Understanding The Evolutionary Dynamics Of Transposable Elements in *Drosophila* Via *de novo* Identification and Classification

Ellison, Christopher PhD and Rele, Chinmay
Department of Genetics, Rutgers University

Introduction

1. Transposons (TEs) are mobile genomic elements.
2. TEs and other repeat elements are very abundant in genomes.
3. Can insert into genes, alter gene expression, may be adaptive or deleterious, or may cause chromosomal rearrangements via ectopic recombination.
4. Sequencing allows us to find these TEs and see their effects on the genome.

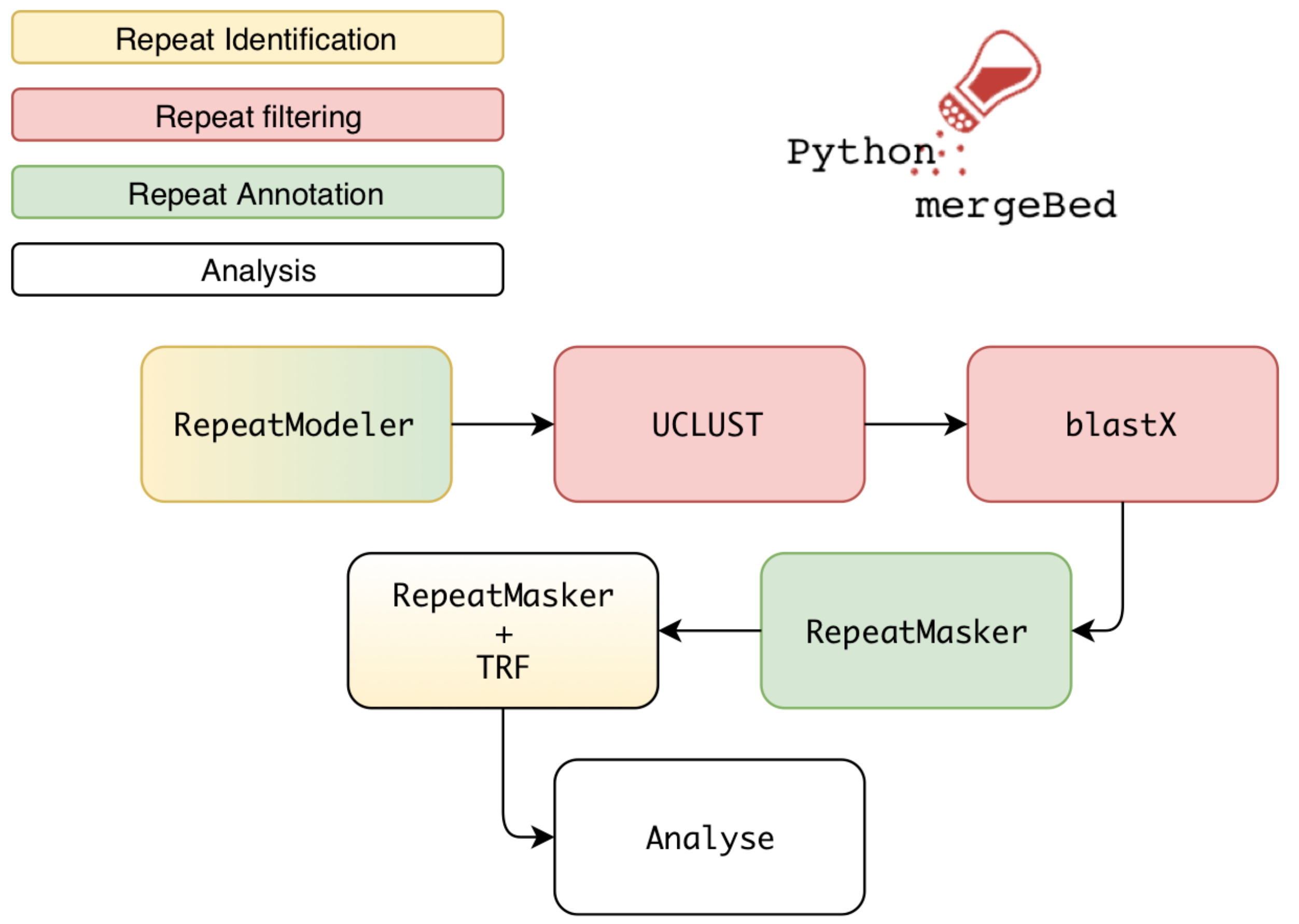
Shortcomings of Previous Work

