GENOME ASSEMBLY WEEK 6

POST-COURSE SURVEY

▸ https://drive.google.com/open?id=1z8np3TavOo82WPbyzUg1c7YRXgU_n1k-Rc4ee6Peals

VISUALIZING KRAKEN RESULTS

▸ see tutorial markdown file for instructions.

SOFTWARE YOU SHOULD FEEL COMFORTABLE WITH NOW…

- SRA Toolkit
- FastQC
- TrimGalore!

Pre-processing

- SPAdes
- DISCOVAR
- MaSuRCA

Assembly

- Paul's fasta metadata parser
- Genomescope
- Jellyfish
- QUAST
- LASTZ
- BUSCO
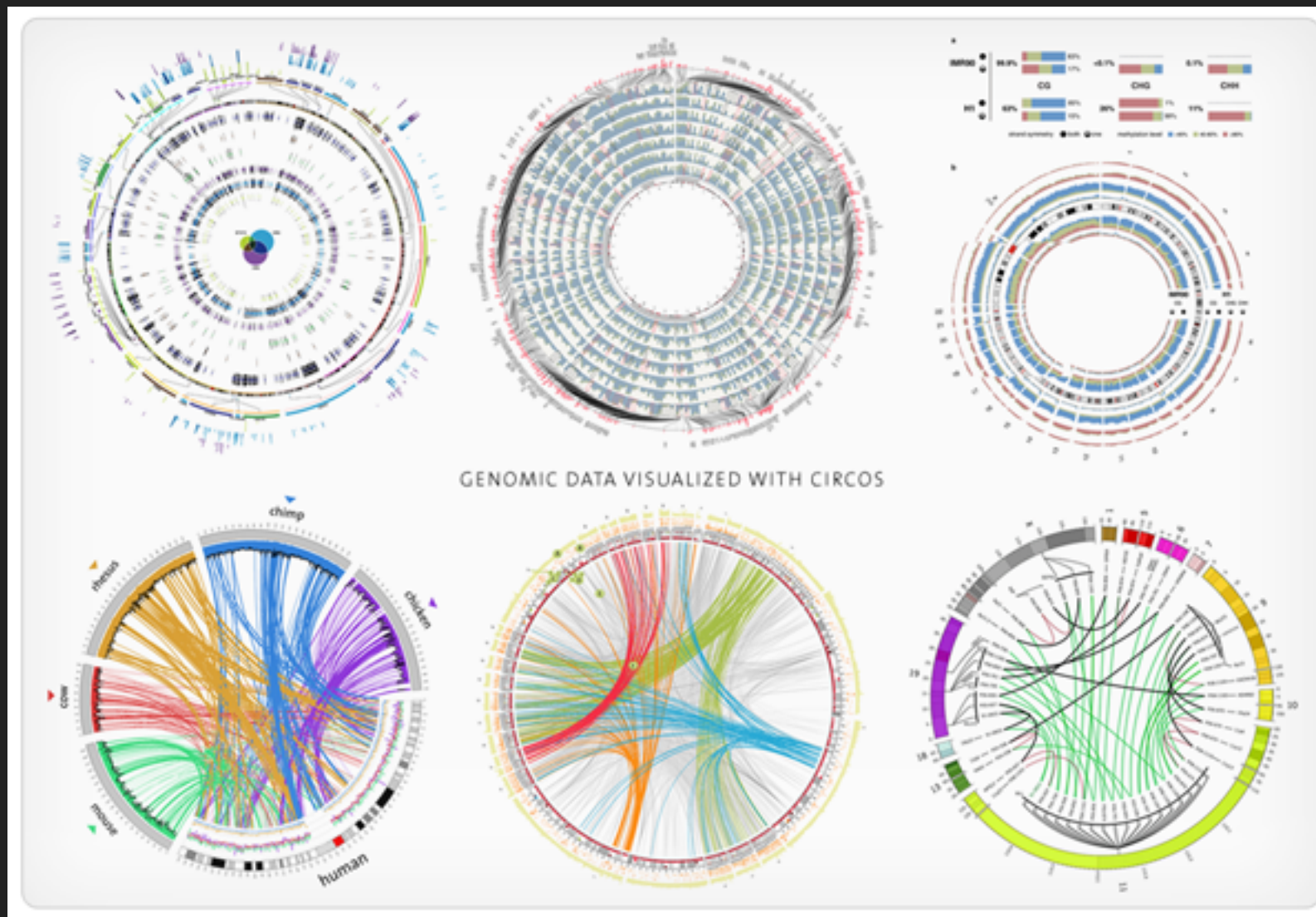- Kraken
- Krona

