modifications: 6N unique molecular identifiers were added to the 5' end of the 3' sequencing adaptor; primer and adaptor concentrations were reduced to match the low RNA input; Agencourt RNAClean XP beads (Beckman Coulter) were used for cleanup and size separation; preadenylated 5' and 3' adaptors were used; and all ligation reactions were carried out without ATP to reduce ligation artifacts. All libraries and controls went through 13 PCR cycles with KAPA HiFi HotStart Ready Mix (KAPA Biosystems). PCR products were size-selected on a 1.8% agarose gel before being loaded on a HiSeq 1500 sequencer (Illumina).

Cell culture. JEG-3 placental trophoblasts (ATCC) and HeLa cells (ATCC) were cultured in minimum essential medium supplemented with 10% FBS, 1 mM sodium pyruvate, GlutaMAX, non-essential amino acids, and penicillin-streptomycin. Vero cells (Sigma-Aldrich) were cultured in DMEM supplemented with 10% FBS, GlutaMAX, and penicillin-streptomycin. All cell lines were cultured in a humidified CO₂ incubator at 37 °C and were regularly examined to exclude mycoplasma contamination.

Virus inoculation. ZIKV isolate PE243 originated in Recife, Brazil, in 2015. We propagated the virus in Vero cells, and determined titers by measuring the TCID₅₀ in JEG-3 cells. For measurements of virus replication, JEG-3 or HeLa cells were inoculated with ZIKV at a multiplicity of infection of 0.1 TCID₅₀ per cell for 3 h, after which cells were washed three times with PBS and supplemented with fresh growth medium. 24 h after inoculation, the medium was removed, cells were washed three times with PBS, and RNA was extracted with the RNeasy kit (Qiagen). Virus copy number was determined via TaqMan real-time PCR assay (Primerdesign) and was normalized to GAPDH and ribosomal RNA. All virus work was handled in a containment level 2 facility registered with the UK Health and Safety Executive under the Control of Substances Hazardous to Health Regulations.

Replicon assay. ZIKV wild-type and 5' CS-3' CS double-mutant replicons were described previously²³. Replicon RNA was synthesized in vitro with the MEGAScript T7 transcription kit (Ambion). Replicon RNA was capped with the ScriptCap m7G capping system (CellsScript) and transfected into HeLa cells with the TransIT-mRNA transfection kit (Mirus). Replicon levels were analyzed after 24–48 h with a microplate luminometer (Promega) and normalized to baseline luminescence values measured at 6 h post-transfection.

MIR21 knockout. JEG-3 cells were transfected with CAS9-gRNA riboprotein complexes using Lipofectamine RNAiMAX (Life Technologies) according to the Alt-R CRISPR-Cas9 user guide (IDT). *MIR21*-knockout clone 1 was generated with the guide RNA 5'-TCATGGCAACACCAGTCGATGGG-3' and contained a homozygous deletion at positions 59841310–59841326 on chromosome 17 (GRCh38/hg38 assembly). *MIR21*-knockout clone 2 was generated with a mixture of two guide RNAs (5'-ATGTCAGACAGCCCATCGACTGG-3' and 5'-CTACCATCGTGACATCTCCATGG-3') and contained a homozygous deletion at positions 59841249–59841321 on chromosome 17. *MIR21*-knockout and control clones were validated by Sanger sequencing and by TaqMan advanced miRNA assay targeting mature miR-21 (Life Technologies).

miR-21 inhibition. HeLa cells were transfected with inhibitors targeting human miR-21 or nontargeting control A (Power Inhibitors, Exiqon) at a final concentration of 25 nM using Lipofectamine RNAiMAX. 6 h post-transfection, the medium was replaced and cells were inoculated with ZIKV or retransfected with ZIKV replicons as described above. miR-21 inhibition was validated with a psiCHECK-2 reporter (Promega) carrying a fully complementary miR-21 site at the 3' UTR of a *Renilla* luciferase reporter along with a firefly reporter to normalize transfection efficiency. The miR-21 psiCHECK-2 reporter was deposited in Addgene (plasmid # 114206). Luminescence was assessed by dual-reporter assay (Promega) and normalized to control psiCHECK-2 without the miR-21-binding site.

Gel-based reverse-transcription stalling (RTS) assay. RTS assay was performed as previously described²⁴ with a Cy5-labeled primer targeting the human 5.8S ribosomal RNA: 5'-Cy5-AAGCGACGCTCAGACAGG-3'.

