nature portfolio

Corresponding author(s): Utpal Pal

Last updated by author(s): 11/11/2022

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the 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
☐	☒ A statement on 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 statistical parameters 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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Software used for genomic data collection are given in the methods, e.g., SMRTLink v8.0 for CCS and SMRTLink v9.0 for Iso-Seq. References for all such software are given. Protein identification was carried out using Proteome Discoverer software (Version 2.4) (Thermo Fisher Scientific), as indicated in the Supplementary Methods section.
Data analysis	Software used for data analysis are given in the methods, e.g., hifiasm 0.5-dirty-r247, hifiasm version 0.14.2 (r315), hicanu v2.0, and IPA v 1.3.1 for assembly. References for all software are given in the manuscript.

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All sequence data and genome assembly data are deposited into NCBI under BioProject accession number PRJNA678334. The data is also available through VEuPathDB (VectorBase release 59, 08/30/2022). All other datasets generated and analyzed during the current study are available in the supplementary information or the source data.

Field-specific reporting

Please select the one below that is 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/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	The sample size for each of the experiments was determined based on parallel studies in the field, including ours (Smith, A.A. et al. Cell Host Microbe 20, 91-8, 2016; Narasimhan et. al., Nature Comm 8: 184, 2017; Kingan SB et al. Genes Basel, 10:62.2019), in accordance to the common practice to generate statistically meaningful data, as appropriate. RAD-Seq sample sizes were based on simulations and empirical or other studies demonstrating that samples of ≥8 individuals are sufficient (e.g., Nazareno et al. 2017, Minimum sample sizes for population genomics... Mol. Ecol. Res. 17:1136 & references therein).
Data exclusions	As stated in the Supplementary Methods, for the genome assembly, we ran purge dups on the resulting primary contigs to remove duplicated haplotypes. The HiFi reads with residual SMRTbell adapters or Ultra Low Primer Adapter Sequences for both libraries were removed prior to downstream analysis. For scaffolding, alignments were filtered with samtools using the -F 2304 filtering flag to remove non-primary and secondary alignments. Bacterial reads were removed prior to the final assembly of the tick genome contigs. For assembled Rickettsia genome, the 16S sequences extracted from the HiFi reads were filtered to remove short sequences (<200 bp); singleton indels were also ignored, as they are the primary error profile for HiFi reads. RAD-Seq analyses include only loci that were present in >60% of individuals in at least 2 populations; SNPs were excluded if minor allele frequency <5%, loci missing in >20% of individuals, or coverage <6x or >200x. For proteomic analysis, only proteins identified in Ixodes protein database with <1% FDR were included in down stream analysis. Parsimony was applied to protein list to generate the smallest list of proteins that explained all peptides identified.
Replication	The replication was built into our sample size; all experiments with quantitative data included at least three independent biological replications, all with similar results.
Randomization	Genomic and proteomic analyses were conducted in a non-randomized order as we do not expect batch variations. All experimental studies were carried out in a randomized order, such as for control and experimental groups to avoid biased outcomes. The analyses of RAD-Seq data were structured by collection location.
Blinding	The nature of the study precludes the need for blinding, as the results are expected to yield objective outcomes, and therefore are not subject to individual bias for interpretation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Polyclonal antisera against recombinant Ash2 were generated in mice and used in a dilution of 1:100. Anti-PH3 (Phospho-Histone H3 antibody, catalog no. 9701; Cell Signaling Technology, Inc.) was used in the dilution of 1:200.
Validation	The primary antibodies were validated using standard Western blotting and confocal immunofluorescence methods, as presented in the manuscript. For validation purposes, we checked the immunoreactivity of the antibody against target antigens by comparing with normal sera or isotype control antibodies. The commercially-obtained antibody is further validated by the manufacturer.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Aliquots of a frozen stock of ISE6 cells were used at the University of Georgia; the ticks used to create this cell line originated in Georgia, as described in Munderloh et. al, J. Parasitol 80, (1994).
Authentication	Cells were authenticated by aligning millions of DNA sequencing reads to the tick genome
Mycoplasma contamination	The cell line was not tested for mycoplasma contamination but this would not alter our sequencing results
Commonly misidentified lines (See ICLAC register)	None

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	We used Ixodes scapularis obtained from the Oklahoma State University colony. We used four- to six-week-old female C3H/HeN mice and approximately 20-week-old female Hartley guinea pigs, which were bred in-house or purchased from Charles River Laboratories. The animals were housed under 12:12 dark-light cycle, 30-70% relative humidity, and 68-79F ambient temperature. These information are also detailed in the Methods.
Wild animals	For the collection of wild populations of mature I. scapularis ticks, flat (unfed or minimally fed) adults were collected from horses in Oklahoma (latitude: 36.1450, longitude: -97.0068; 12 males) and from white-tailed deer in Rhode Island (lat: 41.4568, long: -71.6673; 9 males, 4 females), Pennsylvania (lat: 40.2354, long: -79.4704; 11 males, 2 females), and South Carolina (lat: 33.2464, long: -81.6679; 9 males, 4 females), Louisiana (lat: 32.7551, long: -93.6623; 5 males, 8 females), Wisconsin (lat: 45.3582, long: -88.0650; 6 males, 6 females). All ticks were killed by dipping in 70% ethanol, stored and sent to the University of Georgia for morphological identification, and subsequently underwent DNA extraction and sequencing required for RAD-Seq analysis.
Field-collected samples	Ticks were collected from the field; as described above, no live field-collected ticks were used, rather they were killed and preserved in 70% ethanol for subsequent analysis.
Ethics oversight	All animal experiments were performed in accordance with the ethical and animal care guidelines approved by the University of Maryland Institutional Animal Care and Use Committee and the Institutional Biosafety Committee, as stated in the Methods section of the manuscript.

Note that full information on the approval of the study protocol must also be provided in the manuscript.