Post-processing

OTHER KINDS OF ASSEMBLIES

▸ Transcriptome

    ▸ Tools: Trinity

    ▸ Future RNA-seq workshop planned

OTHER KINDS OF ASSEMBLIES

▸ Organellar genomes:

▸ You can map WGS reads to a reference mitochondrial or chloroplast genome, pull out those reads and assemble *de novo*
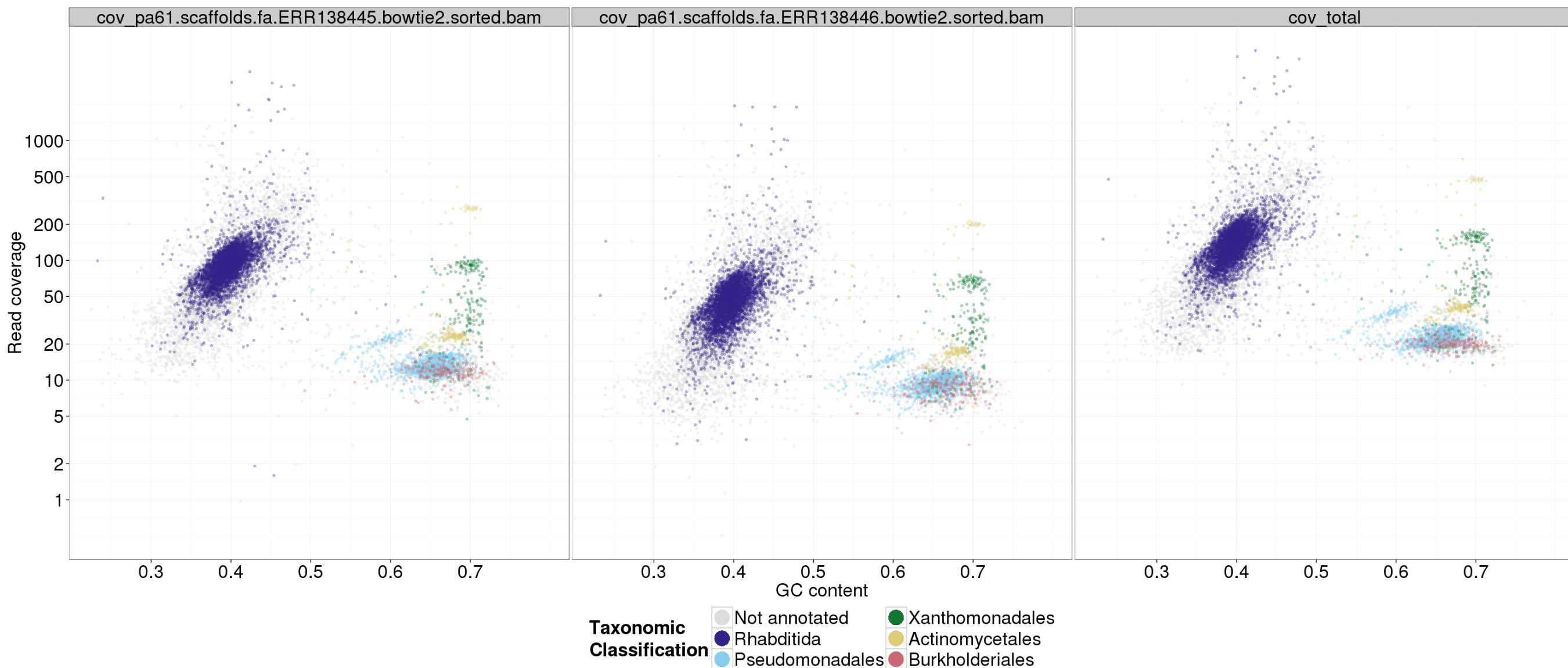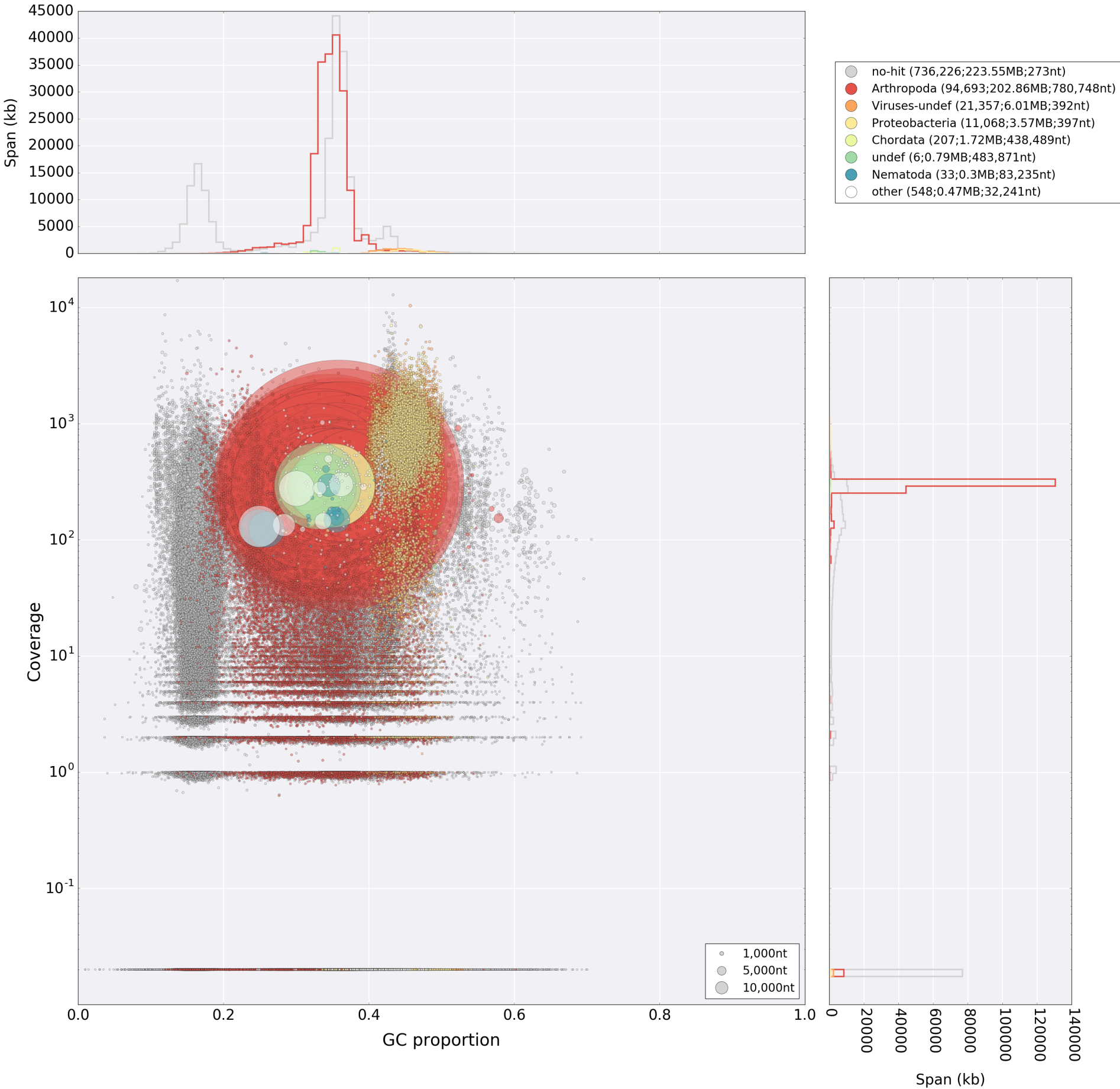
▸ Tools: mitobim, BWA