Dot blot analysis. 50 ng of cross-linked RNA, or the indicated amount of 50-nt-long biotinylated standards, was spotted onto a Biodyne B nylon membrane (Life Technologies) and dried by baking at 80 °C for 10 min. Biotinylated RNA was detected with the chemiluminescent nucleic acid detection module kit (Life Technologies) and visualized with the ChemiDoc MP imaging system (Bio-Rad).

Purification of human AGO2 loaded with miR-21. Human AGO2 homogeneously loaded with miR-21 was prepared according to a published protocol²⁵. Human AGO2 was expressed in Sf9 cells via the Bac-to-Bac baculovirus expression system (Thermo Fisher Scientific). Sf9 cells were lysed and human AGO2 was purified by Ni-NTA affinity chromatography using a His tag. Human AGO2 was loaded with synthetic 5'-phosphorylated miR-21 (IDT), and the His tag was removed with tobacco etch virus protease. Human AGO2 loaded with miR-21 was captured with an antisense oligonucleotide (IDT), eluted, and purified by size-exclusion chromatography on an ÄKTA FPLC (GE Healthcare Life Science). Protein concentration was measured on the basis of absorption at 280 nm, using a combined (protein and miRNA guide) extinction coefficient of 192,760 M⁻¹ cm⁻¹.

Target RNA labeling. Synthetic RNA oligonucleotides (IDT) were radiolabeled at the 5' end with γ-³²P ATP (PerkinElmer) and T4 polynucleotide kinase (NEB), and purified by denaturing polyacrylamide gel and ethanol precipitation. RNA concentration was determined on the basis of absorption at 260 nm, using the following extinction coefficients: wild type, 452,800 M⁻¹ cm⁻¹; mutant 1, 442,900 M⁻¹ cm⁻¹; mutant 2, 456,900 M⁻¹ cm⁻¹; perfect seed match, 455,700 M⁻¹ cm⁻¹; no cHP loop, 273,400 M⁻¹ cm⁻¹; wild-type long, 624,900 M⁻¹ cm⁻¹; 8-mer (seed match), 137,900 M⁻¹ cm⁻¹.

Electrophoretic mobility shift assay. Binding reactions were prepared in reaction buffer (28 mM Tris, pH 8.0, 20 mM KCl, 80 mM KOAc, 1.6 mM Mg(OAc)₂, 0.5 mM TCEP, 0.004% NP-40, 0.01 g/l baker's yeast tRNA) with a final volume of 20 μl

and a final concentration of the labeled RNAs of 10 nM and of the nonlabeled RNA or AGO2-miR-21 of 100 nM. Reactions were incubated for 10 min at room temperature and analyzed on a 15% acrylamide native gel in 0.5×TBE.

K_d measurements. Binding experiments were conducted according to the protocol published in ref. ²⁵. AGO2-miR-21 (0–200 nM) was incubated with 0.1 nM radiolabeled target in reaction buffer (28 mM Tris, pH 8.0, 20 mM KCl, 80 mM KOAc, 1.6 mM Mg(OAc)₂, 0.5 mM TCEP, 0.004% NP-40) with a total volume of 25 µl for 45 min at room temperature. Filter-binding was performed with a dot-blot apparatus (GE Healthcare Life Sciences) with Protran nitrocellulose membrane (Amersham, GE Healthcare Life Sciences) and Hybond-N+ nylon membrane (Amersham, GE Healthcare Life Sciences). Samples were applied with vacuum and washed with 50 µl of wash buffer (30 mM Tris, pH 8.0, 100 mM KOAc, 2 mM Mg(OAc)₂, 0.5 mM TCEP). After air-drying, the membrane strips were used to expose phosphor screens (GE Healthcare Life Sciences) for visualization. Screens were imaged on a Typhoon phosphorimager (GE Healthcare Life Science), and signals were quantified with ImageQuant (GE Healthcare Life Sciences). We calculated dissociation constants by fitting the data to a single site binding equation:

$$F = \frac{B_{\max} [\text{AGO2}]}{[\text{AGO2}] + K_d}$$

where F is the fraction of target RNA bound, B_{\max} is the maximum number of binding sites, $[\text{AGO2}]$ is the total concentration of the AGO2-miR-21 complex, and K_d is the calculated dissociation constant, obtained with Prism (GraphPad Software). For weakly binding RNAs, B_{\max} was constrained to ≤ 1.

Data analysis and statistical testing. *Processing and visualization of sequencing data.* Sequencing data were preprocessed to combine FASTQ files of two sequencing lanes (cat) and to remove adaptors (cutadapt). Paired-end reads were merged by paired-end read merger (pear). Unique molecular identifiers (UMIs) were collapsed by collapse.py (T. Di Domenico; <https://github.com/tdido/tstk/blob/master/tstk/collapse.py>). Chimeric reads were called and annotated with the hyb package²⁶, using the following command: `hyb analyse in=data.fasta db=hOH7_and_Zika format=comp eval=0.001`.

