

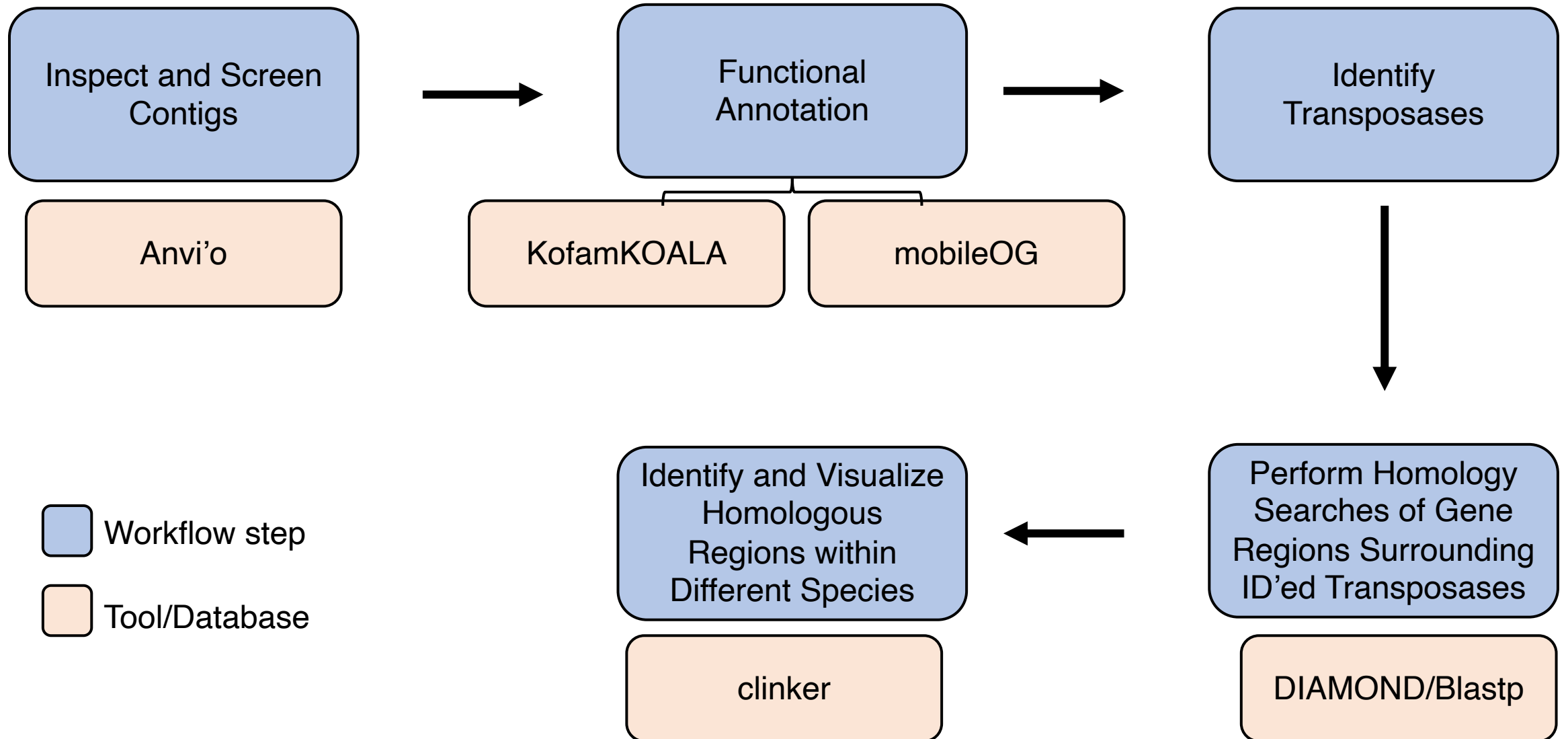
Detecting Signatures of Transposase-Mediated Inter-Species Horizontal Gene Transfer in the Infant Gut Microbiome

Elizabeth McDaniel
Coding Assessment

Hypothesis 1: Transposons are prevalent mobile elements of the infant gut microbiome

Hypothesis 2: Inter-species horizontal gene transfer in the infant gut microbiome could be mediated by transposons

