An Introduction to the MetaLonDA Package

Ahmed Metwally

2018-10-28

Introduction

MetaLonDA (METAgenomic LONgitudinal Differential Abundance method) is a method that identifies significant time intervals of microbial features in longitudinal studies. MetaLonDA has the ability to handle inconsistencies and common challenges associated with human studies, such as variable sample collection times and uneven number of time points along the subjects' longitudinal study. The method employs a negative binomial distribution in conjunction with a semi-parametric SS-ANOVA to model the count reads. Then, it performs the significance testing based on unit time intervals using permutation testing procedure.

Example

An example dataset derived from a subset of stress-induced Yeast is available by running

```
library(MetaLonDA)

## Load read counts of 8 features from 100 samples. Samples are from 2 groups, 5 subjects per group, and
data(metalonda_test_data)
```

Test one feature

```
## Create Group, Time, and ID annotation vectors
n.group = 2
n.sample = 5
n.timepoints = 10
Group = factor(c(rep("A", n.sample*n.timepoints), rep("B",n.sample*n.timepoints)))
Time = rep(rep(1:n.timepoints, times = n.sample), 2)
ID = factor(rep(1:(2*n.sample), each = n.timepoints))

## Define the prediction timeponits
points = seq(1, 10, length.out = 100)
output.metalonda.f5 = metalonda(Count = metalonda_test_data[5,], Time = Time, Group = Group,
                                ID = ID, n.perm = 20, fit.method = "nbinomial", points = points,
                                text = rownames(metalonda_test_data)[5], parall = FALSE, pvalue.thresho
                                adjust.method = "BH", time.unit = "hours", ylabel = "Read Counts", col =
```

```
## Start MetaLonDA
## Visualizing Feature =  OTU_5

## Start Curve Fitting
## Fitting: NB SS
## Visualizing Splines of Feature =  OTU_5

## Calculate Area Under the Fitted Curves
## Start Permutation
##
```

```
  |
  |                                                                      |   0%
  |
  |===                                                                   |   5%
  |
  |======                                                                |  10%
  |
  |==========                                                            |  15%
  |
  |=============                                                         |  20%
  |
  |================                                                      |  25%
  |
  |====================                                                  |  30%
  |
  |========================                                              |  35%
  |
  |============================                                          |  40%
  |
  |================================                                      |  45%
  |
  |====================================                                  |  50%
  |
  |========================================                              |  55%
  |
  |============================================                          |  60%
  |
  |================================================                      |  65%
  |
  |====================================================                  |  70%
  |
  |========================================================              |  75%
  |
  |============================================================          |  80%
  |
  |================================================================      |  85%
  |
  |====================================================================  |  90%
  |
  |======================================================================|  95%
  |
  |======================================================================| 100%
## Visualizing AR Distribution for Feature =  OTU_5

## p-value Adjustment Method =  BH
## Visualizing Significant Intervals of Feature =  OTU_5

## Visualizing Volcano Plot of Feature =  OTU_5
```

In our example, we used 20 permutations just to showcase how MetaLonDA works. In real analysis, this number should be at least 1000.

