* 마크로젠 분석 목록

I. 16S rRNA 분석

* 1. 군집 내 분류군(Taxa) 차이 분석
     + Wilcoxon Rank-Sum Test: 그룹 간 미생물 분류군 존재비 차이 검증 (p-value)

II. Shotgun Metagenoimic 분석

* 1. Read 기반 분석
     + 군집 구성 및 다양성 분석
       1. Shogun을 이용한 분류군(Taxonomic) 및 기능(Functional) 프로파일링 (Read count data)
       2. PCoA (Principal Coordinates Analysis): 샘플 간 군집 구조 비교 (베타 다양성 시각화)
     + 그룹 간 유의미한 차이 분석 (Differential Abundance)
       1. Wilcoxon Rank-Sum Test: 그룹 간 특정 분류군 또는 유전자/기능 존재비 차이 검증 (p-value, Fold Change)
       2. LEfSe (Linear discriminant analysis Effect Size): 그룹을 구분하는 주요 바이오마커(분류군/유전자/기능) 발굴 (LDA Score, p-value)
       3. KEGG Completeness Fold Change(Graphical Abstract)
     + 기능 분석 (Functional Analysis)
       1. KEGG Orthology (KO) 분석: 그룹 간 유전자 기능(KO) 차이 분석 (Wilcoxon p-value, Fold Change)
       2. KEGG Pathway Enrichment 분석: 그룹 간 유의하게 차이나는 대사 경로 분석 (p-value)
       3. Kegg Classification, eggNOG 기능 분류 (Functional Classification)
       4. KEGG Pathway Completeness(선택 MAG, Module Classs를 통한 Heatmap)
       5. Completeness Top 10 module
     + 샘플 분류 및 예측 모델링 (Machine Learning)
       1. PAM (Partitioning Around Medoids) 클러스터링: 유사한 군집 패턴을 가진 샘플 그룹핑
       2. Random Forest: 그룹 분류 모델 및 중요 변수(Feature) 확인
       3. PLS-DA (Partial Least Squares Discriminant Analysis): 그룹 분리 및 관련 변수 확인 (CLR 변환 데이터 사용)
  2. MAG 기반 분석
     + 군집 구성 및 다양성 분석
       1. Shogun을 이용한 분류군(Taxonomic) 및 기능(Functional) 프로파일링 (Read count data)
       2. PCoA (Principal Coordinates Analysis): 샘플 간 군집 구조 비교 (베타 다양성 시각화)
       3. MAG Completeness, Heterogenity
     + 그룹 간 유의미한 차이 분석 (Differential Abundance)
       1. Wilcoxon Rank-Sum Test: 그룹 간 특정 분류군 또는 유전자/기능 존재비 차이 검증 (p-value, Fold Change)
       2. KEGG Pathway Enrichment 분석: 그룹 간 유의하게 차이나는 대사 경로 분석 (p-value)
       3. KEGG Completeness Fold Change(Graphical Abstract)
     + MAG 기능 주석 및 분석 (Functional Annotation & Analysis)
       1. KEGG Pathway 분석: 각 MAG 내 대사 경로 존재 유무 및 완전성(Completeness) 분석
       2. KEGG Module Completeness: 특정 기능 모듈의 완성도 평가
* NCC 분석 목록

II. Shotgun Metagenoimic 분석

* 1. Read 기반 분석
     + Machine Learning(XGBoost, Random Forest, LightGBM, CatBoost)
     + 개별 AUC 분석(모델이 아닌 feature의 value를 score로 활용하여 AUC 계산)
     + Multivariate AUC 분석(AUC > Threshold Species) -GC(0.63), OC(0.7(
  2. MAG 기반 분석
     + 개별 Feature AUC 분석 (NCC): 단일 변수(Feature)의 그룹 예측 성능 평가 (AUC score)
  3. MAG과 기능 간의 연관성 분석
     + 선별된 MAG - KEGG Orthology 연관성 분석: 특정 MAG과 연관된 유전자 기능(KO) 관계 시각화 (Heatmap)
     + MAG - Pathway 네트워크 시각화: MAG과 주요 대사 경로 간의 관계 및 경로 내 유전자 발현 증감(Up/Down) 패턴 시각화 (Cytoscape, Pathway 노드를 Pie chart로 표현)