1. Previous studies:
    1. Large variation in sequencing efforts
    2. Variable protocols/parameters for assembly generation.
    3. Short read
    4. Homology-based search
2. We needed to account for all of these:
    1. Same sequencing methods and protocols: Used single dataset for analysis.
    2. Longer reads from Nanopore data
    3. *De novo* based search allows us to identify novel repeats

Research Questions

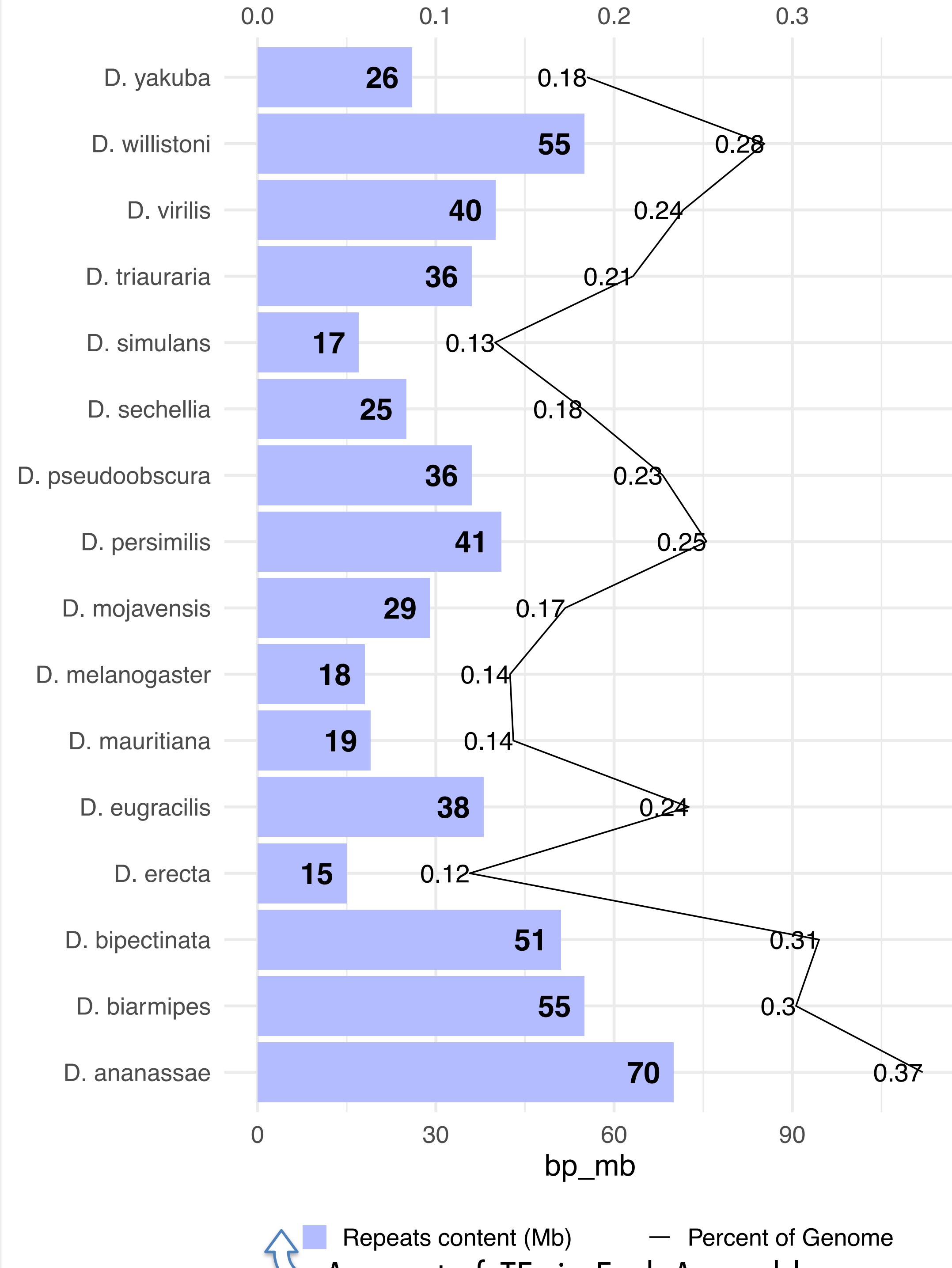
1. Questions:
    1. Any undiscovered TEs?
    2. What drives genome size?
    3. Correlation between Satellite/Simple Repeats and genome size
    4. Reduction in melanogaster group.