Test all features

```
## Identify significant time intervals for all features:
output.metalonda.all = metalondaAll(Count = metalonda_test_data, Time = Time, Group = Group,
                                    ID = ID, n.perm = 20, fit.method = "nbinomial", num.intervals = 100
                                    parall = FALSE, pvalue.threshold = 0.05, adjust.method = "BH", time
                                    norm.method = "none", prefix = "Test", ylabel = "Read Counts", col =
```

```
## Dimensionality check passed
## Prediction Points =   [1]  1.00  1.09  1.18  1.27  1.36  1.45  1.54  1.63  1.72  1.81  1.90
##  [12]  1.99  2.08  2.17  2.26  2.35  2.44  2.53  2.62  2.71  2.80  2.89
##  [23]  2.98  3.07  3.16  3.25  3.34  3.43  3.52  3.61  3.70  3.79  3.88
##  [34]  3.97  4.06  4.15  4.24  4.33  4.42  4.51  4.60  4.69  4.78  4.87
##  [45]  4.96  5.05  5.14  5.23  5.32  5.41  5.50  5.59  5.68  5.77  5.86
##  [56]  5.95  6.04  6.13  6.22  6.31  6.40  6.49  6.58  6.67  6.76  6.85
##  [67]  6.94  7.03  7.12  7.21  7.30  7.39  7.48  7.57  7.66  7.75  7.84
##  [78]  7.93  8.02  8.11  8.20  8.29  8.38  8.47  8.56  8.65  8.74  8.83
##  [89]  8.92  9.01  9.10  9.19  9.28  9.37  9.46  9.55  9.64  9.73  9.82
## [100]  9.91 10.00
##
## Feature  =  OTU_1
## Start MetaLonDA
## Visualizing Feature =  OTU_1
```

```
## Start Curve Fitting
## Fitting: NB SS
## Visualizing Splines of Feature =  OTU_1
```

```
## Calculate Area Under the Fitted Curves
## Start Permutation
##
  |
  |                                                                      |   0%
  |
  |===                                                                   |   5%
  |
  |======                                                                |  10%
  |
  |==========                                                            |  15%
  |
  |=============                                                         |  20%
  |
  |================                                                      |  25%
  |
  |====================                                                  |  30%
  |
  |=======================                                               |  35%
  |
  |===========================                                           |  40%
  |
  |===============================                                       |  45%
  |
  |==================================                                    |  50%
  |
```

```
  |==================================                                    |  55%
  |
  |====================================                                  |  60%
  |
  |======================================                                |  65%
  |
  |========================================                              |  70%
  |
  |==========================================                            |  75%
  |
  |============================================                          |  80%
  |
  |==============================================                        |  85%
  |
  |================================================                      |  90%
  |
  |==================================================                    |  95%
  |
  |======================================================================| 100%
## Visualizing AR Distribution for Feature =  OTU_1

## p-value Adjustment Method =  BH
## Visualizing Significant Intervals of Feature =  OTU_1

## Visualizing Volcano Plot of Feature =  OTU_1

##
##
## Feature  =  OTU_2
## Start MetaLonDA
## Visualizing Feature =  OTU_2

## Start Curve Fitting
## Fitting: NB SS
## Visualizing Splines of Feature =  OTU_2

## Calculate Area Under the Fitted Curves
## Start Permutation
##
  |
  |                                                                      |   0%
  |
  |===                                                                   |   5%
  |
  |======                                                                |  10%
  |
  |==========                                                            |  15%
  |
  |=============                                                         |  20%
  |
  |================                                                      |  25%
  |
  |====================                                                  |  30%
  |
  |========================                                              |  35%
  |
```

```
  |=========================                                             |  40%
  |
  |============================                                          |  45%
  |
  |================================                                      |  50%
  |
  |===================================                                   |  55%
  |
  |=======================================                               |  60%
  |
  |==========================================                            |  65%
  |
  |==============================================                        |  70%
  |
  |==================================================                    |  75%
  |
  |======================================================                |  80%
  |
  |==========================================================            |  85%
  |
  |==============================================================        |  90%
  |
  |==================================================================    |  95%
  |
  |======================================================================| 100%
## Visualizing AR Distribution for Feature =  OTU_2

## p-value Adjustment Method =  BH
## Visualizing Significant Intervals of Feature =  OTU_2

## Visualizing Volcano Plot of Feature =  OTU_2

##
##
## Feature  =  OTU_3
## Start MetaLonDA
## Visualizing Feature =  OTU_3

## Start Curve Fitting
## Fitting: NB SS
## Visualizing Splines of Feature =  OTU_3

## Calculate Area Under the Fitted Curves
## Start Permutation
##
  |
  |                                                                      |   0%
  |
  |===                                                                   |   5%
  |
  |======                                                                |  10%
  |
  |==========                                                            |  15%
  |
