



Probability distributions of

- sequence pairwise similarities within species
- sequence pairwise similarities within genera

Clustering thresholds

- Intersection
- Species 10% quantile
- Genus 90% quantile
- Midpoint between modes