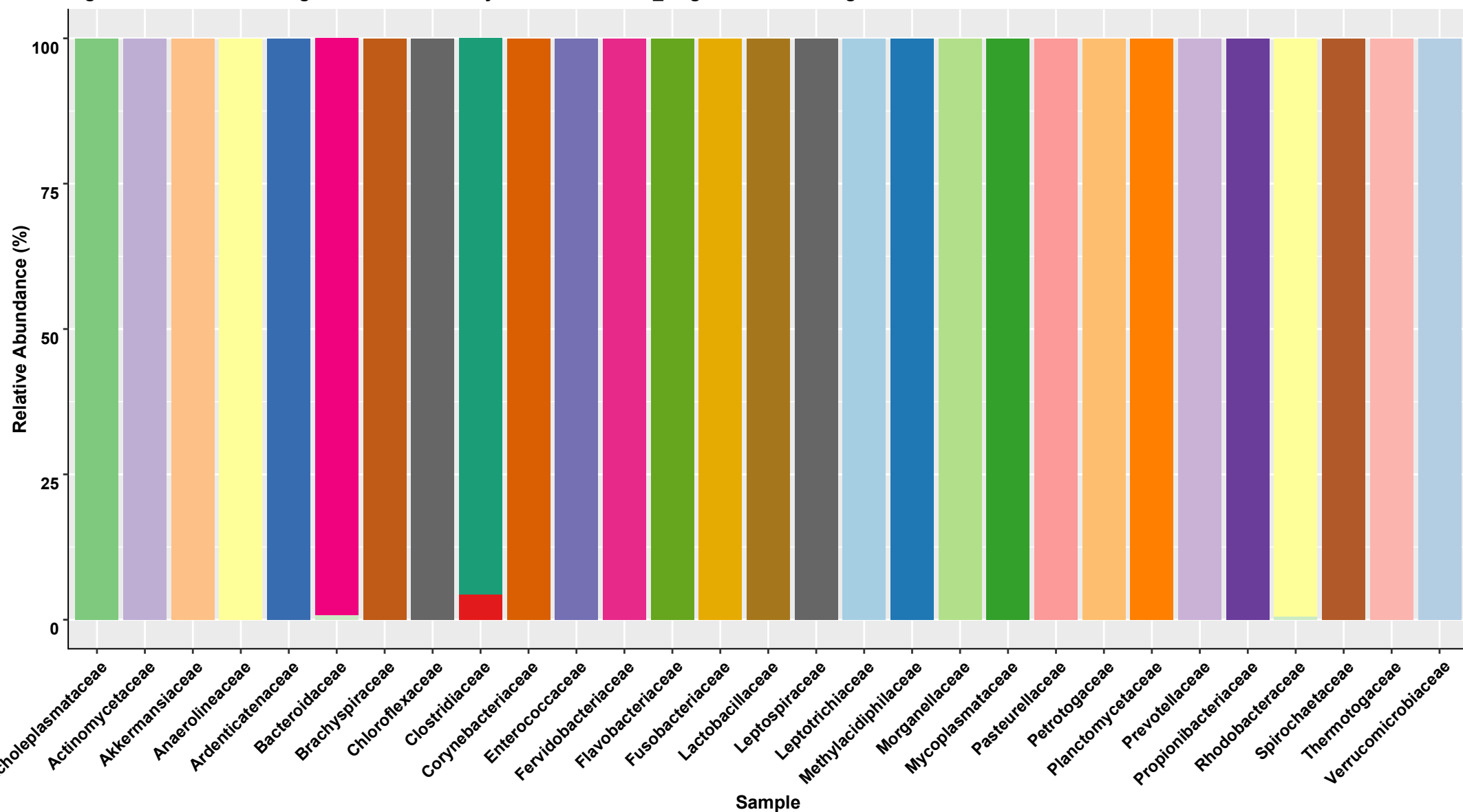


Figure F1: Taxonomic assignments for the family-level dataset “full\_length” at a clustering threshold of 80 %