  |=============                                                         |  20%
  |
```

```
  |================                                                 |  25%
  |
  |====================                                             |  30%
  |
  |========================                                         |  35%
  |
  |============================                                     |  40%
  |
  |================================                                 |  45%
  |
  |====================================                             |  50%
  |
  |========================================                         |  55%
  |
  |============================================                     |  60%
  |
  |================================================                 |  65%
  |
  |====================================================             |  70%
  |
  |========================================================         |  75%
  |
  |============================================================     |  80%
  |
  |================================================================ |  85%
  |
  |====================================================================  90%
  |
  |======================================================================== |  95%
  |
  |==========================================================================| 100%
## Visualizing AR Distribution for Feature =  OTU_3

## p-value Adjustment Method =  BH
## Visualizing Significant Intervals of Feature =  OTU_3

## Visualizing Volcano Plot of Feature =  OTU_3

##
##
## Feature  =  OTU_4
## Start MetaLonDA
## Visualizing Feature =  OTU_4

## Start Curve Fitting
## Fitting: NB SS
## Visualizing Splines of Feature =  OTU_4

## Calculate Area Under the Fitted Curves
## Start Permutation
##
  |
  |                                                                 |   0%
  |
  |===                                                              |   5%
  |
```

```
  |======                                                                | 10%
  |
  |==========                                                            | 15%
  |
  |=============                                                         | 20%
  |
  |================                                                      | 25%
  |
  |====================                                                  | 30%
  |
  |========================                                              | 35%
  |
  |============================                                          | 40%
  |
  |================================                                      | 45%
  |
  |====================================                                  | 50%
  |
  |========================================                              | 55%
  |
  |============================================                          | 60%
  |
  |================================================                      | 65%
  |
  |====================================================                  | 70%
  |
  |========================================================              | 75%
  |
  |============================================================          | 80%
  |
  |================================================================      | 85%
  |
  |====================================================================  | 90%
  |
  |======================================================================| 95%
  |
  |======================================================================| 100%
## Visualizing AR Distribution for Feature =  OTU_4

## p-value Adjustment Method =  BH
## Visualizing Significant Intervals of Feature =  OTU_4

## Visualizing Volcano Plot of Feature =  OTU_4

##
##
## Feature  =  OTU_5
## Start MetaLonDA
## Visualizing Feature =  OTU_5

## Start Curve Fitting
## Fitting: NB SS
## Visualizing Splines of Feature =  OTU_5

## Calculate Area Under the Fitted Curves
## Start Permutation
```

```
##
  |
  |                                                                      |   0%
  |
  |===                                                                   |   5%
  |
  |======                                                                |  10%
  |
  |==========                                                            |  15%
  |
  |=============                                                         |  20%
  |
  |================                                                      |  25%
  |
  |====================                                                  |  30%
  |
  |========================                                              |  35%
  |
  |============================                                          |  40%
  |
  |================================                                      |  45%
  |
  |====================================                                  |  50%
  |
  |========================================                              |  55%
  |
  |============================================                          |  60%
  |
  |================================================                      |  65%
  |
  |====================================================                  |  70%
  |
  |========================================================              |  75%
  |
  |============================================================          |  80%
  |
  |================================================================      |  85%
  |
  |====================================================================  |  90%
  |
  |======================================================================|  95%
  |
  |======================================================================| 100%
## Visualizing AR Distribution for Feature =  OTU_5

## p-value Adjustment Method =  BH
## Visualizing Significant Intervals of Feature =  OTU_5

## Visualizing Volcano Plot of Feature =  OTU_5

##
##
## Feature  =  OTU_6
## Start MetaLonDA
## Visualizing Feature =  OTU_6
```

```
## Start Curve Fitting
## Fitting: NB SS
## Visualizing Splines of Feature =  OTU_6
## Calculate Area Under the Fitted Curves
## Start Permutation
##
  |
  |                                                                      |   0%
