**질내 미생물 알고리즘**

----------------------------질내 미생물 MIrROR 분석 파이프라인-----------------------

1. **새로운 Sample sequencing data Input**
2. **Basecalling/Read correction**
3. **De-multiplexing/Adapter trimming**
4. **MIrROR process**
   1. Long read mapping
   2. Classification (Read count)
   3. Output = Mapping output file(paf 파일), Read count classification file(.txt, .std, mpa 파일)
5. **질환 별 MRS 지수 도출을 위한 taxonomy proportion 계산**

---------------------------------질환 별 MRS 계산 논리 및 프로세스-------------------------------

**참고:** <https://github.com/Gyungbu/vaginal_microbiome>

[vaginal\_percentile\_rank\_new.py](https://github.com/Gyungbu/vaginal_microbiome/blob/main/vaginal_percentile_rank_new.py) **파일**

1. **파일 load**
   1. /input/phenotype\_microbiome.xlsx – 각 질환 별 미생물 리스트가 저장된 파일
   2. /input/db\_abundance.xlsx – 91명 sample에 대한 Relative Abundance 값이 저장된 파일
   3. /input/vaginal\_mrs.xlsx – 91명 sample에 대한 각 질환 별 MRS(Microbiome Risk Score) 값이 저장된 파일
   4. /input/VALENCIA\_output\_merged.csv – 91명 sample에 대한 VALENCIA 결과 파일
2. **질환 별 유익균/유해균 비율 계산**
   1. 질환 별 유익균 비율: /input/phenotype\_microbiome.xlsx 파일의 각 질환과 연관된 미생물 리스트 중 환자군에서 많이 감소한 미생물에 해당되는 Relative abundance의 합
   2. 질환 별 유해균 비율: /input/phenotype\_microbiome.xlsx 파일의 각 질환과 연관된 미생물 리스트 중 환자군에서 많이 증가한 미생물에 해당되는 Relative abundance의 합
3. **MRS 계산**

: Microbiome Risk Score

: Relative abundance of microbiome associated with specific diseases

: The number of microbiome associated with a specific disease

: Health sign +1 for 유해/증가, -1 for 유익/감소

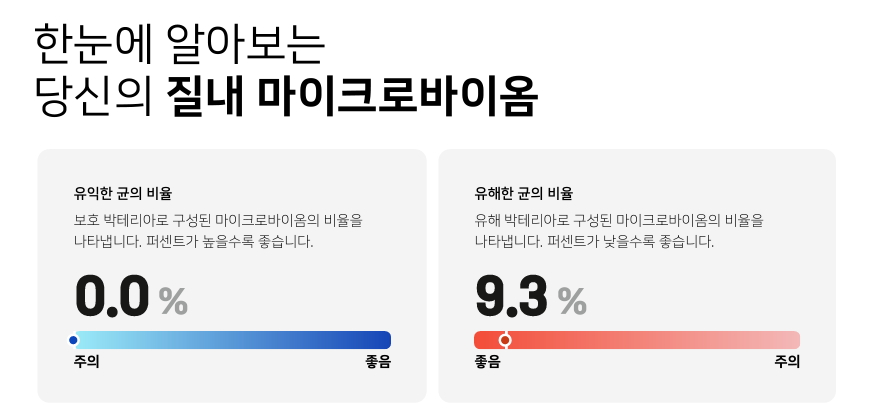
1. **최종 Percentile Rank 계산**

앞서 구한 샘플의 MRS를 기존 /input/vaginal\_mrs.xlsx 파일을 기준으로 Percentile Rank 를 계산

scipy.stats.percentileofscore를 사용.

* 1. 백분위 값이 95 이상의 outliers 95로 처리, 5 이하의 outliers 5로 처리

1. **질내 미생물 보고서**
   1. 총 유익균 유해균 비율



: Beneficial bacteria ratio

: Relative abundance of beneficial microbiome

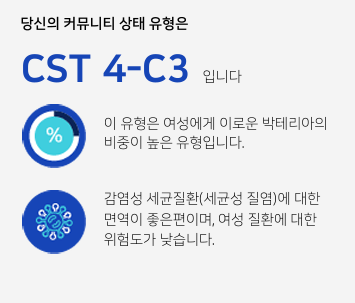
: The number of beneficial microbiome

: Harmful bacteria ratio

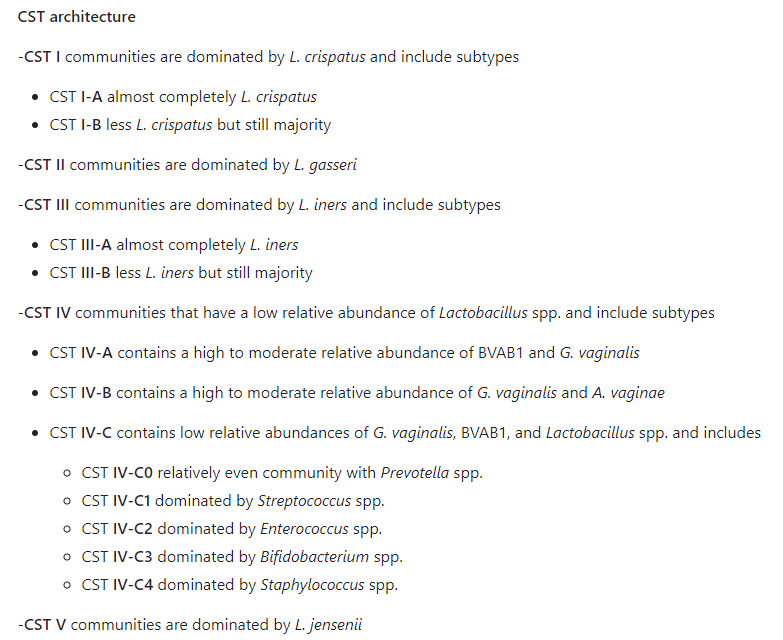
: Relative abundance of harmful microbiome