### Family

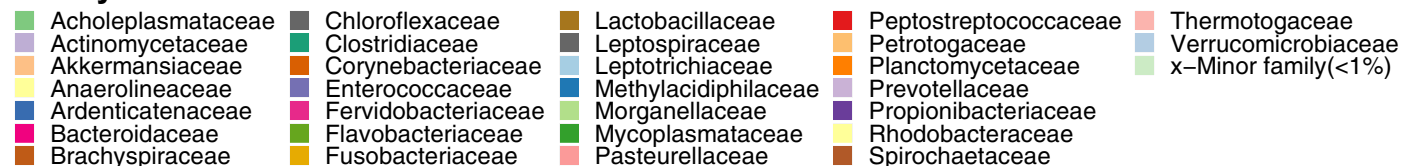
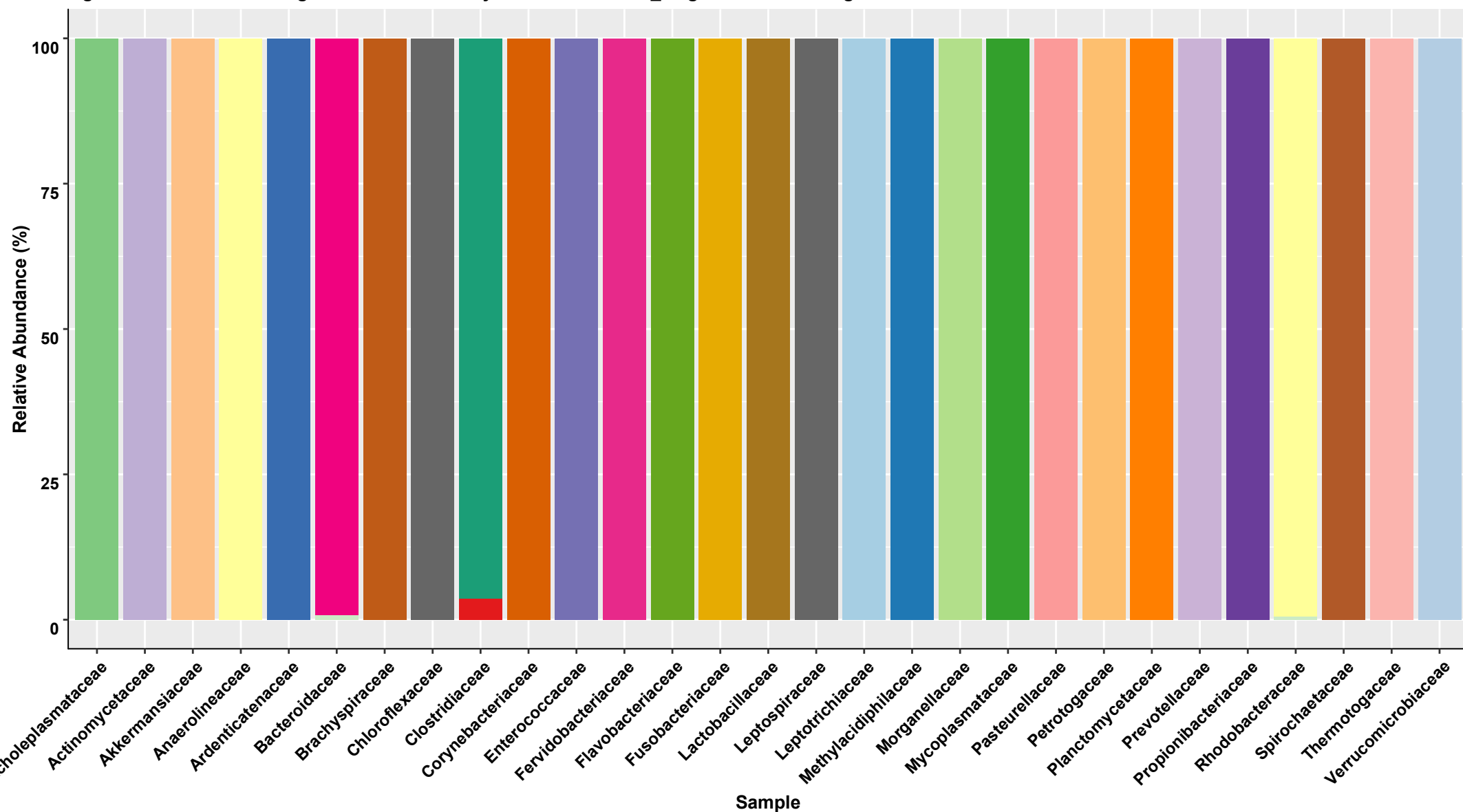


