Analyze omics from SLIMM-T2D trial

jd

2019-10-02

# 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 = 0.15 is at [soma\_FDR15\_venn.csv](./results/soma_FDR15_venn.csv).

A heatmap of differential proteins is at [soma\_heat.pdf](./results/soma_heat.pdf).

### Text

We read out that the number of proteins significant:  
\* At baseline: 0  
\* At 3 mo in unadj: 14; whereas in adj: 8. In both, we have

|  |  |  |  |
| --- | --- | --- | --- |
| RYGBvsDWMat3mo.sig | RYGBvsDWMat3mo\_bmi.sig | EntrezGeneSymbol | TargetFullName |
| 1 | 1 | IBSP | Bone sialoprotein 2 |
| 1 | 1 | IGFBP2 | Insulin-like growth factor-binding protein 2 |
| 1 | 1 | ESM1 | Endothelial cell-specific molecule 1 |
| 1 | 1 | MMP12 | Macrophage metalloelastase |
| -1 | -1 | CNDP1 | Beta-Ala-His dipeptidase |
| 1 | 1 | SERPINA3 | Alpha-1-antichymotrypsin complex |
| 1 | 1 | CCL22 | C-C motif chemokine 22 |
| -1 | -1 | FETUB | Fetuin-B |

* After baseline correction at any time point without BMI adjustment: 56

## 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 = 0.15 is at [met\_FDR15\_venn.csv](./results/met_FDR15_venn.csv).

A heatmap of the 62 metabolites differential at FDR of 5% (b/c too many at 15%) is at [met\_heat.pdf](./results/met_heat.pdf).

### Text

We read out that the number of mets significant:  
\* At baseline: 0  
\* At 3 mo in unadj: 96; whereas in adj: 74. In both, we have 72:

|  |  |  |
| --- | --- | --- |
| RYGBvsDWMat3mo.sig | RYGBvsDWMat3mo\_bmi.sig | BIOCHEMICAL |
| 1 | 1 | pro-hydroxy-pro |
| -1 | -1 | choline |
| -1 | -1 | stearidonate (18:4n3) |
| 1 | 1 | indolin-2-one |
| -1 | -1 | xanthurenate |
| -1 | -1 | isovalerylcarnitine |
| -1 | -1 | kynurenate |
| -1 | -1 | gamma-CEHC |
| -1 | -1 | gamma-glutamylleucine |
| -1 | -1 | gamma-CEHC glucuronide\* |
| 1 | 1 | trimethylamine N-oxide |
| 1 | 1 | phenylacetylglutamate |
| 1 | 1 | phenylacetylglutamine |
| -1 | -1 | gamma-glutamylvaline |
| -1 | -1 | eicosapentaenoate (EPA; 20:5n3) |
| -1 | -1 | valine |
| -1 | -1 | 3-methyl-2-oxobutyrate |
| -1 | -1 | leucine |
| 1 | 1 | 2-hydroxydecanoate |
| -1 | -1 | N-acetyl-beta-alanine |
| -1 | -1 | 3-hydroxyisobutyrate |
| -1 | -1 | retinol (Vitamin A) |
| -1 | -1 | propionylcarnitine |
| -1 | -1 | uridine |
| 1 | 1 | imidazole propionate |
| -1 | -1 | histidine |
| -1 | -1 | creatine |
| -1 | -1 | 1-stearoyl-2-oleoyl-GPE (18:0/18:1) |
| -1 | -1 | N-acetylleucine |
| -1 | -1 | 2-aminophenol sulfate |
| -1 | -1 | 2-methylmalonyl carnitine |
| 1 | 1 | sphingomyelin (d18:1/18:1, d18:2/18:0) |
| -1 | -1 | 4-methyl-2-oxopentanoate |
| 1 | 1 | N-acetyl-cadaverine |
| -1 | -1 | beta-hydroxyisovalerate |
| -1 | -1 | glucuronate |
| 1 | 1 | pantothenate |
| -1 | -1 | 2-acetamidophenol sulfate |
| -1 | -1 | alpha-CEHC glucuronide\* |
| 1 | 1 | 3-hydroxybutyrate (BHBA) |
| -1 | -1 | methionine sulfoxide |
| -1 | -1 | pyrraline |
| 1 | 1 | stearoyl sphingomyelin (d18:1/18:0) |
| -1 | -1 | phenylalanine |
| -1 | -1 | asparagine |
| -1 | -1 | N-acetylphenylalanine |
| -1 | -1 | prolylglycine |
| -1 | -1 | 1,2-dilinoleoyl-GPC (18:2/18:2) |
| 1 | 1 | pregnen-diol disulfate\* |
| -1 | -1 | hydantoin-5-propionic acid |
| -1 | -1 | betonicine |
| -1 | -1 | 4-vinylphenol sulfate |
| 1 | 1 | 3-hydroxyhexanoate |
| -1 | -1 | 1-linolenoyl-GPC (18:3)\* |
| 1 | 1 | glycoursodeoxycholate |
| 1 | 1 | 3-hydroxybutyrylcarnitine (1) |
| 1 | 1 | palmitoylcarnitine |
| 1 | 1 | palmitoyl dihydrosphingomyelin (d18:0/16:0)\* |
| 1 | 1 | sphingomyelin (d18:1/24:1, d18:2/24:0)\* |
| -1 | -1 | 1-(1-enyl-palmitoyl)-2-linoleoyl-GPE (P-16:0/18:2)\* |
| -1 | -1 | N-alpha-acetylornithine |
| -1 | -1 | 1-stearoyl-2-oleoyl-GPS (18:0/18:1) |
| -1 | -1 | alpha-CEHC sulfate |
| -1 | -1 | 5alpha-androstan-3alpha,17beta-diol disulfate |
| -1 | -1 | 1-linoleoyl-2-linolenoyl-GPC (18:2/18:3)\* |
| 1 | 1 | oleoylcarnitine |
| 1 | 1 | 3-methoxytyrosine |
| -1 | -1 | N-acetylvaline |
| -1 | -1 | homostachydrine\* |
| -1 | -1 | N-methylpipecolate |
| -1 | -1 | imidazole lactate |
| 1 | 1 | N-palmitoyl-sphingosine (d18:1/16:0) |