Overall Workflow



Contigs Inspection and Functional Annotation

- Initial contigs inspection in Anvi'o identifying single copy core genes for assessing quality of genomes and the taxonomical composition of the community
 - All 71 bacterial single copy core genes found in the metagenome
 - 4 bacterial species contained most of the 71 bacterial single copy core genes, the “unknown” bin only contained 4 single copy core genes
 - Classifications based on the Ribosomal L2 gene matched the given classifications, “unknown” bin classified as Desulfobacterota but thrown out of downstream analysis because low-quality
- KofamKOALA pipeline resulted in 3785 significant annotations
- mobileOG pipeline resulted in 517 significant annotations

Transposase Identification and Homology Search

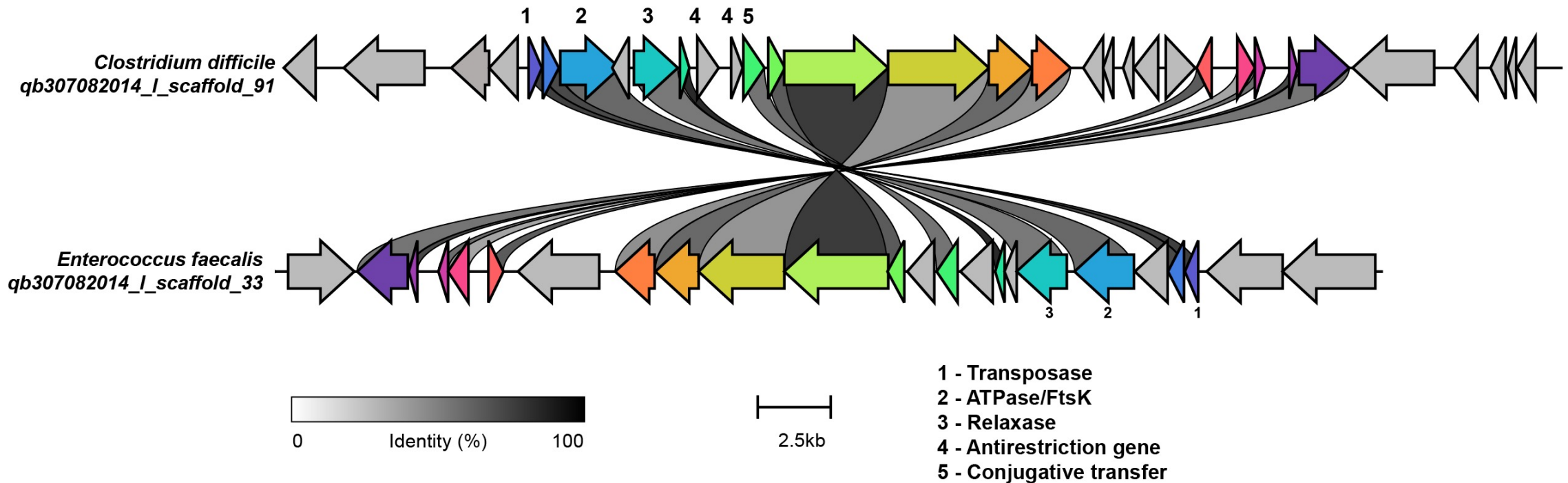
The mobileOG pipeline is a curated database of mobile elements consisting of:

1. ICEBerg 2.0 (ICEs, AICEs, CIMEs, IMEs)
2. COMPASS (plasmids)
3. NCBI Plasmid RefSeq (plasmids)
4. Gut Phage Database (phages)
5. ACLAME (various)
6. immedb (integrative elements)
7. Prokaryotic viral orthologous groups (pVOG)

From this pipeline, I identified 36 transposases among the 4 bacterial species

I then retrieved the gene neighborhoods of these transposases (10 genes upstream and downstream of transposase depending on contig length) and performed a BLASTP search against the contigs of the infant gut microbiome, focusing on results that showed homology to neighborhoods of other species

C. difficile and *E. faecalis* transposase neighborhood



Color refers to the homologous group between the two species, shade of lines between the plots refers to percent identity of the regions, and numbered regions are functional annotations retrieved from either the KofamKOALA or mobileOG databases – absence of number infers no confident annotation could be made

Conclusions and Caveats

Main conclusions:

- Transposases are prevalent in the 4 species of bacteria in this infant gut microbiome
- Homology analysis infers a putative transposase-mediated inter-species horizontal gene transfer between the *Clostridium difficile* and *Enterococcus faecalis* species that shares a high percent identity and is poorly functionally characterized by the KofamKOALA and mobileOB databases

Caveats:

- Initial contig inspection would also include read mapping to check uniform coverage, assessing bins with CheckM and GTDB
- Additional annotation databases can be ran, parameters can be relaxed to find more distant homologs, and annotations need to be confirmed by inspecting conserved domains