Figure F2: Taxonomic assignments for the family-level dataset “full\_length” at a clustering threshold of 100 %



### Family

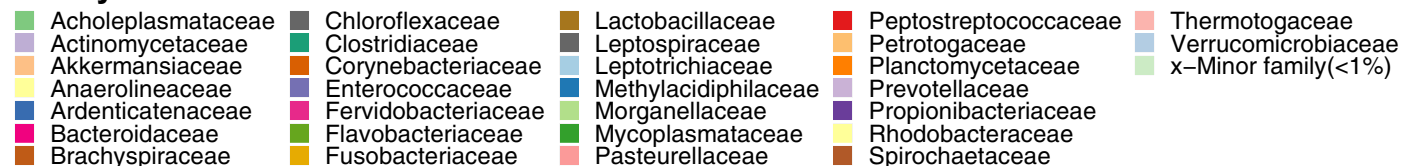
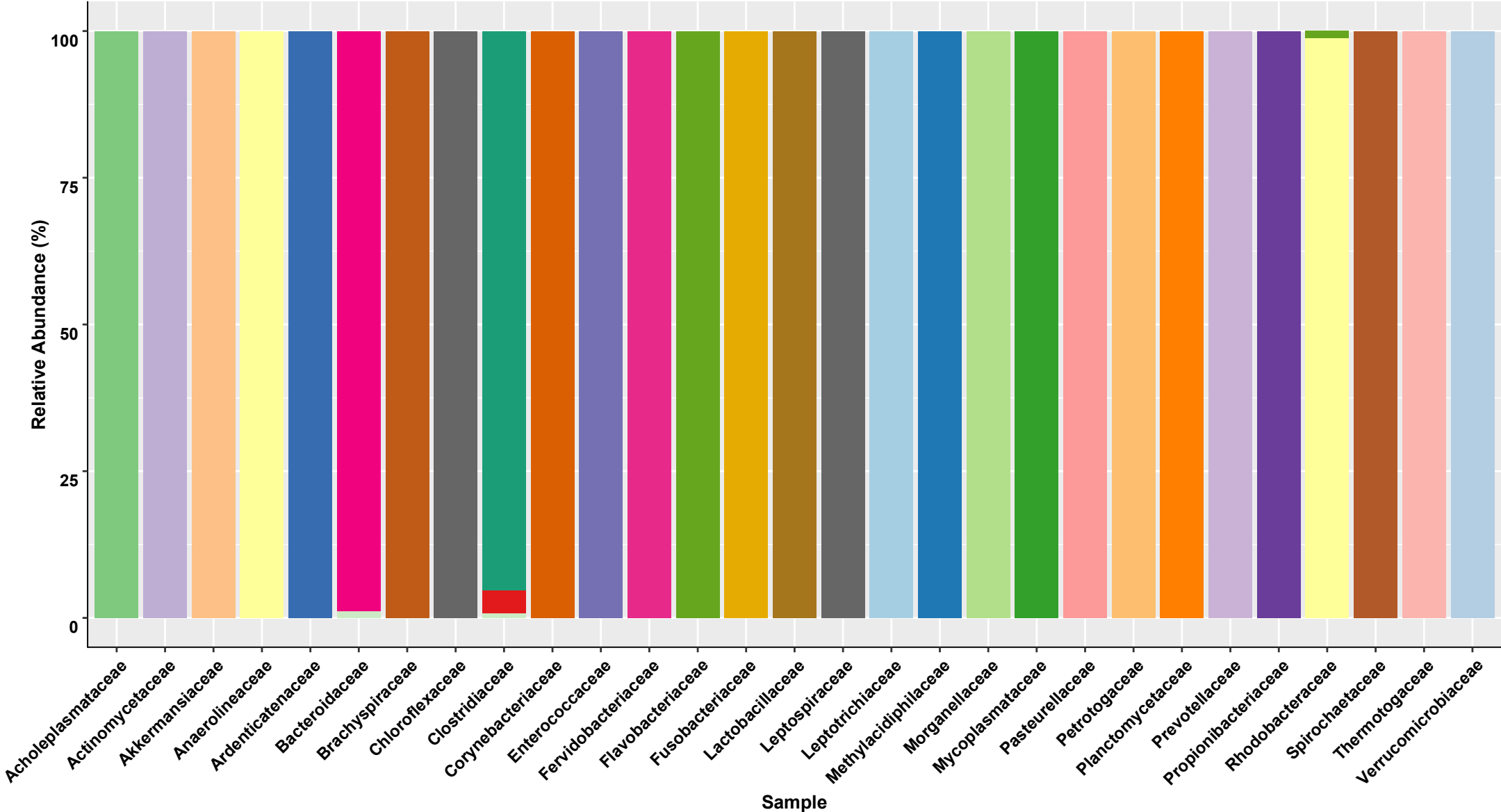