Hyb uses bowtie2²⁷ in local mapping mode to map reads to a transcriptome database and to identify chimeras, and it annotates the chimeras with RNA base-pairing information generated by hybrid-min²⁸. The transcriptome database used by hyb, hOH7_and_Zika, consists of human spliced mRNAs and noncoding RNAs described in ref. ⁴, and the genome sequence of ZIKV (Zika virus isolate ZIKV/*Homo sapiens*/Brazil/PE243/2015, complete genome). To evaluate the folding energy of chimeric reads, we used hybrid-min²⁸ with the default settings. We then randomly reassigned (shuffled) pairs of fragments found in chimeric reads, and repeated the folding-energy analysis. The folding energies of experimentally identified and shuffled chimeras were compared by Wilcoxon test.

Virus interaction heat maps were plotted in Java Treeview²⁹, such that color intensity represented the coverage of chimeric reads at every pair of positions. We plotted the first mapped fragment of each chimera along the x axis, and the second fragment along the y axis. As a result, chimeras found in the 5′–3′ orientation were shown above the diagonal, and chimeras in the 3′–5′ orientation were below the diagonal. Viewpoint histograms were plotted with ggnplot, and arc plots were plotted with R-chie³⁰.

For every pair of positions (i, j) along the virus genome, we calculated the COMRADES score C_{ij} , the number of chimeric reads that, when analyzed with the program hybrid-min with default settings, indicated base-pairing between positions i and j . We used COMRADES scores to calculate the per-base Shannon entropy for each nucleotide position along the virus. The Shannon entropy of position i is defined as

$$\text{Entropy}_i = - \sum_{j=1}^n P(C_{ij}) \log_2 P(C_{ij})$$

where n is the length of the genome (10,807 nt) and

$$P(C_{ij}) = \frac{C_{ij}}{\sum_{k=1}^n C_{ik}}$$

High entropy indicates flexible positions that may form multiple alternative base pairs, whereas low entropy indicates positions that always pair with the same nucleotide partner. We visualized RNA structures using VARNA³¹, in which the color scale represented the COMRADES score for each base pair.

RNA structure prediction. For RNA structure predictions, we collected all potential base pairs with a nonzero C_{ij} value, assembled sets of adjacent base pairs into uninterrupted stem structures, and calculated the base-pairing score of each stem as the sum of C_{ij} values of individual base pairs. We then ranked these stem elements by their scores. In a preliminary analysis, we folded the 10,807-nt virus

genome in a set of 50 overlapping 1,000-nt fragments, using the hybrid-ss-min program²⁸. Each fragment was folded using a set of 250 top-ranked in vivo probed stem elements as folding constraints. On the basis of this preliminary analysis, we identified high-scoring stem-loop structures that were reproducibly predicted across multiple fragments, and we defined new fragment boundaries to prevent the disruption of these reproducible structural elements. As a result, we obtained fragments that varied in size but were approximately 1,000 nt long each.

We then performed full folding analysis using the following fragment boundaries: 5′ UTR, 1–107; fragment 1 (F1), 108–1,275; F2, 1,276–2,287; F3, 2,288–3,323; F4, 3,324–4,521; F5, 4,522–5,551; F6, 5,552–6,810; F7, 6,811–7,757; F8, 7,758–8,755; F9, 8,756–9,543; F10, 9,544–10,379; 3′ UTR, 10,380–10,807.

For each fragment, we assembled a set of folding constraints that represented the 75 top-scoring stem elements in that fragment, and we then randomly shuffled this set of constraints 1,000 times and used the shuffled constraints for folding prediction by hybrid-ss-min. The resulting individual structures typically incorporated 25–40% of these constraints. We recorded the folding energy of each structure, as predicted by hybrid-ss-min, and we used the sum of C_{ij} values to calculate an overall score for each structure. To assemble the top-scoring full-genome structure shown in Supplementary Fig. 4a, we assembled the top-scoring structures for each coding sequence fragment (F1–F10) and the previously proposed structures of the 5′ and 3′ UTRs. Additional analyses of folding within and between the 5′ and 3′ UTRs are shown in Fig. 2a–c and Supplementary Fig. 4c.

We also repeated the folding analysis with shuffled sets of 50–250 top-scoring constraints per fragment. This yielded similar results, but we found that either reducing or increasing the numbers of constraints tended to reduce the number of high-scoring structures.

