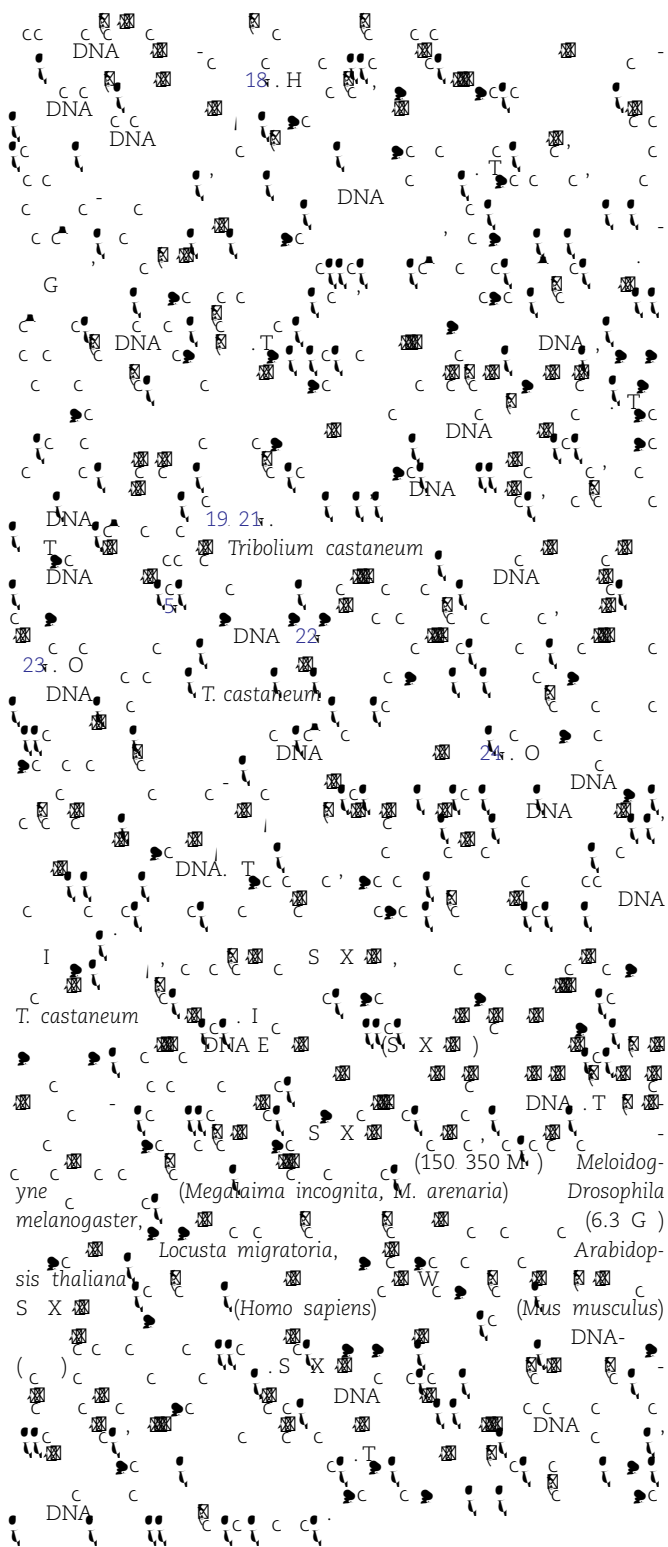
























This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com



Materials and methods

Genome and satDNA data

S X                      

T. castaneum, D. melanogaster, L. migratoria, M. incognita, M. arenaria,
A. thaliana, H. sapiens, and M. musculus

Diagram illustrating the structure of a DNA molecule, showing a double helix with various components labeled: S, C, T, I, and DNA. The diagram includes a legend for the components: S (blue square), C (blue circle), T (black circle), I (black square), and DNA (black line). The structure shows a double helix with a central core and various components attached to the strands.