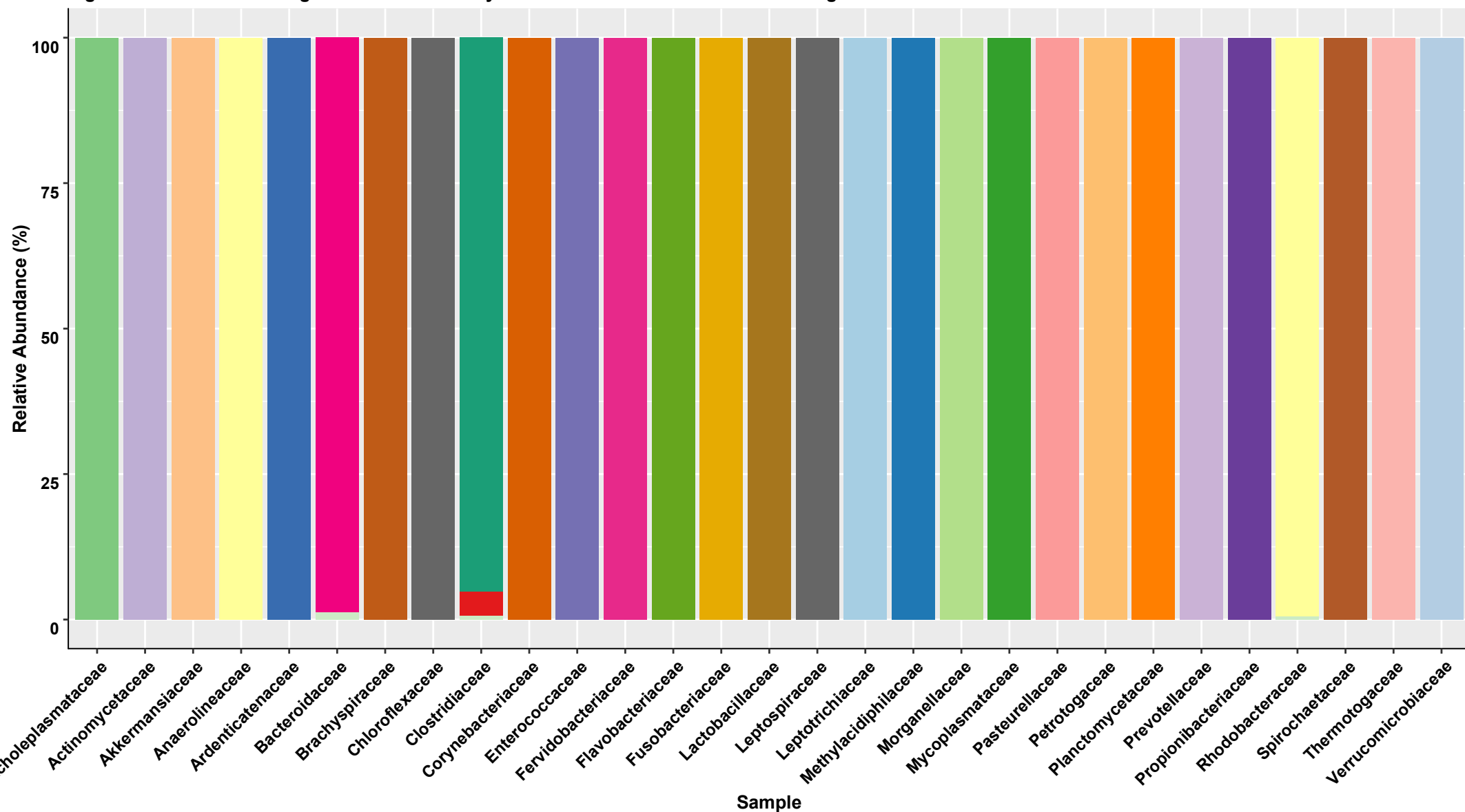
Figure F3: Taxonomic assignments for the family-level dataset “trimmed” at a clustering threshold of 80 %