  |
  |===                                                                   |   5%
  |
  |======                                                                |  10%
  |
  |==========                                                            |  15%
  |
  |=============                                                         |  20%
  |
  |================                                                      |  25%
  |
  |====================                                                  |  30%
  |
  |=======================                                               |  35%
  |
  |===========================                                           |  40%
  |
  |===============================                                       |  45%
  |
  |===================================                                   |  50%
  |
  |=======================================                               |  55%
  |
  |===========================================                           |  60%
  |
  |===============================================                       |  65%
  |
  |===================================================                   |  70%
  |
  |=======================================================               |  75%
  |
  |===========================================================           |  80%
  |
  |===============================================================       |  85%
  |
  |===================================================================   |  90%
  |
  |=======================================================================|  95%
  |
  |=======================================================================| 100%
## Visualizing AR Distribution for Feature =  OTU_6
## p-value Adjustment Method =  BH
## Visualizing Significant Intervals of Feature =  OTU_6
## Visualizing Volcano Plot of Feature =  OTU_6
```

```
##
##
## Feature  =  OTU_7
## Start MetaLonDA
## Visualizing Feature =  OTU_7

## Start Curve Fitting
## Fitting: NB SS
## Visualizing Splines of Feature =  OTU_7

## Calculate Area Under the Fitted Curves
## Start Permutation
##
  |
  |                                                                      |   0%
  |
  |===                                                                   |   5%
  |
  |======                                                                |  10%
  |
  |==========                                                            |  15%
  |
  |=============                                                         |  20%
  |
  |================                                                      |  25%
  |
  |====================                                                  |  30%
  |
  |========================                                              |  35%
  |
  |============================                                          |  40%
  |
  |================================                                      |  45%
  |
  |====================================                                  |  50%
  |
  |========================================                              |  55%
  |
  |============================================                          |  60%
  |
  |================================================                      |  65%
  |
  |====================================================                  |  70%
  |
  |========================================================              |  75%
  |
  |============================================================          |  80%
  |
  |================================================================      |  85%
  |
  |====================================================================  |  90%
  |
  |======================================================================|  95%
  |
  |======================================================================| 100%
```

```
## Visualizing AR Distribution for Feature =  OTU_7
## p-value Adjustment Method =  BH
## No Significant Intevals Found
## Visualizing Volcano Plot of Feature =  OTU_7
##
##
## Feature  =  OTU_8
## Start MetaLonDA
## Visualizing Feature =  OTU_8
## Start Curve Fitting
## Fitting: NB SS
## Visualizing Splines of Feature =  OTU_8
## Calculate Area Under the Fitted Curves
## Start Permutation
##
  |
  |                                                                      |   0%
  |
  |===                                                                   |   5%
  |
  |======                                                                |  10%
  |
  |==========                                                            |  15%
  |
  |=============                                                         |  20%
  |
  |================                                                      |  25%
  |
  |====================                                                  |  30%
  |
  |=======================                                               |  35%
  |
  |===========================                                           |  40%
  |
  |===============================                                       |  45%
  |
  |===================================                                   |  50%
  |
  |======================================                                |  55%
  |
  |==========================================                            |  60%
  |
  |==============================================                        |  65%
  |
  |==================================================                    |  70%
  |
  |======================================================                |  75%
  |
  |==========================================================            |  80%
  |
  |==============================================================        |  85%
  |
```

```
  |=============================================================       |  90%
  |
  |=================================================================   |  95%
  |
  |=====================================================================| 100%
## Visualizing AR Distribution for Feature =  OTU_8

## p-value Adjustment Method =  BH
## No Significant Intevals Found
## Visualizing Volcano Plot of Feature =  OTU_8
```

