

---

```

36: Mapping training genome or user genome from GTDB taxonomy to NCBI
    taxonomy
37: Initializing final hit list as  $\mathcal{L}_{hit_{final}}$ 
38: for  $rank = r_{species}, r_{kindom}$  do
39:   if  $\mathcal{D}_{gtdb\_taxa}[rank]$  is unassigned then
40:     Continue
41:   end if
42:   Initializing hit list for  $\mathcal{D}_{ncbi\_taxa}$  with  $\mathcal{L}_{hit_{genome}}$ 
43:   Mapping genome accession gid to tree leaf node  $\mathcal{NODE}_{map}$ 
44:   traversing leaf nodes  $\mathcal{NODE}_{map}$  to find NCBI decendants with
     $\mathcal{N4G}_{cluster}[rid]$ 
45:   Storing  $\mathcal{N4G}_{cluster_{hit}}$  as list  $\mathcal{L}_{cluster_{hit}}$ 
46:   for  $rid$  in  $\mathcal{L}_{cluster_{hit}}$  do
47:     appending  $\mathcal{N4G}_{cluster}[rid][rank]$  into  $\mathcal{L}_{hit_{genome}}$ 
48:     Counting element occurence in  $\mathcal{L}_{hit_{genome}}$ 
49:     Applying heapsort to  $\mathcal{L}_{hit_{genome}}$  and finding top 1 element
50:      $tax_{top}, count_{top} = \mathcal{HEAPSORT}(\mathcal{L}_{hit_{genome}})$ 
51:     if  $count_{top} \geq \theta' * \mathcal{L}_{hit_{genome}}.size()$  &&  $tax_{top}$  is NOT unassigned
    then
52:        $\mathcal{L}_{hit_{final}} = \mathcal{D}_{ncbi\_lineage}[tax_{top}]$ 
53:       break
54:     end if
55:   end for
56:   Storing  $\mathcal{L}_{hit_{final}}$ 
57: end for

```

---