Family

- |                    |                     |                  |                       |                     |
|--------------------|---------------------|------------------|-----------------------|---------------------|
| Acholeplasmataceae | Chloroflexaceae     | Lactobacillaceae | Peptostreptococcaceae | Thermotogaceae      |
| Actinomycetaceae   | Clostridiaceae      | Leptospiraceae   | Petrogorgaceae        | Verrucomicrobiaceae |
| Akkermansiaceae    | Corynebacteriaceae  | Leptotrichiaceae | Planctomycetaceae     | x-Minor family(<1%) |
| Anaerolineaceae    | Enterococcaceae     | Methylophilaceae | Prevotellaceae        |                     |
| Ardenticatenaceae  | Fervidobacteriaceae | Morganellaceae   | Propionibacteriaceae  |                     |
| Bacteroidaceae     | Flavobacteriaceae   | Mycoplasmataceae | Rhodobacteraceae      |                     |
| Brachyspiraceae    | Fusobacteriaceae    | Pasteurellaceae  | Spirochaetaceae       |                     |

Figure F4: Taxonomic assignments for the family-level dataset "trimmed" at a clustering threshold of 100 %



### Family

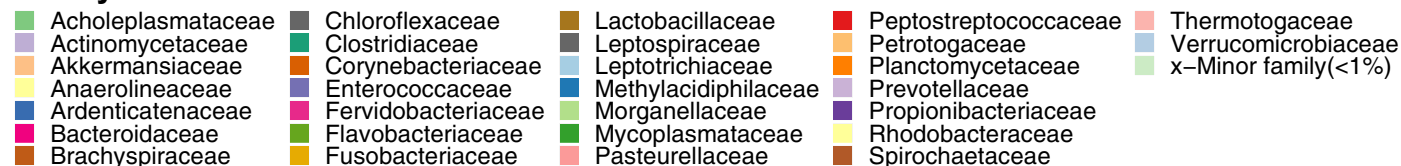
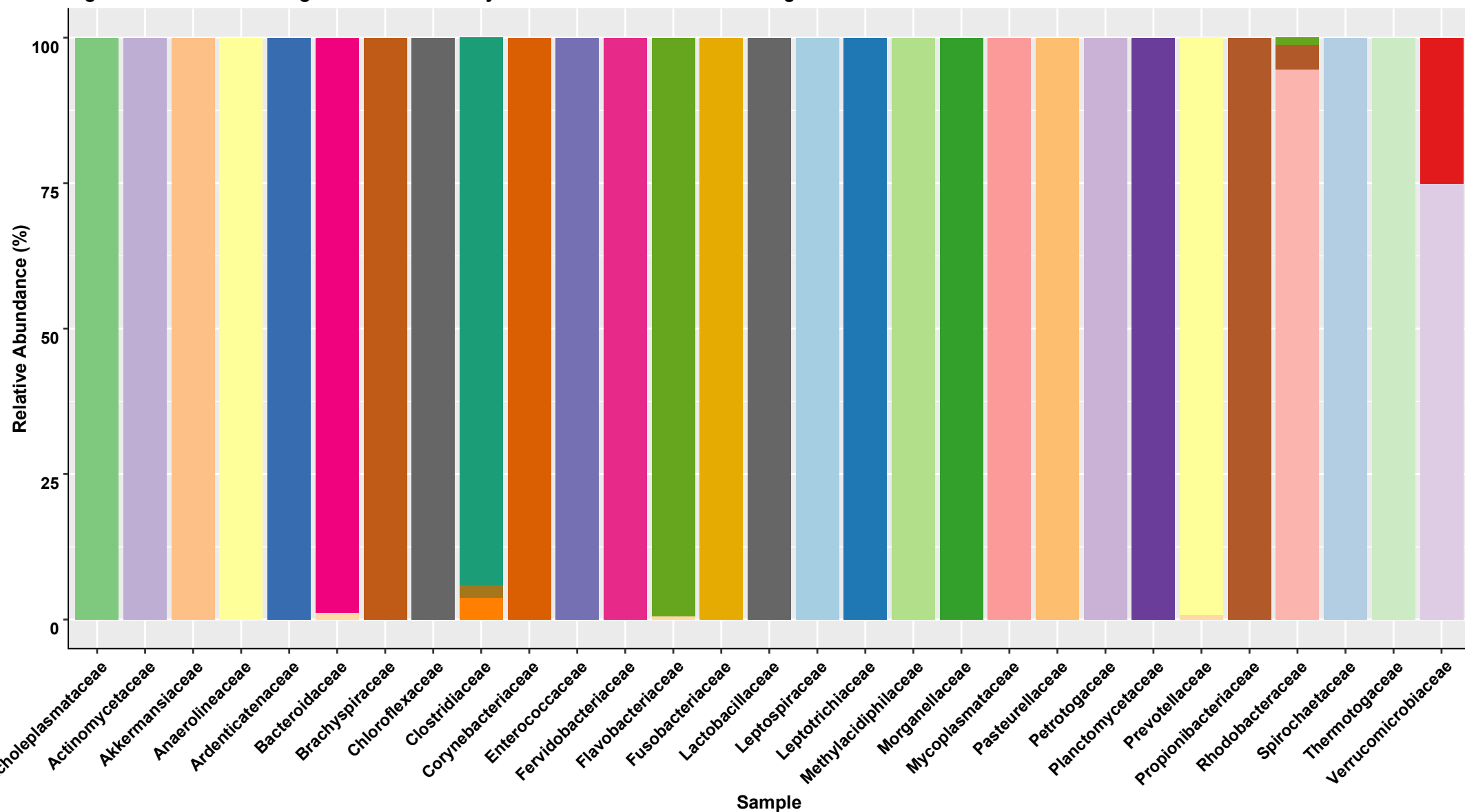