Session information

```
sessionInfo()
```

```
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-apple-darwin13.4.0 (64-bit)
## Running under: macOS Sierra 10.12.4
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] parallel  stats     graphics  grDevices utils     datasets  methods
## [8] base
##
## other attached packages:
##  [1] MetaLonDA_1.1.2   rmarkdown_1.8     zoo_1.8-1
##  [4] doParallel_1.0.11 iterators_1.0.9   foreach_1.4.4
##  [7] pracma_1.9.9      plyr_1.8.4        gss_2.1-7
## [10] ggplot2_2.2.1
##
## loaded via a namespace (and not attached):
##  [1] Biobase_2.34.0             edgeR_3.16.5
##  [3] splines_3.3.2              gtools_3.5.0
##  [5] assertthat_0.2.0           Formula_1.2-1
##  [7] stats4_3.3.2               latticeExtra_0.6-28
##  [9] yaml_2.1.16                RSQLite_1.1-2
## [11] backports_1.1.2            lattice_0.20-34
## [13] limma_3.30.11              quadprog_1.5-5
## [15] digest_0.6.13              GenomicRanges_1.26.3
## [17] RColorBrewer_1.1-2         XVector_0.14.0
## [19] checkmate_1.8.2            colorspace_1.3-2
## [21] htmltools_0.3.6            Matrix_1.2-7.1
## [23] XML_3.98-1.5               DESeq2_1.14.1
## [25] devtools_1.12.0            genefilter_1.56.0
## [27] zlibbioc_1.20.0            xtable_1.8-2
## [29] scales_0.5.0               gdata_2.17.0
## [31] BiocParallel_1.8.1         tibble_1.3.3
## [33] htmlTable_1.9              annotate_1.52.1
## [35] IRanges_2.8.1              withr_1.0.2
## [37] SummarizedExperiment_1.4.0 nnet_7.3-12
## [39] BiocGenerics_0.20.0        lazyeval_0.2.0
```

```
## [41] crayon_1.3.2              survival_2.40-1
## [43] magrittr_1.5              evaluate_0.10.1
## [45] memoise_1.0.0             gplots_3.0.1
## [47] xml2_1.1.1                foreign_0.8-67
## [49] tools_3.3.2               data.table_1.10.0
## [51] matrixStats_0.51.0        stringr_1.2.0
## [53] S4Vectors_0.12.1          locfit_1.5-9.1
## [55] munsell_0.4.3             glmnet_2.0-5
## [57] cluster_2.0.5             AnnotationDbi_1.36.2
## [59] GenomeInfoDb_1.10.3       caTools_1.17.1
## [61] rlang_0.1.6               grid_3.3.2
## [63] RCurl_1.95-4.8            htmlwidgets_0.8
## [65] bitops_1.0-6              base64enc_0.1-3
## [67] labeling_0.3              metagenomeSeq_1.16.0
## [69] gtable_0.2.0              codetools_0.2-15
## [71] DBI_0.5-1                 roxygen2_6.0.1
## [73] R6_2.2.2                  gridExtra_2.2.1
## [75] knitr_1.18                rprojroot_1.3-2
## [77] commonmark_1.2            Hmisc_4.0-2
## [79] desc_1.1.0                KernSmooth_2.23-15
## [81] stringi_1.1.2             Rcpp_0.12.14
## [83] geneplotter_1.52.0        rpart_4.1-10
## [85] acepack_1.4.1
```

References

Metwally, Ahmed A., Jie Yang, Christian Ascoli, Yang Dai, Patricia W. Finn, and David L. Perkins. "MetaLonDA: a flexible R package for identifying time intervals of differentially abundant features in metagenomic longitudinal studies", *Microbiome*, 2018.

Metwally, Ahmed A., Patricia W. Finn, Yang Dai, and David L. Perkins. "Detection of Differential Abundance Intervals in Longitudinal Metagenomic Data Using Negative Binomial Smoothing Spline ANOVA." *ACM BCB*, 2017.

Bugs and Suggestions

MetaLonDA is under active research development. Please report any bugs/suggestions to Ahmed Metwally (ametwall@stanford.edu).