

The diagram illustrates a workflow for protein homology search. On the left, under the heading "Focal genome", there are two document icons: the top one is labeled "GFF3" and the bottom one is labeled "FASTA". An arrow labeled "Protein homology search" points from the "FASTA" file to the right. On the right, under the heading "Neighbor genomes", there is a 2x3 grid of document icons. Below this grid is a 3D cube icon labeled "NR".

Search for orphan genes or TRGs

NO protein hits

NO protein hits

NO protein hits

NO protein hits

TRGs

Focal genome

Search for homology in the noncoding regions of outgroup species

Hits in noncoding regions of outgroup species

NO hits

Focal genome

Investigate micro-synteny for non-genic hits

The diagram illustrates a phylogenetic tree on the left and a corresponding micro-synteny analysis on the right. The tree shows a root branching into two main lineages. The left lineage leads to a node marked with a purple circle (non-coding) and then to a node marked with a dark green circle (de novo gene birth). The right lineage leads to a node marked with a dark green circle. An arrow points to the tip of the tree, labeled 'Focal genome'. The micro-synteny analysis on the right shows gene clusters for different lineages. The top three lineages are labeled 'NO hits'. The next three lineages are labeled 'Hit not in synteny', 'Hit in flexible synteny', and 'Hit in synteny'. The bottom four lineages are labeled 'validated de novo gene'. A legend on the right indicates that a purple circle represents 'non-coding' and a dark green circle represents 'de novo gene birth'.

NO hits

Hit not in synteny

Hit in flexible synteny

Hit in synteny

validated
de novo gene

Focal genome

● non-coding

● *de novo* gene birth