VISUALIZATION TOOLS

▸ Circos (http://circos.ca)

GENOMIC DATA VISUALIZED WITH CIRCOS

VISUALIZATION

▸ Blob tools: https://github.com/blaxterlab/blobology

BlobDB.json.bestsum.sam0

Legend:
- no-hit (736,226;223.55MB;273nt)
- Arthropoda (94,693;202.86MB;780,748nt)
- Viruses-undef (21,357;6.01MB;392nt)
- Proteobacteria (11,068;3.57MB;397nt)
- Chordata (207;1.72MB;438,489nt)
- undef (6;0.79MB;483,871nt)
- Nematoda (33;0.3MB;83,235nt)
- other (548;0.47MB;32,241nt)

VISUALIZATION TOOLS

▸ JBROWSE: (http://jbrowse.org)

▸ Fast, embeddable javascript genome browser

VISUALIZATIONS WE ARE WORKING ON

▸ https://hydra-3.si.edu/genomes:

   ▸ not yet available outside SI-staff

WEB APOLLO

▸ http://genomearchitect.org

▸ Community annotation

WHAT'S NEXT ONCE YOU HAVE AN ASSEMBLY?

▸ Annotation!

▸ Vanessa gave a Peer-Led workshop recently on using Augustus. See here: https://github.com/SmithsonianWorkshops/Peer-Led-Bioinformatics/tree/master/2016

▸ Paul recently installed the MAKER (http://www.yandell-lab.org/software/maker.html) pipeline on Hydra. Feel free to test it and tell us how it goes.