Methods



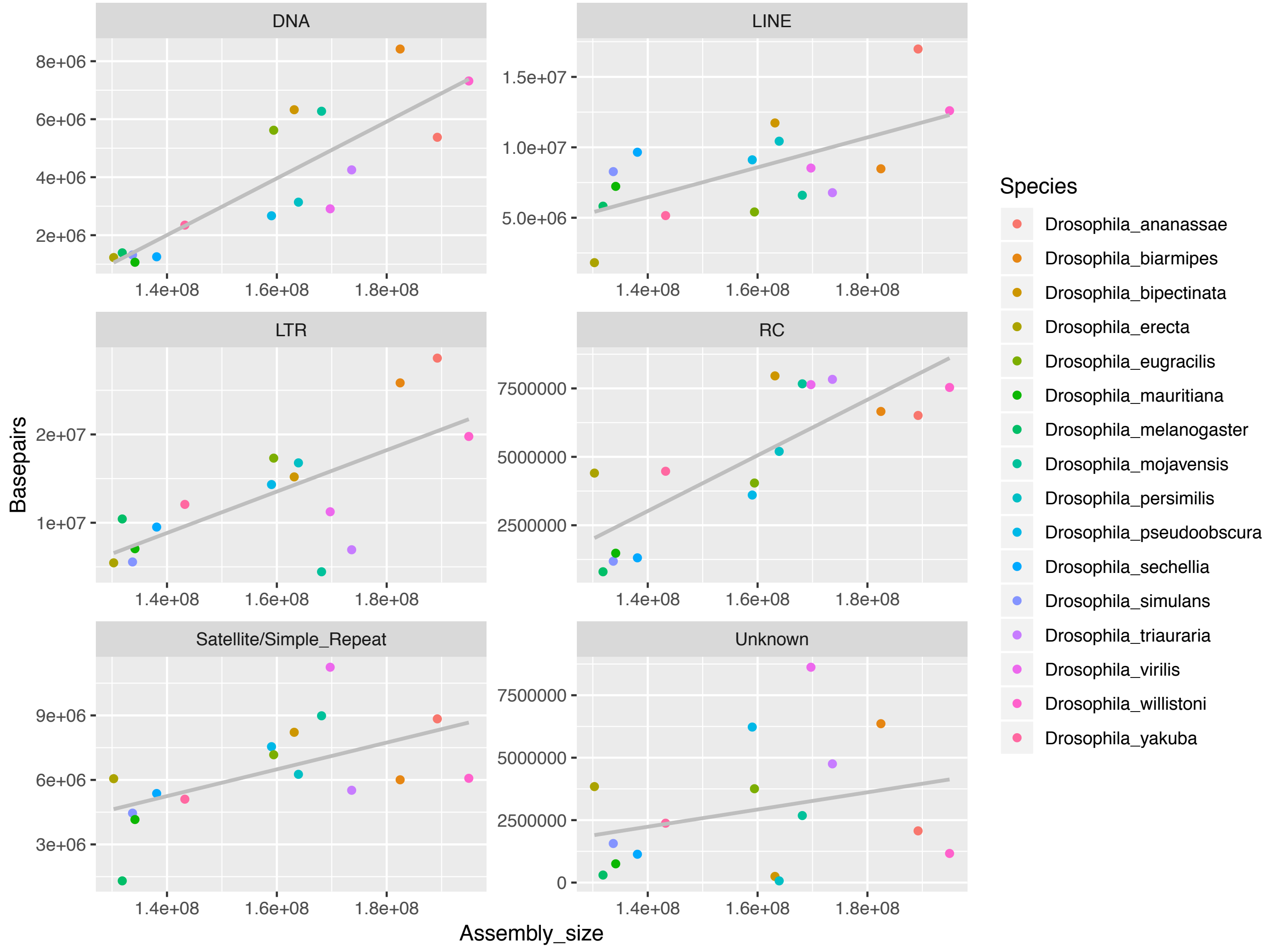
Results

TES RANGE FROM 12% TO 37% OF FLY GENOMES.



