

菌株ゲノムデータ解析手順

1. Flye [v2.9.4] を用いてアセンブルを行いました。オプションは--genome-size Xm -asm-coverage 200です。(Xは各サンプルの推定ゲノムサイズです。)
2. seqkit [v2.6.1]を用いて、(1)で取得したアセンブリ配列から1000塩基未満のコンティグを削除しました。
3. Prokka [version 1.14.6] を用いて、取得したアセンブリ配列から遺伝子領域を推定しました。指定しているオプションは、--rawproduct --mincontiglen 1000 となります。
4. QUAST [v5.2.0] を用いて、取得したアセンブリ配列のコンティグ数、全長、GC含量などを評価しました。指定しているオプションは、デフォルトとなります。
5. CheckM [v1.1.3] を用いて、取得したアセンブリ配列の完全性および汚染度を評価しました。指定しているオプションは、lineage_wf -r --ali --genes --tab_table となります。
6. GTDBTk [version 2.3.2] を用いて、取得したアセンブリ配列の生物系統情報を推定しました。オプションはデフォルトとなります。生物系統推定に使用したデータベースのバージョンは、Release 214 となります。