Analyze omics from SLIMM-T2D trial

jd

2019-03-31

# Read data and take the difference from baseline per person

Clinical metadata at [clin\_metadata.csv](./data/clin_metadata.csv). This file also used for uploads to GEO. Here, change colnames in workflow, so can use them more easily in ggplot2. After taking differences of phenotypes, we remove individual points from R object *diff.pheno*, such as Topaz\_3, who have NA in every measure.

The original SOMAscan dataset was parsed to write [soma\_annot.csv](./data/soma_annot.csv) and log2-transformed to write [soma\_mat.csv](./data/soma_mat.csv).

The original metabolomics dataset was parsed, filtered, and log2-transformed to write [met\_mat\_norm.csv](./data/met_mat_norm.csv). Annotations for metabolites did not originally have CHEBI annotations, but these were needed for Pathaway analysis via network-smoothing to match our network (Pathway Commons PC9) and pathway database (SMPDB), so we added these automatically using CTSgetR. However, CTSgetR wasn’t given the chirality of many compounds, so we found that 3-hydroxyisobutyrate had CHEBI:11805, whereas SMPDB had (S)-3-Hydroxyisobutyric acid (CHEBI:37373), so we manually fixed 3-hydroxyisobutyrate to have CHEBI:37373, and wrote [met\_annot\_with\_chebi.csv](./data/met_annot_with_chebi.csv).

Samples in 12mo metabolomics and proteomics are identical, so create diff\_mats at 1 year. Also, create a combined annotation.

Calculate metadata differences from baseline and remove samples with too many NAs, who shouldn’t be used in Hitman.

# Differential abundance between groups

Analyze differences between groups at each time point, and these differences after taking each person’s difference from baseline, with and without adjusting for each person’s change in BMI.

## Proteome

The Soma supp table with means per group per time point and statistics on differences between groups on differences from baseline is at [soma\_supp\_table.csv](./results/soma_supp_table.csv).

A table representing a Venn diagram of proteomics using FDR = 0.05 is at [soma\_FDR05\_venn.csv](./results/soma_FDR05_venn.csv). A table representing a heatmap of the logFC of each of the proteins that ever has FDR < 0.05 without BMI adjustment over time is at [soma\_heat.csv](./results/soma_heat.csv).

## Metabolome

Analyze differences between groups at each time point, and these differences after taking each person’s difference from baseline. The met supp table with means per group per time point and statistics on differences between groups on differences from baseline is at [met\_supp\_table.csv](./results/met_supp_table.csv).

A table representing a Venn diagram of metabolomics using FDR = 0.05 is at [met\_FDR05\_venn.csv](./results/met_FDR05_venn.csv). A table representing a heatmap of the logFC of each of the metabolites that ever has FDR < 0.05 without BMI adjustment over time is at [met\_heat.csv](./results/met_heat.csv).

# Hitman/Lotman

We apply Hitman/Lotman to 12mo outcome change from baseline using change at 3mo from baseline as mediators and arm as exposure. The results of these mediation tests are written as CSVs to [results](./results/) folder using the naming convention outcome-timepoint\_vs\_mediator-timepoint\_hitman, such as “HbA1c12\_vs\_soma3\_hitman.csv,” whose *EMY* columns hold the overall p-values and FDRs.

## HbA1c

We apply Hitman/Lotman to 12mo HbA1c change using change at 3mo from baseline in analytes and clinical parameters as mediators.

### Compare clinical vs analytes

The protein analytes more significant in p-value & FDR than any clinical mediator are:

|  |  |  |  |
| --- | --- | --- | --- |
| probe | ratio3 | nm | EntrezGeneSymbol |
| SL005168 | -0.163 | Growth hormone receptor | GHR |
| SL006694 | -0.372 | Beta-Ala-His dipeptidase | CNDP1 |
| SL010375 | -0.485 | Neutral ceramidase | ASAH2 |
| SL000573 | -0.184 | Serum amyloid P-component | APCS |

.

The metabolite mediators are pro-hydroxy-pro

## HOMA-IR

We apply Hitman to 12mo HOMA-IR change using 3mo analyte change as mediators.

## dins030

We apply Hitman to dins030 [change in insulin from 0 to 30 min in mixed-meal tolerance test (mcU/ml)] change at 12mo from baseline using 3mo analyte change from baseline as mediators.

# Correlate soma vs. met

We correlate top somalogic proteins with top metabolites.

# Pants

For a network integration analysis of differences, we use Pathway Commons PC9 and SMPDB files downloaded around 2016.

We map ChEBI IDs in PC9 to chemical names, so that we can show the names in our network plots.

For input data to Pants, differences from baseline are combined for people who have both somascan & metabolomics. Analytes are subset and summarized (by averaging over analytes with the same ID) to match network.

## Pants between-group

We write the z-score for the between-arm comparison of each analyte’s 3 month difference from baseline from ezlimma::limma\_contrasts to use in results.

We executed the chunk below on 2019-02-17 on a Linux cluster parallelized over 20 cores, and it took only several minutes. Here we execute it on a Windows machine with 8 cores with less RAM, which takes much longer. You can see how many cores you have available with parallel::detectCores().

We plot the top pathways shown in the paper. The t-statistics from above have sufficiently many degrees of freedom to be approximately z-scores, which are better known, so we annotate them as z-scores.

## Pants Hitman homair

We write the hitman results for each analyte’s 3mo change from baseline as a mediator of 12mo HOMA-IR change.