Amount of TEs in Each Assembly

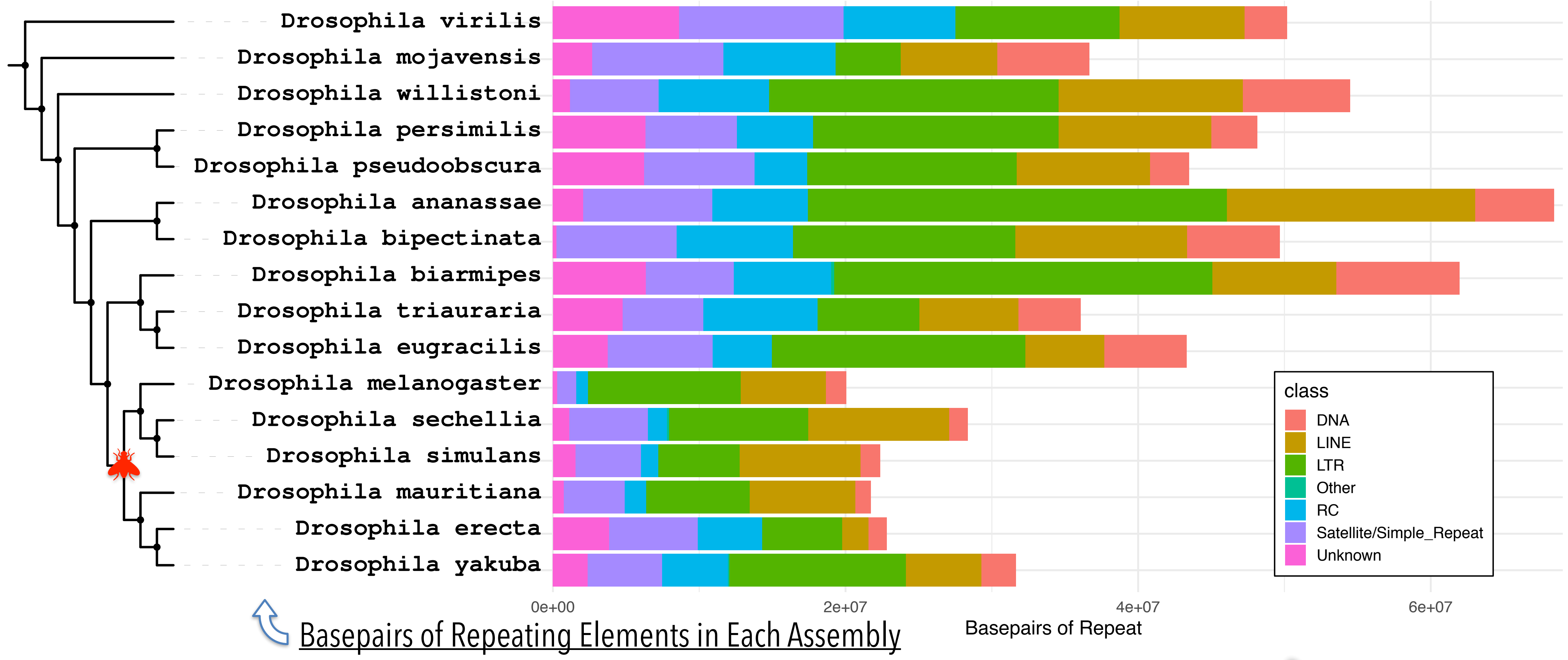
DNA ELEMENTS ARE MOST CORRELATED WITH GENOME SIZE



Element	r_s	p	**Element**	r_s	p
DNA	0.829 412	0.000 080	LINE	0.567 647	0.024 020