: The number of harmful microbiome

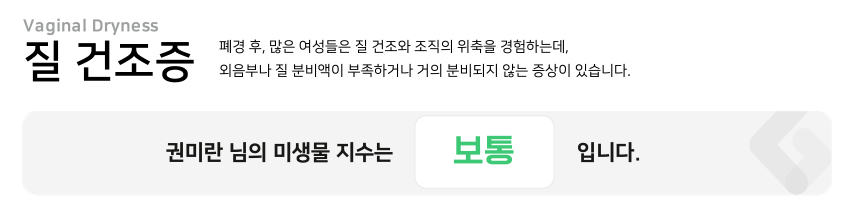
* 1. CST 타입



<https://github.com/ravel-lab/VALENCIA> 프로그램 결과



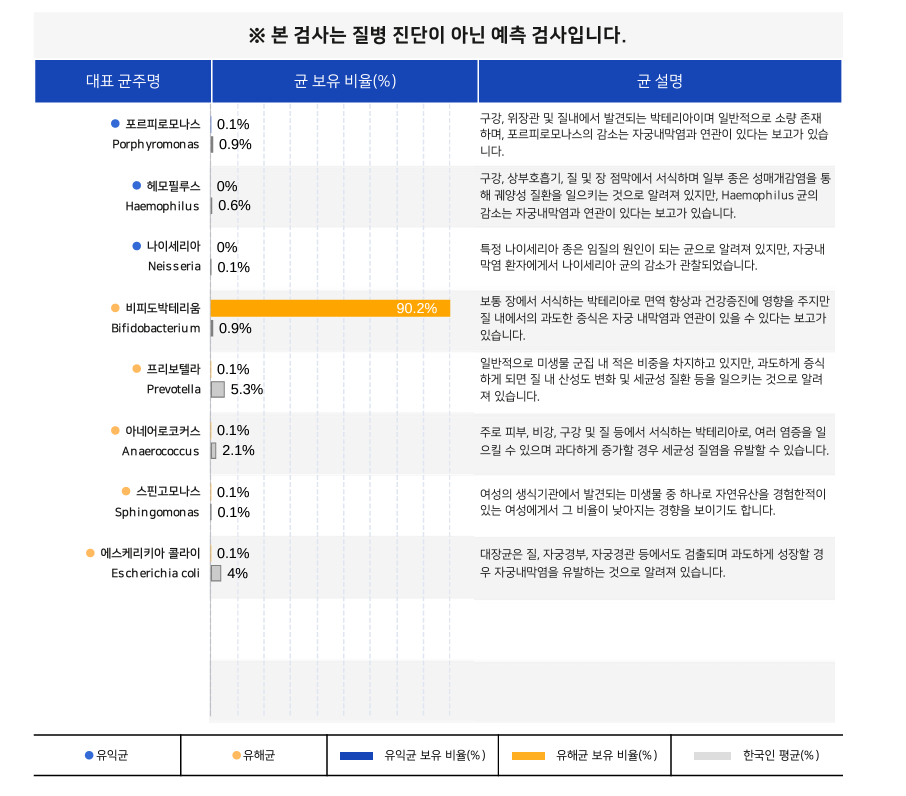
* 1. 미생물 지수



Microbiome Risk Score의 백분위 값()을 기준으로 평가를 함.

|  |  |
| --- | --- |
| 점수 기준 | 평가 항목 |
| 나쁨 | 90초과 |
| 주의 | 70초과 90이하 |
| 보통 | 50초과 70이하 |
| 양호 | 30초과 50이하 |
| 우수 | 30이하 |

* 1. 질환 별 균 비율



1. 유익균 보유 비율(%): 환자군에서 많이 감소한 미생물의 RA[%]
2. 유익균 한국인 평균(%): 환자군에서 많이 감소한 미생물에 대한 91개 샘플의 평균 RA[%]
3. 유해균 보유 비율(%): 환자군에서 많이 증가한 미생물의 RA[%]
4. 유해균 한국인 평균(%): 환자군에서 많이 증가한 미생물에 대한 91개 샘플의 평균 RA[%]

★ 유익균 유해균 리스트는 각각 최대 5개까지 표시됨

★ 유익균 유해균 리스트는 각각 유익균 보유 비율(%), 유해균 보유 비율(%)의 내림차순 순으로 표시됨