**JASA ACS Reproducibility Initiative - Author Contributions Checklist Form**

This form is for the manuscript “IFAA: Robust association identification and Inference For Absolute Abundance in microbiome analyses”

## Data

**Abstract (Mandatory)**

There are two data sets used in this manuscript. The first one is the New Hampshire Birth Cohort Study (NHBCS) data set and the other one is the VSL#3 mouse study data set. The NHBCS study is an ongoing longitudinal human study. The VSL#3 study is a mouse model study that has been finished. Both data sets are from the projects of some of the co-authors in this manuscript and not publicly available.

**Availability (Mandatory)**

The NHBCS human study data set will not be made publicly available in the near future due to confidentiality and that the study is still collecting data. The VSL#3 mouse study data set is now available at https://github.com/gitlzg/IFAA/blob/master/data/data\_VSL3.xlsx.

## Code

**Abstract (Mandatory)**

The submitted R codes include a zipped IFAA package “IFAA\_0.0.0.9000.tar.gz”, real data analyses, and codes for running simulations. A hypothetical example of using the main function IFAA() is also included where the data was randomly generated to mimic one of the real data sets.

**Description (Mandatory)**

* How delivered: the IFAA package in R
* Licensing information: MIT License
* Link to code/repository: the IFAA package can be installed from <https://github.com/gitlzg/IFAA>
* Version information: commit 1b85b44c90662cbc37a87bb06072f0e9907690fe
* R package dependencies:

picasso (>= 1.2.0), glmnet (>= 2.0-16), expm (>= 0.999-3),

foreach (>= 1.4.3), snow (>= 0.4-2), doSNOW (>= 1.0.15),

rlecuyer (>= 0.3-3), Matrix (>= 1.2-14), HDCI (>= 1.0-2),

doParallel (>= 1.0.11), future (>= 1.12.0)

## Instructions for Use

**Reproducibility (Mandatory)**

* What is to be reproduced: the point estimates and 95% confidence intervals in the two real data applications
* How to reproduce analyses: For the NHBCS study, the analysis results can be reproduced by the submitted R file “Rcode For NHBCS study.R”. For the VSL#3 mouse study, the analysis results can be reproduced by the submitted “Rcode For VSL3 study.R”.
* Expected run-time of the workflow: The “Rcode For NHBCS study.R” will run about 73 minutes to finish the analysis on a 8-core Windows 10 machine. Most of the time is used by the IFAA() function. The “Rcode For VSL3 study.R” will run about 125 minutes on a 8-core Windows 10 machine and again most of the time is used by the IFAA() function.