* After baseline correction at any time point without BMI adjustment: 129
* at 3 years w/o BMI: 7

# Differential abundance remission at 3 years within RYGB

Analyze each person’s analyte difference from baseline at 3mo between those who achieved glycemic endpoints (coded as *dmcure*) at 36mo and those who did not within RYGB.

I think we should drop, because doesn’t account for beginning.

## Proteome

The Soma table with statistics on differences in remission is at [soma\_dmcure\_table.csv](./results/soma_dmcure_table.csv). The minimum FDR is 0.998.

## Metabolome

The met table with statistics on differences is at [met\_dmcure\_table.csv](./results/met_dmcure_table.csv). The minimum FDR is 0.392.

# Hitman

We apply Hitman 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 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 |

The metabolite mediators are pro-hydroxy-pro

### TO DO Compare Hitman to joint significance analyte mediation

## 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.

# Pathways

For pathway input data, 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, for network plotting.

We plot the top pathways. The t-statistics from ezlimma have sufficiently many degrees of freedom to be approximately z-scores, which are better known, so we annotate them as z-scores. For a network plotting 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.

## Roast 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.

Run roast.

Roast barplot.

Plot Roast pwys.

# Pathway mediation

We apply Hitman to combined (metabolite + protein) dataset, then test pathways.

We need to know inter-gene correlation for method cameraPR.

## HOMA-IR

We write the hitman results for each analyte’s 3mo change from baseline as a mediator of 12mo HOMA-IR change at [homair12\_vs\_analytes3\_hitman\_stats.csv](./results/homair12_vs_analytes3_hitman_stats.csv).

## HbA1c

We find no significant pathways.

## dins030

We find caffeine metabolism significant.