# Session info

## - Session info ----------------------------------------------------------  
## setting value   
## version R version 3.5.3 (2019-03-11)  
## os Windows 7 x64 SP 1   
## system x86\_64, mingw32   
## ui RTerm   
## language (EN)   
## collate English\_United States.1252   
## ctype English\_United States.1252   
## tz America/New\_York   
## date 2019-03-31   
##   
## - Packages --------------------------------------------------------------  
## package \* version date lib source   
## assertthat 0.2.1 2019-03-21 [1] CRAN (R 3.5.3)   
## backports 1.1.3 2018-12-14 [1] CRAN (R 3.5.2)   
## callr 3.2.0 2019-03-15 [1] CRAN (R 3.5.3)   
## cli 1.1.0 2019-03-19 [1] CRAN (R 3.5.3)   
## colorspace 1.4-1 2019-03-18 [1] CRAN (R 3.5.3)   
## crayon 1.3.4 2017-09-16 [1] CRAN (R 3.5.3)   
## desc 1.2.0 2018-05-01 [1] CRAN (R 3.5.3)   
## devtools \* 2.0.1 2018-10-26 [1] CRAN (R 3.5.3)   
## digest 0.6.18 2018-10-10 [1] CRAN (R 3.5.3)   
## dplyr 0.8.0.1 2019-02-15 [1] CRAN (R 3.5.3)   
## evaluate 0.13 2019-02-12 [1] CRAN (R 3.5.3)   
## ezlimma \* 0.2.3.9003 2019-03-28 [1] Github (jdreyf/ezlimma@cc62771)  
## ezlimmaplot \* 0.0.1.9007 2019-03-28 [1] local   
## ezlimmautils \* 0.0.0.9001 2019-03-31 [1] local   
## fs 1.2.7 2019-03-19 [1] CRAN (R 3.5.3)   
## ggplot2 \* 3.1.0 2018-10-25 [1] CRAN (R 3.5.3)   
## glue 1.3.1 2019-03-12 [1] CRAN (R 3.5.3)   
## gtable 0.3.0 2019-03-25 [1] CRAN (R 3.5.3)   
## highr 0.8 2019-03-20 [1] CRAN (R 3.5.3)   
## htmltools 0.3.6 2017-04-28 [1] CRAN (R 3.5.3)   
## igraph 1.2.4 2019-02-13 [1] CRAN (R 3.5.3)   
## knitr \* 1.22 2019-03-08 [1] CRAN (R 3.5.3)   
## labeling 0.3 2014-08-23 [1] CRAN (R 3.5.2)   
## lattice 0.20-38 2018-11-04 [1] CRAN (R 3.5.3)   
## lazyeval 0.2.2 2019-03-15 [1] CRAN (R 3.5.3)   
## limma \* 3.38.3 2018-12-02 [1] Bioconductor   
## magrittr 1.5 2014-11-22 [1] CRAN (R 3.5.3)   
## Matrix 1.2-17 2019-03-22 [1] CRAN (R 3.5.3)   
## memoise 1.1.0 2017-04-21 [1] CRAN (R 3.5.3)   
## munsell 0.5.0 2018-06-12 [1] CRAN (R 3.5.3)   
## PANTS \* 0.0.3.9022 2019-03-30 [1] local   
## pillar 1.3.1 2018-12-15 [1] CRAN (R 3.5.3)   
## pkgbuild 1.0.3 2019-03-20 [1] CRAN (R 3.5.3)   
## pkgconfig 2.0.2 2018-08-16 [1] CRAN (R 3.5.3)   
## pkgload 1.0.2 2018-10-29 [1] CRAN (R 3.5.3)   
## plyr 1.8.4 2016-06-08 [1] CRAN (R 3.5.3)   
## prettyunits 1.0.2 2015-07-13 [1] CRAN (R 3.5.3)   
## processx 3.3.0 2019-03-10 [1] CRAN (R 3.5.3)   
## ps 1.3.0 2018-12-21 [1] CRAN (R 3.5.3)   
## purrr 0.3.2 2019-03-15 [1] CRAN (R 3.5.3)   
## R6 2.4.0 2019-02-14 [1] CRAN (R 3.5.3)   
## RColorBrewer 1.1-2 2014-12-07 [1] CRAN (R 3.5.2)   
## Rcpp 1.0.1 2019-03-17 [1] CRAN (R 3.5.3)   
## remotes 2.0.2 2018-10-30 [1] CRAN (R 3.5.3)   
## rlang 0.3.2 2019-03-21 [1] CRAN (R 3.5.3)   
## rmarkdown 1.12 2019-03-14 [1] CRAN (R 3.5.3)   
## rprojroot 1.3-2 2018-01-03 [1] CRAN (R 3.5.3)   
## scales 1.0.0 2018-08-09 [1] CRAN (R 3.5.3)   
## sessioninfo 1.1.1 2018-11-05 [1] CRAN (R 3.5.3)   
## stringi 1.4.3 2019-03-12 [1] CRAN (R 3.5.3)   
## stringr 1.4.0 2019-02-10 [1] CRAN (R 3.5.3)   
## testthat 2.0.1 2018-10-13 [1] CRAN (R 3.5.3)   
## tibble 2.1.1 2019-03-16 [1] CRAN (R 3.5.3)   
## tidyselect 0.2.5 2018-10-11 [1] CRAN (R 3.5.3)   
## usethis \* 1.4.0 2018-08-14 [1] CRAN (R 3.5.3)   
## withr 2.1.2 2018-03-15 [1] CRAN (R 3.5.3)   
## xfun 0.5 2019-02-20 [1] CRAN (R 3.5.3)   
## yaml 2.2.0 2018-07-25 [1] CRAN (R 3.5.2)   
## zeallot \* 0.1.0 2018-01-28 [1] CRAN (R 3.5.3)   
##   
## [1] C:/Program Files/R/R-3.5.3/library