

About QMD@XYL_Lab

Development of DNA sequencing techniques have enabled our insights into composition and dynamics of complex microbial communities from human gut to soils and oceans. In practice, the microbiome data can be divided into four categories.

1. The *Microbial Load* refers to the number of a taxa in a sample.
2. The *Microbial Density* refers to the microbial loads per unit sample mass (e.g. volume or weight).
3. The *Absolute Abundance* refers to the microbial feature raw counts of samples generated from the sequencing platforms.
4. The *Relative Abundance* are compositional data and summed to a constant.

Microbial density is the most unbiased data presenting the real microbiome world. Quantification of Microbial Density Changes (QMD) is aiming for estimation of the microbial density changes between two experimental groups. Based on the result from QMD, statistical tests for differentially abundant taxa (QMDD) are constructed and p value are provided to nail the statistically significant differentially abundant taxa.

Assumption of QMD and QMDD

The microbial density changes of a considerable part of taxa are relatively small.

Mathematical foundation of QMD and QMDD

QMD model was built using L1 regularization:

$$\min: \sum_{Taxa_i} \frac{D_i^T + D_i^C}{2} |\Delta\Psi + E(\log(o_i^T)) - E(\log(o_i^C))| \quad (1)$$

Where D_i^T and D_i^C is the detection rate of $Taxa_i$ in Group K. $\Delta\Psi$ is the fold change of total microbial density of all taxa between groups. o_i^C and o_i^T are the relative Abundance of $Taxa_i$ in control group and treatment group.

With the quantified $\Delta\Psi$, the microbial density change ΔM_i for each $Taxa_i$ is calculated as

$$\Delta M_i = \Delta\Psi + E(\log(o_i^T)) - E(\log(o_i^C)) \quad (2)$$

QMDD make hypothesis H_0 : *the microbial density changed between groups* to identify differentially abundant taxa, i.e.

$$H_0: \Delta\Psi + E\left(\log(o_i^T)\right) = E\left(\log(o_i^C)\right) \quad (3)$$

After correction of $\Delta\Psi$, Mann-Whitney U Test could be adopted to conduct the statistical test in formula 3.

QMD@XYL_Lab

QMD@XYL_Lab helps researchers to conduct QMD and QMDD in differentially abundant taxa identification and different abundance estimations between experimental groups. This software was implemented by python and qt5. QMD@XYL_Lab is open source software. One can find source code of QMD@XYL_Lab at <http://www.github.com>. The CUI version of the software us also provided.

Citing QMD@XYL_Lab

If you use QMD@XYL_Lab for any published research, please include the following citation:

**QMD: Quantification of microbial density changes and its
application on microbiome differential abundance analysis**