SatXplor pipeline overview

Detection and extraction of monomers

▲ — evaluate 10 outfmt 6 max target_seqs 10000 task blastn
num_threads 2 — dust no — soft_masking false ▲

Array creation

Monomer density profiles

T 2D BLAST C C C C
G, P F , 2D C C C C
scipy.stats.gaussian_kde().F
2D NKERNEL_BINS

K-mer-based profiling of arrays

T

DNA

1. E

2. C

3. E

4. C

5. F

6. C

7. D

87%

DNA

+/-500

32

H

+/- 5
























32

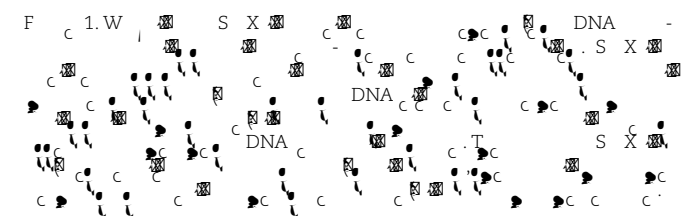
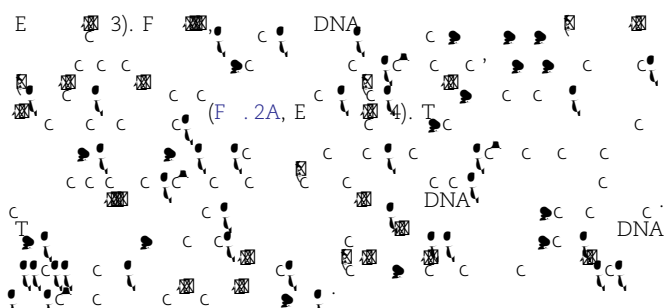
87%