To explore the sets of alternative structures, we computed pairwise distances between structures as the number of positions with discordant base-pairing. This resulted in a 1,000 × 1,000 matrix of distances, which we then represented on a two-dimensional surface via multidimensional scaling (using the R function cmdscale). Multidimensional scaling, also known as principal coordinate analysis³², maps multidimensional objects (in this case, RNA structures) to a set of points on a plane, such that the distances between RNA structures are well approximated by Euclidean distances between points, by minimization of a stress function:

$$\text{Stress}_D(x_1, \dots, x_n) = \left(\sum_{i \neq j=1 \dots n} (D_{ij} - \|x_i - x_j\|)^2 \right)^{1/2}$$

We calculated the statistical significance of host–virus RNA–RNA interactions in DESeq³³ by comparing counts of chimeric reads from three COMRADES and three control datasets.

Discovery of ZIKV miR-21 interaction via an independent analysis pipeline.

Alignment. The first read of each pair was processed with UMI-tools³⁴ to extract the 6-nt UMI at the start of the read. Processed reads were aligned using the STAR aligner³⁵ in three modes: all reads in their original order (–outSAMtype BAM Unsorted –outSAMunmapped Within), only unique alignments (–outFilterMultimapNmax 1), and alignments to individual segments of chimeric reads (–chimOutType WithinBAM –chimSegmentMin 20 –chimScoreJunctionNonGTAG 0 –chimMainSegmentMultNmax 1). The reference consisted of the hg38 build of the human genome combined with the genome sequence of the PE243 strain of ZIKV. Each read of the pair was aligned separately to avoid any preference for alignment to the same genomic locus.

For each library, the pair of BAM files was collated and pair information was fixed using samtools³⁶. PCR duplicates were removed on the basis of their UMIs, using UMI-tools in paired mode.

Detecting significant interactions. We considered the ‘interaction space’ between the human and ZIKV genomes, which consisted of pairs of 1-kbp bins (one on each genome). For each replicate library in each condition (COMRADES and control), we counted the number of read pairs with one read in each bin using diffHic³⁷. This yielded a count matrix that was normalized via the trimmed mean-of-M-values method³⁸ to correct for composition biases, under the assumption that most read pairs mapping across the ZIKV and human genomes were caused by nonspecific ligation. We then applied the quasi-likelihood framework in edgeR³⁹ with two residual degrees of freedom for dispersion estimation to test for significant differences between the read pair counts for COMRADES and control. This was done with an additive design matrix that blocked on the batch to reflect the paired-sample design of the experiment.

Robust empirical Bayes shrinkage⁴⁰ was also used to stabilize the dispersion estimates in the presence of limited replication. Bin pairs were aggregated into clusters on the basis of whether they overlapped the same human gene.

We combined test statistics for each gene–Zika interaction using Simes’ method⁴¹ before applying the Benjamini–Hochberg method. Interactions that were significantly enriched in COMRADES compared with interactions in the control were defined at a false discovery rate threshold of 5%.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability. The RNA structure prediction pipeline can be downloaded from <https://github.com/gkudla/comrades>. The independent interactions prediction pipeline can be downloaded from <https://github.com/MarioniLab/ZikaPsolaren2017>.

Data availability

All sequencing datasets have been deposited in ArrayExpress under accession number [E-MTAB-6427](https://www.ebi.ac.uk/ena/browser/view/E-MTAB-6427). Base-pairing prediction, structure prediction, and clustering data are available in the Supplementary Data files. Source data for Figs. 1 and 3b are available online. Additional data that support the findings of this study are available from the corresponding authors upon request. A step-by-step protocol is available as a Supplementary Protocol and will be provided as an open resource in *Protocol Exchange*⁴².