LTR	0.585 294	0.019 310	RC	0.752 941	0.001 147
Sat/SR	0.576 471	0.021 560	Unknown	0.261 765	0.326 200

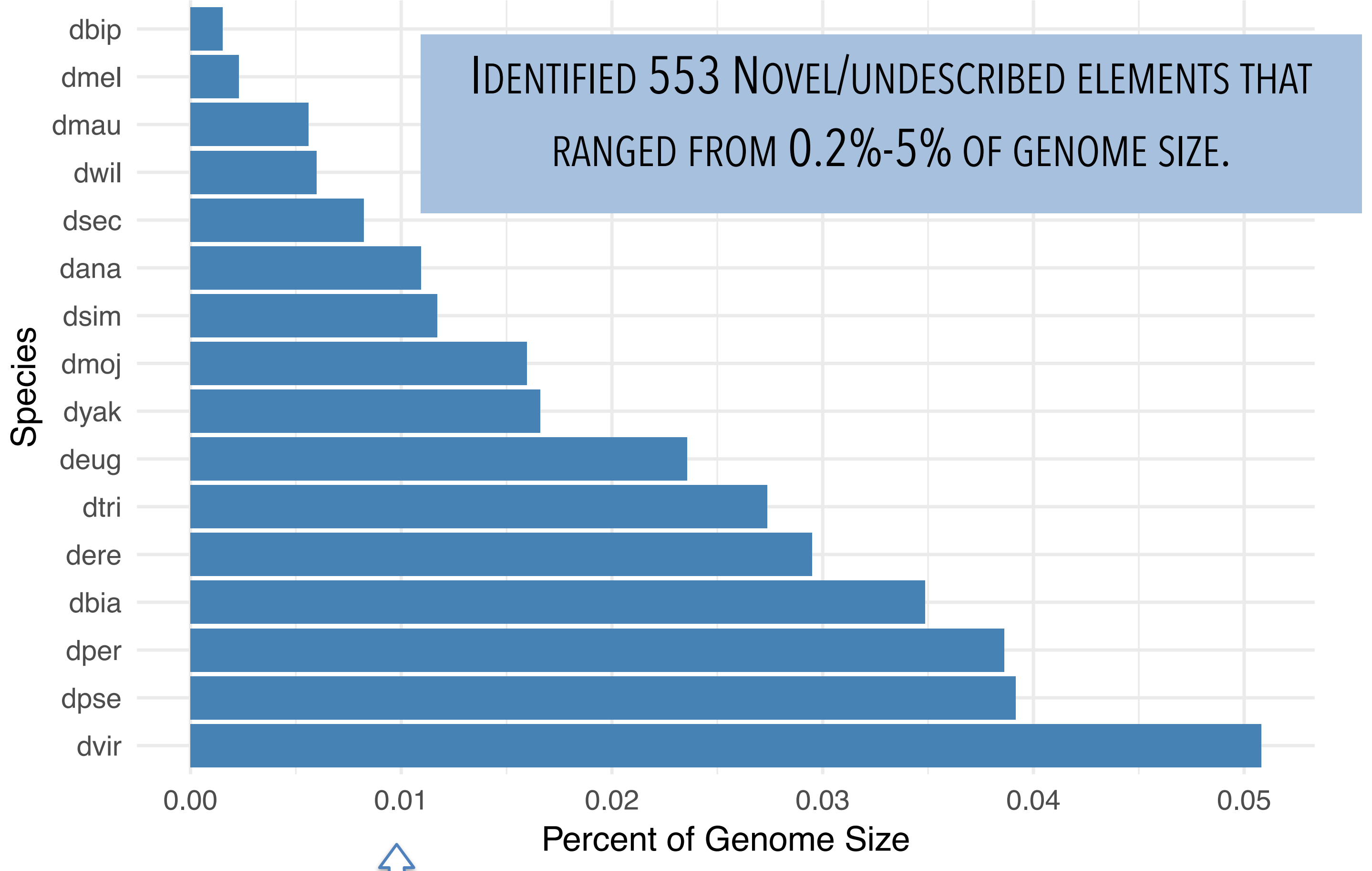
Correlation of all Classes of Repeat Elements