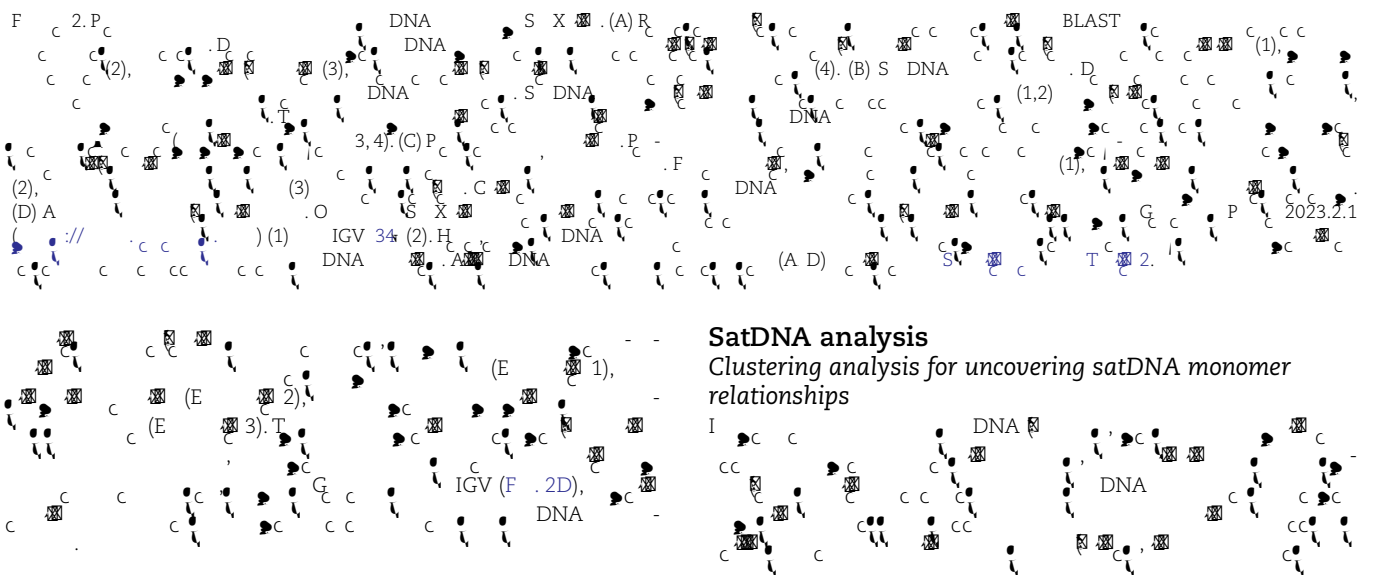
Alignment and distance calculations

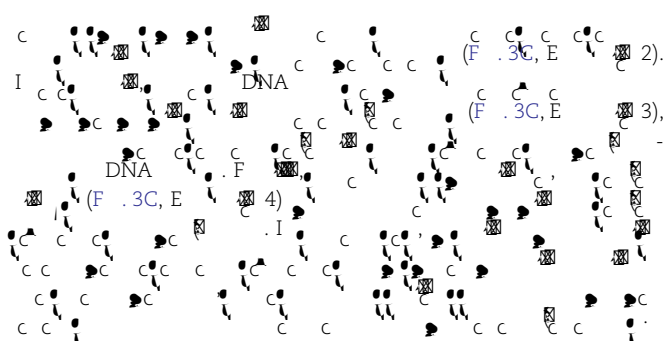
T C MAFMT 26 P

`mafft --adjustdirection reorder --threads str(half_threads)`


W                       

[illegible]





- (T, A) . Sci Rep 2019;9:1 10.
8. S P, T W, S K. et al. H DNA N W . Sci Rep 2015;5:10315.
9. N S, K S, R A. et al. T . Science 2022;376:44 53.
10. N P, R L, K A. et al. TAREAN: DNA Nucleic Acids Res 2017;45: 111.
11. B G, T : DNA Nucleic Acids Res 1999;27:573 80.
12. O D, W T. ULTRA: A M T D T R . I : Proc. 2018 ACM Int. Conf. Bioinformatics, Comput. Biol. Heal. Informatics, A M , N , US, . 37 46, 2018. // /10.1145/3233547.3233604.
13. K R, B G, K G. M : DNA Nucleic Acids Res 2003;31: 3672 8.
14. B W, K KK, K O. R : Mob DNA 2015;6:4 9.
15. S AFA, H R, G P. RepeatMasker Open-4.0. 2013 2015. //
16. W P, H M, H IR. TRASH: Bioinformatics 2023;39: 308.
17. K I, K E, D M. et al. A N TRF de novo DNA Plan Theory 2022;11:2103.
18. V MR, K P, P AM. et al. S G : Bioinformatics 2022;38:2049 51. // /10.1093/ / 018.
19. R -R FJ, L -L MD, C J. et al. H DNA Sci Rep 2016;6:28333.
20. V T, R L, N P. et al. C DNA Plant J 2020;101:484 500. // /10.1111/14546.
21. S JS, K DE, E DG. et al. D Mbl Biol Evol 2020;37: 2241 56. // /10.1093/ / 078.
22. G T, D -S E, M N. et al. C H3 PLoS Genet 2020;16: 1009115.
23. G T, D M, D -S E. et al. T DNA (Basel) 2023;14:999.
24. V M, D -S E, V D. et al. L Tribolium castaneum DNA Genome Res 2024;34:1878 94. // /10.1101/279225.124.
25. C C, G, A V. et al. BLAST+: BMC Bioinformatics 2009;10:421.
26. K K, M K, K KI. et al. MAFFT: Nucleic Acids Res 2002;30:3059 68.
27. R C, T : R: A Language and Environment for Statistical Computing. R F S , V , A , 2023. // .R- /.
28. L S, J, H F. F M R: R J Stat Softw 2008;25:1 18.
29. W H. ggplot2: Elegant Graphics for Data Analysis. S V N . 2016. ISBN 978-3-319-24277-4. // 2.
30. C L LG, H SL, G JL. O 1.688. DNA G3 Genes Genomes Genet 2020;10:4129 46.
31. D -S E, M B, S S. et al. T 19- Meloidogyne N Mol Biol Evol 2021;38:1943 65. // /10.1093/ / 336.
32. D -S E, S S, M B. et al. S DNA BMC Biol 2022;20:1 19.
33. S CR, G J, M M. et al. C A Nucleic Acids Res 1988;16:6753 66.
34. R JT, T H, W W. et al. I Nat Biotechnol 2011;29:24-6. // /10.1038/1754.
35. L, J R, G T. et al. E DNA: Signal Transduct Target Ther 2022;7:342.
36. L S, L M, F D. et al. D DNA 0 0004T 05.) T M C P (.) T M -0004T QS

- 
 BRIEFINGS IN BIOINFORMATICS 2025, 26(1), c 660
 DOI: 10.1093/bib/bbaf001