Figure F5: Taxonomic assignments for the family-level dataset "300" at a clustering threshold of 80 %



### Family

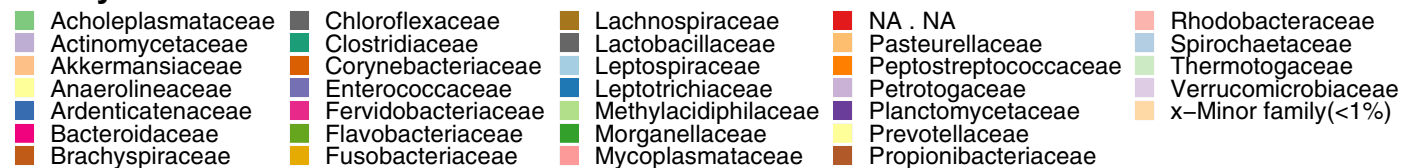
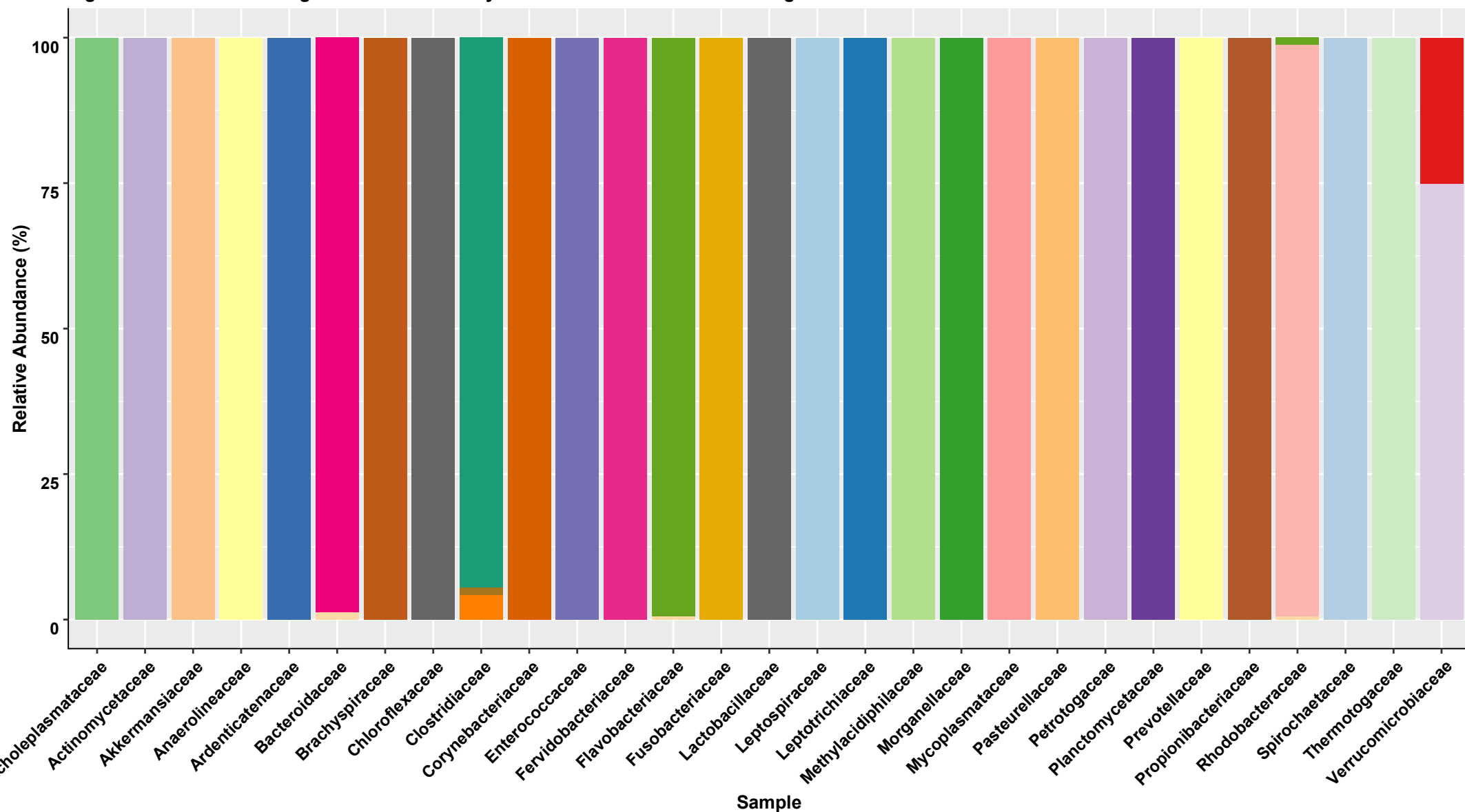