REDUCTION IN GENOME SIZE IN *MELANOGASTER* GROUP DUE TO REDUCTION IN ALL TE CLASSES



Basepairs of Repeating Elements in Each Assembly

IDENTIFIED 553 NOVEL/UNDESCRIBED ELEMENTS THAT RANGED FROM 0.2%-5% OF GENOME SIZE.



Novel Elements as Percent of Genome Size

Conclusions

- Previous studies using short-read data cannot identify TEs efficiently.
- Homology-based searches cannot identify NOVEL/undescribed TEs.
- Reduction of genome size in *melanogaster* group attributed to reduction in TE content.

Theories for reduction in TE content

- Stochastic reduction of TEs (unlikely).
- Arrival of gene that more effectively controls for TE content.
- Larger effective population size.

Acknowledgements

- Ellison Lab
- Xing, Jinchuan PhD
- Shah, Premal PhD
- Office of Advanced Research Computing
- Aresty Research Center

References

1. D. 12 G. Consortium, Evolution of genes and genomes on the Drosophila phylogeny, Nature. 450 (2007) 203–218. doi:10.1038/nature06341.
2. D.E. Miller, C. Staber, J. Zeitlinger, R.S. Hawley, Highly Contiguous Genome Assemblies of 15 Drosophila Species Generated Using Nanopore Sequencing., G3 (Bethesda). 8 (2018) 3131–3141. doi:10.1534/g3.118.200160.
3. J. Gonzalez, K. Lenkov, M. Lipatov, J.M. Macpherson, D.A. Petrov, High rate of recent transposable element-induced adaptation in Drosophila melanogaster., Plos Biol. 6 (2008) e251. doi://dx.doi.org/10.1371/journal.pbio.0060251.
4. A.S. Wilson, B.E. Power, P.L. Molloy, DNA hypomethylation and human diseases, Biochim. Biophys. Acta - Rev. Cancer. 1775 (2007) 138–162. doi:10.1016/J.BBCAN.2006.08.007.

Questions I have:

1. Using Keynote — Tried exporting to PPTX, but graphs are getting pixelated. Not an issue in KeyNote or a PDF. Please advise.
   1. I checked, and this PDF prints to the correct size.
2. Made title larger and got rid of the “R”.
   1. I feel like something Rutgers related should be there, so added the knight.
   2. Want to keep the fly/Transposons image as it shows our model organism and the field of genetics we are in..
3. Defined TEs and rechecked any grammatical or clerical errors.
4. I can see the lines when I’m zoomed out as well, but when I zoom in, I don’t see them anymore.
   1. I do not intend for the lines to be shown, but they seem to be appearing in the PDF.
5. I have attempted to shrink crowded panels, but I still need that data to be shown.

TEs range from 12% to 37% of fly genomes.

Add your information, graphs and images to this section.

Add your information, graphs and images to this section.

Add your information, graphs and images to this section.

Basepairs of Repeating Elements in Each Assembly

TEs range from 12% to 37% of fly genomes.

TEs range from 12% to 37% of fly genomes.