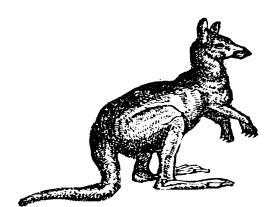## Genome Directory

Define the number of genomes to pick (default = 10) and number of reads
to generate per file (default = 100,000)

Execute **sim** mode

PD on concatenated tree

Determines PD contributions for taxa present in concatentated guide tree in PhyloSift marker directory

Select Taxa
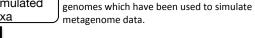
Two separate approaches used:
1. Select a set of taxa that contribute to PD (user input, default = 10)
2. Same number of taxa are randomly select taxa across PD list

Generate Simulated Reads

Grinder algorithm randomly generates reads from selected genomes, outputs simulated PE-Illumina and 454 datasets

Mask Simulated Taxa

Workflow plugs into updateDB to mask genomes which have been used to simulate metagenome data.

Cloned Marker Directory

A cloned marker directory is created, where simulated genomes have been removed from marker packages.