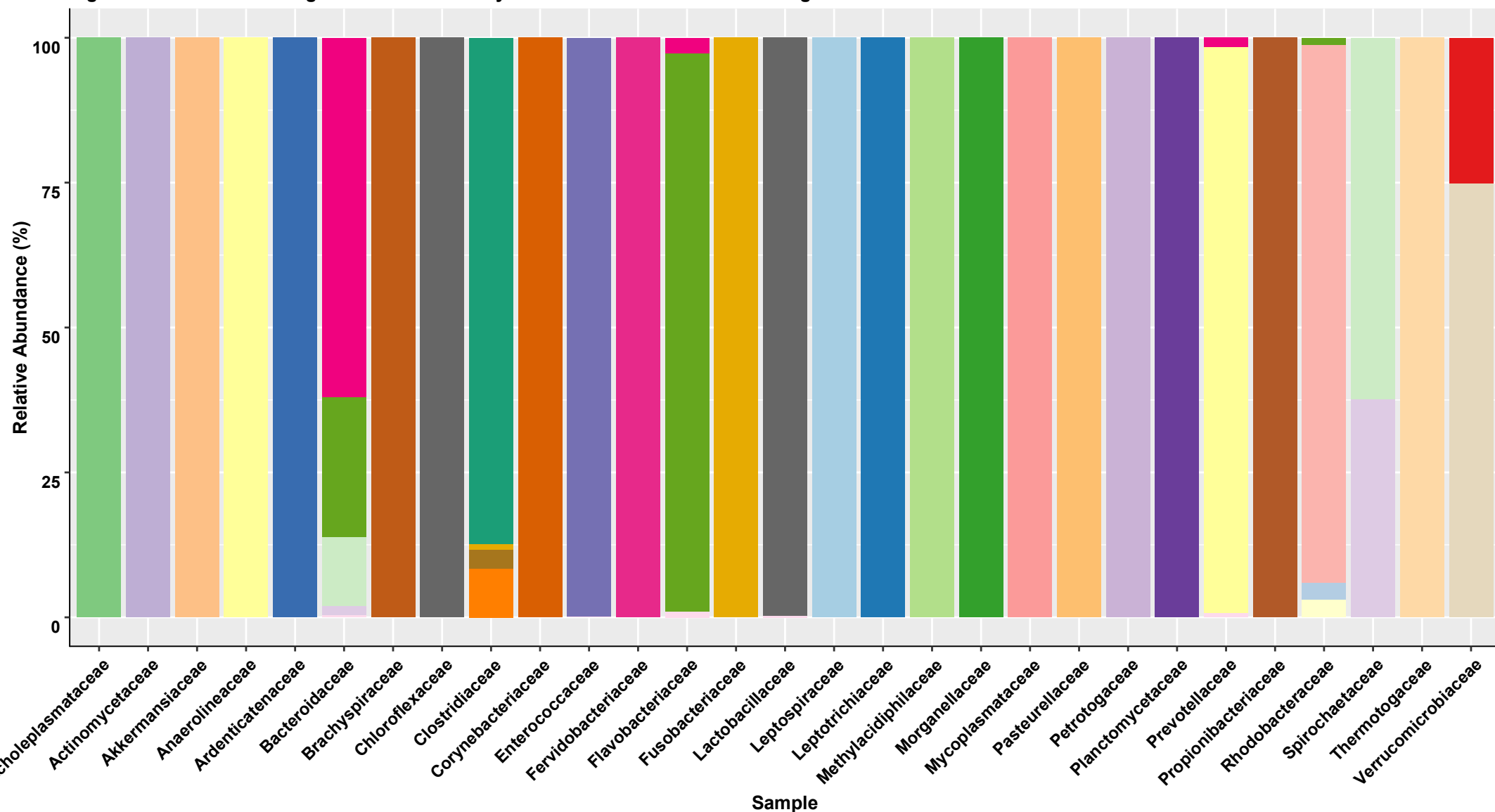
Figure F6: Taxonomic assignments for the family-level dataset "300" at a clustering threshold of 100 %