# Session info

## - Session info ----------------------------------------------------------  
## setting value   
## version R version 3.6.1 (2019-07-05)  
## 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 2019-10-02   
##   
## - Packages --------------------------------------------------------------  
## package \* version date lib source   
## assertthat 0.2.1 2019-03-21 [1] CRAN (R 3.6.1)  
## backports 1.1.4 2019-04-10 [1] CRAN (R 3.6.0)  
## callr 3.3.0 2019-07-04 [1] CRAN (R 3.6.1)  
## cli 1.1.0 2019-03-19 [1] CRAN (R 3.6.1)  
## colorspace 1.4-1 2019-03-18 [1] CRAN (R 3.6.1)  
## crayon 1.3.4 2017-09-16 [1] CRAN (R 3.6.1)  
## desc 1.2.0 2018-05-01 [1] CRAN (R 3.6.1)  
## devtools 2.2.0 2019-09-07 [1] CRAN (R 3.6.1)  
## digest 0.6.20 2019-07-04 [1] CRAN (R 3.6.1)  
## dplyr \* 0.8.3 2019-07-04 [1] CRAN (R 3.6.1)  
## DT 0.7 2019-06-11 [1] CRAN (R 3.6.1)  
## ellipsis 0.2.0.1 2019-07-02 [1] CRAN (R 3.6.1)  
## evaluate 0.14 2019-05-28 [1] CRAN (R 3.6.1)  
## ezlimma \* 0.2.3.9028 2019-09-26 [1] local   
## ezlimmaplot \* 0.0.1.9015 2019-09-28 [1] local   
## fs 1.3.1 2019-05-06 [1] CRAN (R 3.6.1)  
## ggplot2 \* 3.2.0 2019-06-16 [1] CRAN (R 3.6.0)  
## glue 1.3.1 2019-03-12 [1] CRAN (R 3.6.1)  
## gtable 0.3.0 2019-03-25 [1] CRAN (R 3.6.1)  
## highr 0.8 2019-03-20 [1] CRAN (R 3.6.1)  
## Hitman \* 0.0.0.9005 2019-10-02 [1] local   
## htmltools 0.3.6 2017-04-28 [1] CRAN (R 3.6.1)  
## htmlwidgets 1.3 2018-09-30 [1] CRAN (R 3.6.1)  
## igraph 1.2.4.1 2019-04-22 [1] CRAN (R 3.6.1)  
## knitr \* 1.23 2019-05-18 [1] CRAN (R 3.6.1)  
## lattice 0.20-38 2018-11-04 [1] CRAN (R 3.6.1)  
## lazyeval 0.2.2 2019-03-15 [1] CRAN (R 3.6.1)  
## limma \* 3.40.2 2019-05-17 [1] Bioconductor   
## magrittr 1.5 2014-11-22 [1] CRAN (R 3.6.1)  
## Matrix \* 1.2-17 2019-03-22 [1] CRAN (R 3.6.1)  
## memoise 1.1.0 2017-04-21 [1] CRAN (R 3.6.1)  
## Mesa \* 0.0.0.9002 2019-08-19 [1] local   
## munsell 0.5.0 2018-06-12 [1] CRAN (R 3.6.1)  
## pheatmap 1.0.12 2019-01-04 [1] CRAN (R 3.6.1)  
## pillar 1.4.2 2019-06-29 [1] CRAN (R 3.6.1)  
## pkgbuild 1.0.3 2019-03-20 [1] CRAN (R 3.6.1)  
## pkgconfig 2.0.2 2018-08-16 [1] CRAN (R 3.6.1)  
## pkgload 1.0.2 2018-10-29 [1] CRAN (R 3.6.1)  
## prettyunits 1.0.2 2015-07-13 [1] CRAN (R 3.6.1)  
## processx 3.4.0 2019-07-03 [1] CRAN (R 3.6.1)  
## ps 1.3.0 2018-12-21 [1] CRAN (R 3.6.1)  
## purrr 0.3.2 2019-03-15 [1] CRAN (R 3.6.1)  
## R6 2.4.0 2019-02-14 [1] CRAN (R 3.6.1)  
## RColorBrewer 1.1-2 2014-12-07 [1] CRAN (R 3.6.0)  
## Rcpp 1.0.1 2019-03-17 [1] CRAN (R 3.6.0)  
## remotes 2.1.0 2019-06-24 [1] CRAN (R 3.6.1)  
## rlang 0.4.0 2019-06-25 [1] CRAN (R 3.6.1)  
## rmarkdown 1.14 2019-07-12 [1] CRAN (R 3.6.1)  
## rprojroot 1.3-2 2018-01-03 [1] CRAN (R 3.6.1)  
## scales 1.0.0 2018-08-09 [1] CRAN (R 3.6.1)  
## sessioninfo 1.1.1 2018-11-05 [1] CRAN (R 3.6.1)  
## stringi 1.4.3 2019-03-12 [1] CRAN (R 3.6.0)  
## stringr 1.4.0 2019-02-10 [1] CRAN (R 3.6.1)  
## testthat 2.1.1 2019-04-23 [1] CRAN (R 3.6.1)  
## tibble 2.1.3 2019-06-06 [1] CRAN (R 3.6.1)  
## tidyselect 0.2.5 2018-10-11 [1] CRAN (R 3.6.1)  
## usethis \* 1.5.1 2019-07-04 [1] CRAN (R 3.6.1)  
## withr 2.1.2 2018-03-15 [1] CRAN (R 3.6.1)  
## writexl 1.1 2018-12-02 [1] CRAN (R 3.6.1)  
## xfun 0.8 2019-06-25 [1] CRAN (R 3.6.1)  
## yaml 2.2.0 2018-07-25 [1] CRAN (R 3.6.0)  
## zeallot \* 0.1.0 2018-01-28 [1] CRAN (R 3.6.1)  
##   
## [1] C:/Program Files/R/R-3.6.1/library