References

- Chu, C., Qu, K., Zhong, F. L., Artandi, S. E. & Chang, H. Y. *Mol. Cell* **44**, 667–678 (2011).
- Kwok, C. K., Marsico, G., Sahakyan, A. B., Chambers, V. S. & Balasubramanian, S. *Nat. Methods* **13**, 841–844 (2016).
- Liu, Z.-Y. et al. *J. Virol.* **91**, e00484-17 (2017).
- Kwok, C. K., Ding, Y., Tang, Y., Assmann, S. M. & Bevilacqua, P. C. *Nat. Commun.* **4**, 2971 (2013).
- Schirle, N. T., Sheu-Gruttadauria, J. & MacRae, I. J. *Science* **346**, 608–613 (2014).
- Travis, A. J., Moody, J., Helwak, A., Tollervey, D. & Kudla, G. *Methods* **65**, 263–273 (2014).
- Langmead, B. & Salzberg, S. L. *Nat. Methods* **9**, 357–359 (2012).
- Markham, N. R. & Zuker, M. in *Bioinformatics: Structure, Function and Applications* (ed. Keith, J. M.) 3–31 (Humana Press, New York, 2008).
- Saldanha, A. J. *Bioinformatics* **20**, 3246–3248 (2004).
- Lai, D., Proctor, J. R., Zhu, J. Y. A. & Meyer, I. M. *Nucleic Acids Res.* **40**, e95 (2012).
- Darty, K., Denise, A. & Ponty, Y. *Bioinformatics* **25**, 1974–1975 (2009).
- Gower, J. C. *Biometrika* **53**, 325–338 (1966).
- Love, M. I., Huber, W. & Anders, S. *Genome Biol.* **15**, 550 (2014).
- Smith, T., Heger, A. & Sudbery, I. *Genome Res.* **27**, 491–499 (2017).
- Dobin, A. et al. *Bioinformatics* **29**, 15–21 (2013).
- Li, H. et al. *Bioinformatics* **25**, 2078–2079 (2009).
- Lun, A. T. L. & Smyth, G. K. *BMC Bioinformatics* **16**, 258 (2015).
- Robinson, M. D. & Oshlack, A. *Genome Biol.* **11**, R25 (2010).
- Lun, A. T. L., Chen, Y. & Smyth, G. K. *Methods Mol. Biol.* **1418**, 391–416 (2016).
- Phipson, B., Lee, S., Majewski, I. J., Alexander, W. S. & Smyth, G. K. *Ann. Appl. Stat.* **10**, 946–963 (2016).
- Lun, A. T. L. & Smyth, G. K. *Nucleic Acids Res.* **42**, e95 (2014).
- Ziv, O. & Miska, E. A. COMRADES: crosslinking of matched RNAs and deep sequencing. *Protoc. Exch.* (in the press).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

The code used for sequencing data collection can be found on github: <https://github.com/tdido>. The FACS data collection was done by: BD FACSDiva version 6.1.3.

Data analysis

The code used for analyzing the COMRADES data can be found on github: <https://github.com/gkudla/hyb>. Additional analysis softwares used to analyze the data: hyb package version: Nov 20 2013; Java Treeview version: 1.1.6r2; R-chie version: R4RNA 0.1.4; GraphPad Prism version 7.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All sequencing data sets have been deposited in ArrayExpress under accession number: E-MTAB-6427

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	COMRADES experiment was independently performed 3 times in different days, thus meeting the standards of next-generation sequencing studies.
Data exclusions	No data were excluded from the analyses. all attempts at replication were successful.
Replication	All experiments were repeated at least 3 times
Randomization	The data was not randomized since the experiments did not include allocation of samples to groups
Blinding	Data was not blinded, since the crosslinked and control libraries undergo different order of processing steps

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

Methods

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	Flavivirus group antigen antibody against the NS1 envelope protein, Novus biologicals, D1-4G2-4-15 (4G2), lot: T1650A04, diluted 1:100. Goat anti-Mouse IgG secondary antibody, Alexa Fluor 488, Eugene, A11029, lot: 1550911, dilution: 1:1,000.
-----------------	--

Validation	Flavivirus group antigen antibody against NS1 was validated by (Chavali PL et al, Science. 2017 Jul 7;357(6346):83-88). Secondary antibody is commonly used and established antibody. Both Flavivirus group antigen antibody against NS1 and the secondary antibody were validated by the authors by positive and negative staining of cells inoculated / not inoculated with Zika virus respectively.
------------	--

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	JEG-3 and Hela cells were purchased from ATCC; VERO cells were purchased from Sigma
Authentication	Cell lines were purchased from commercial providers and were not authenticated by the authors
Mycoplasma contamination	All cell lines were tested negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Jeg-3 cells pre-inoculated with ZIKV were stained according to BD Cytofix/Cytoperm kit protocol, using a primary monoclonal antibody: D1-4G2-4-15 (4G2), and a secondary antibody Alexa Fluor 488 (A-11029, Thermo Fisher Scientific).
Instrument	LSRFortessa, BD Biosciences
Software	Data collection was done using BD FACSDiva. Data analysis was done using Flowjo
Cell population abundance	Cell population abundance is shown in supplementary Fig. 9h. Initial gating based on FCS-SSC values resulted in collecting 88 +/-4% of the entire cell population
Gating strategy	Gating strategy is shown in supplementary Fig. 9h-j. ZIKV virus positive cells are defined as ZIKV positive gating divided by all cells gating

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.