### Family

- |                    |                     |                  |                       |                     |
|--------------------|---------------------|------------------|-----------------------|---------------------|
| Acholeplasmataceae | Chloroflexaceae     | Lachnospiraceae  | NA . NA               | Rhodobacteraceae    |
| Actinomycetaceae   | Clostridiaceae      | Lactobacillaceae | Pasteurellaceae       | Spirochaetaceae     |
| Akkermansiaceae    | Corynebacteriaceae  | Leptospiraceae   | Peptostreptococcaceae | Thermotogaceae      |
| Anaerolineaceae    | Enterococcaceae     | Leptotrichiaceae | Petrogagaceae         | Verrucomicrobiaceae |
| Ardenticatenaceae  | Fervidobacteriaceae | Methylococcaceae | Planctomycetaceae     | x-Minor family(<1%) |
| Bacteroidaceae     | Flavobacteriaceae   | Morganellaceae   | Prevotellaceae        |                     |
| Brachyspiraceae    | Fusobacteriaceae    | Mycoplasmataceae | Propionibacteriaceae  |                     |

Figure F7: Taxonomic assignments for the family-level dataset "150" at a clustering threshold of 80 %



### Family

- |                    |                     |                       |                      |                              |
|--------------------|---------------------|-----------------------|----------------------|------------------------------|
| Acholeplasmataceae | Clostridiaceae      | Leptospiraceae        | Petrogaceae          | Thermotogaceae               |
| Actinomycetaceae   | Corynebacteriaceae  | Leptotrichiaceae      | Planctomycetaceae    | unclassified_Rhodobacterales |
| Akkermansiaceae    | Enterococcaceae     | Methylophilaceae      | Prevotellaceae       | Verrucomicrobiaceae          |
| Anaerolineaceae    | Fervidobacteriaceae | Morganellaceae        | Propionibacteriaceae | x-Minor family(<1%)          |
| Ardenticatenaceae  | Flavobacteriaceae   | Mycoplasmataceae      | Rhodobacteriaceae    |                              |
| Bacteroidaceae     | Fusobacteriaceae    | NA . NA               | Rhodospirillaceae    |                              |
| Brachyspiraceae    | Lachnospiraceae     | Pasteurellaceae       | Spirochaetaceae      |                              |
| Chloroflexaceae    | Lactobacillaceae    | Peptostreptococcaceae | Tannerellaceae       |                              |

Figure F8: Taxonomic assignments for the family-level dataset "150" at a clustering threshold of 100 %

