**Algorithm 2** Rebuilding NCBI-Tree Denovo with Custom Taxonomy

---

**Input:** Custom taxonomy table with capacity $\mathcal{C}$ with column taxa as List $\mathcal{L}_{taxa7}$
**Output:** NCBI-TREE nodes.dmp and names.dmp for custom taxonomy

---

1: Initializing NCBI-Tree as Directed Acyclic Graph(DAG) with Dict $\mathcal{D}_{dag}$
2: Initializing vertices and edges in DAG with List $\mathcal{L}_{vertices}$ and $\mathcal{L}_{edges}$
3: Initializing taxa node ever seen with Dict $\mathcal{D}_{taxa_{seen}}$
4: Initializing taxa node ID with Dict $\mathcal{D}_{taxid_{node}}$
5:
6: **for** $genome = 1, \mathcal{C}$ **do**
7:   **for** $\varepsilon, taxon$ in enumerate($\mathcal{L}_{taxa7}$) **do**
8:     Adding Vertex $taxon$ into $\mathcal{D}_{dag}$ when not in $\mathcal{D}_{dag}$, or nothing to do
9:     **if** $taxon\ NOT$ in $\mathcal{D}_{dag}$ **then**
10:       $\mathcal{D}_{dag}[taxon] = []$
11:       Updating Taxa Node ID with occurence order of incrementing by 1
12:       $\mathcal{D}_{taxid_{node}}[taxon] = \text{len}(\mathcal{D}_{taxid_{node}}) + 1$
13:     **end if**
14:
15:     Updating Edges from root to leaf(RTL) in Linked List
16:     **if** $\varepsilon = 0$ **then**
17:       Adding Edge $taxon$ from root of $\mathcal{D}_{dag}$
18:       $\mathcal{D}_{dag}[root] = [taxon]$
19:     **else**
20:       Adding Edge $taxon$ to $taxon_{before}$ denoted with $\mathcal{L}_{taxa7}[\varepsilon - 1]$ of $\mathcal{D}_{dag}$
21:       $\mathcal{D}_{dag}[taxon_{before}] = [taxon]$
22:     **end if**
23:   **end for**
24: **end for**
25:
26: Recursive DFS to generate node.dmp and names.dmp
27: Initializing nodes and names as List $\mathcal{L}_{name}$ and $\mathcal{L}_{node}$
28: DFS($taxon$, $\mathcal{L}_{name}$, $\mathcal{L}_{node}$)
29: **for** $taxon_{child}$ in $\mathcal{D}_{dag}[taxon]$ **do**
30:   **if** $taxon_{child}$ in $\mathcal{D}_{taxa_{seen}}$ **then**
31:     Continue
32:   **end if**
33:   $\mathcal{D}_{taxa_{seen}}[taxon_{child}] = flag$
34:
35:   Appending nodes with $taxid, parent_i d, rank, ...$ format
36:   $\mathcal{L}_{node}.\text{append}(\mathcal{D}_{taxid_{node}}[taxon_{child}], \mathcal{D}_{taxid_{node}}[taxon], \mathcal{D}_{rank}[taxon_{child}], ...)$
37:
38:   Appending names with $taxid, name, ' scientificname'$ format
39:   $\mathcal{L}_{name}.\text{append}(\mathcal{D}_{taxid_{node}}[taxon_{child}], taxon_{child},$'scientific name')
40:   DFS($taxon_{child}$, $\mathcal{L}_{name}$, $\mathcal{L}_{node}$)
41: **end for**
42: