Detecting Differentially Expressed Metabolic Pathways with Adjustment for Macronutrient Intake

Teal Guidici[1], Amy Rothberg[2], Charles Burant[2], George Michailidis[3]

[1]Statistics, University of Michigan, [2] Internal Medicine, University of Michigan, [3]Statistics, University of Florida

Introduction

- Differential expression (DE) testing and set enrichment analysis (SEA) are commonly used to summarize the results of high throughput biological experiments, to generate biologically meaningful hypothesis for further analysis, and to aid in the planning of validation experiments.
- Conventional approaches to differential expression testing and set enrichment analysis do not usually account for individual variation in relevant background features, in many cases due to lack of pertinent data. These features are especially relevant in the context of metabolomics, where blood metabolite levels can react sensitively and quickly to changes in macronutrient intake.
- In this project we introduce a method for detecting differentially expressed metabolites and metabolic pathways, while adjusting for individual variation in the consumption of relevant macronutrients through the integration of macronutrient intake data. We test our method on data gathered from human subjects in a controlled feeding study which featured two distinct diets.

Data and Methods

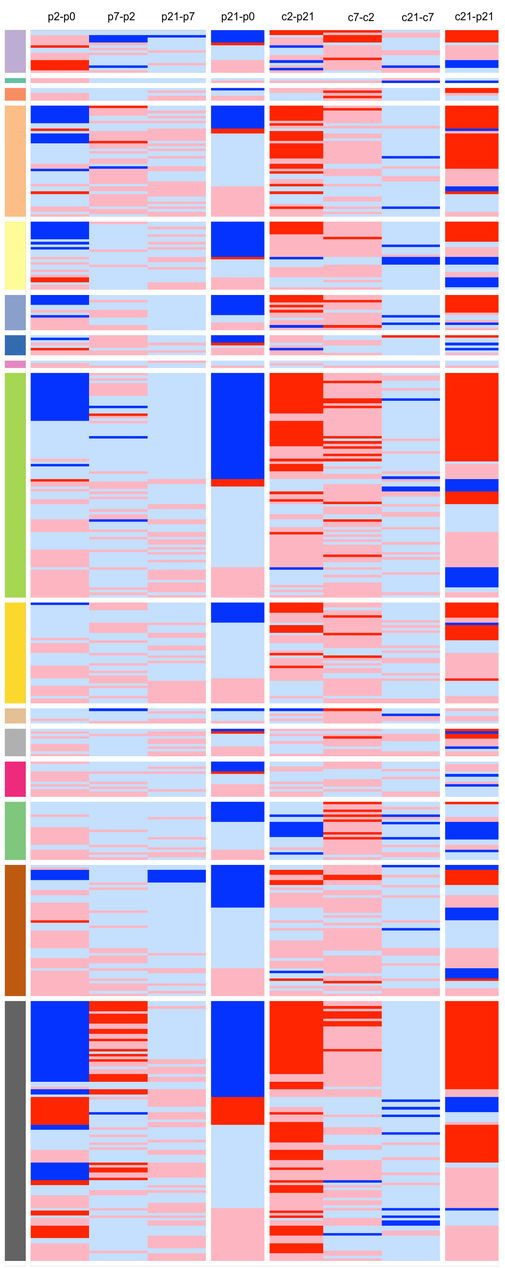
Experimental Design
- 12 healthy adult humans (6m/6f)
- Fed a diet high in Polyunsatruated Fatty Acids (PUFA) for 21 days
- Fed a diet high in Carbohydrates (CARB) for 21 days, immediately after PUFA
- Blood drawn at baseline (day 0) , and days 2, 7, 21 for each diet
  - Day 21 on PUFA designated as baseline for CARB

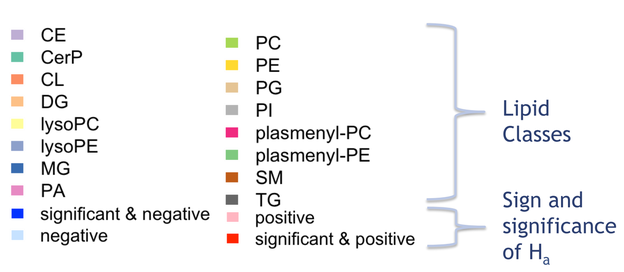
Data
- Macronutrient intake for each subject across diets
- MS Shotgun lipidomics profiling for each subject and each day within each diet. 84 data points on 458 lipids

Analysis – Linear Model
- Main variable: Diet-Day (DD) factor.
  - DD levels: p0, p2, p7, p21, c2, c7,c21 (eg: p2 = PUFA, day 2)
- Model 1: Log_2(Lipid Intensity) ~ 0 + DD
- Model 2: Log_2(Lipid Intensity) ~ 0 + DD + macronutrient intake
- Mixed effects modeling used to account for nested structure of data (days:diets:subjects), and variability between subjects.
- Testing for DE between levels of DD
  - fdr corrected pvalues <0.1
- Tested DE lipids for enrichment in relevant lipid ontologies

Analysis – Networks
- Lipid Correlation Network generated separately under PUFA and CARB
  - Significant correlations* retained, non-significant set to 0. (*based on fdr corrected pvalues via Fisher transformation, cutoff of 0.05.)
- Subnetworks in PUFA and CARB identified by leading eigenvector community detection.
  - Tested for enrichment in relevant lipid ontologies
- Subnetworks tested for overall differential expression
  - Subnetwork summary statistic: Maxmean of lipid level differential expression t-statistics.
  - Significance calculated from permutation distribution

Linear Model





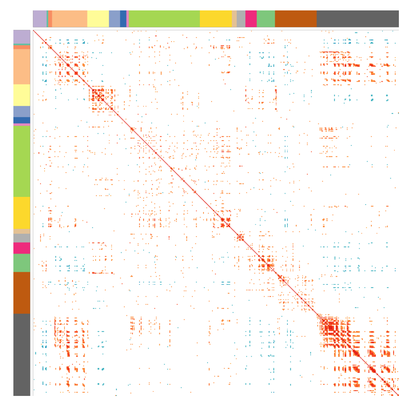
Tested for DE between levels of DD.
- H_o: DD_i-DD_j = 0
- H_a: DD_i-DD_j ≠ 0

- Model 1: Many lipids have significant DE between beginning and end of diet
  - p21-p0≠0
  - p21-c21≠0
- Model 1: For each diet, lipids separate into:
  - Fast: DE between baseline and day 2 AND baseline and day 21
  - Slow: DE between baseline and day 21 only
  - No DE: not DE between baseline and day 21
- Model 1: SEA indicates fast & slow lipids in either diet enriched for certain classes and saturation levels of lipids
- Model 2: Only 2 lipids show significant DE between baseline and day 21 for either diet.
- For many lipids, Model 1 and/or Model 2 significant overall
  - Model 1: 287/458
  - Model 2: 283/458

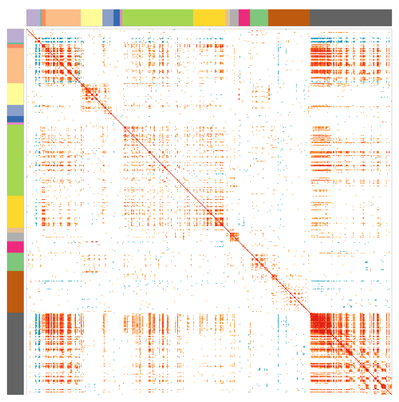
		CARB		
		Fast	Slow	No DE
PUFA	Fast	62	20	11
	Slow	25	23	39
	No DE	39	53	186

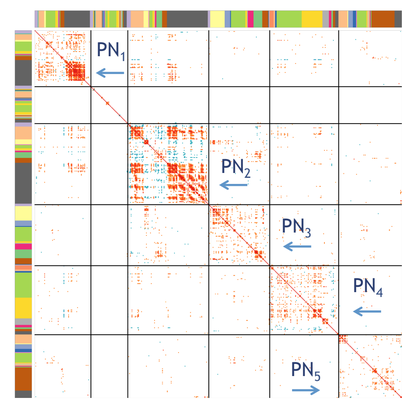
- Distribution of lipids across DE labels (Fast, Slow, No DE) for each diet. PUFA labels have significant association with CARB labels (Pearson's Chi-squared p < 2.2e-16)

Lipid Network

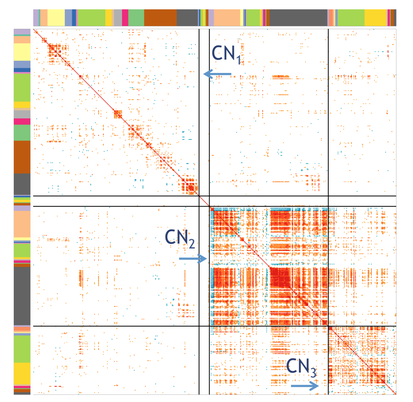


Sorted into lipid classes



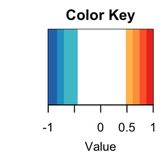


Sorted into lipid subnetworks



PUFA

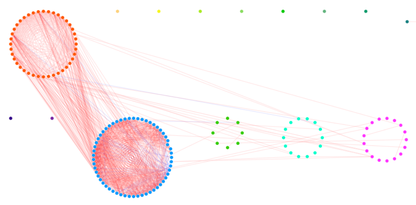
CARB



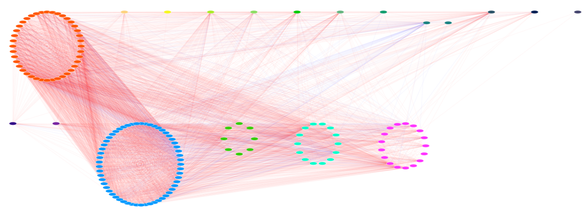
- Lipid Correlation subnetworks, before & after sorting into identified subnetworks.
- Main subnetworks located on diagonal, black outlines
- Subnetworks in both diets enriched for various classes and saturation levels of lipids

- PUFA subnetworks labeled PN_i,CARB subnetworks labeled CN_j

- Subnetworks in PUFA dissolve and coalesce into new subnetworks in CARB



CN_2 nodes, with edges present in PUFA



CN_2 nodes, with edges present in CARB

- Nodes colored and arranged by PUFA subnetwork labels.
- Edges colored by sign of correlation coefficient between nodes (red >0, blue <0)

Detecting Differentially Expressed Networks

- PUFA, CARB subnetworks tested for enrichment (E) or depletion (D) of DE lipids via over representation analysis (ORA).
  - Enrichment: more DE lipids than by chance
  - Depletion: fewer DE lipids than by chance
  - via Fisher's Exact Test
- PUFA, CARB subnetworks tested for subnetwork level DE with modified Gene Set Analysis (GSA) algorithm.
  - Designed to capture joint activity of non-signif DE lipids
  - Maxmean statistic calculated based on t-statistics from:
    - PUFA: H_a: p21-p0≠0
    - CARB: H_a: c21-p21≠0

		ORA by DE label			Subnetwork DE	
		Fast	Slow	No DE	Model 1	Model 2
Subnetwork	PN_1	E	-	D	DE	DE
	PN_2	-	-	-	DE	-
	PN_3	-	-	-	-	-
	PN_4	D	E	-	-	-
	PN_5	D	E	-	-	-
	CN_1	D	E	E	-	-
	CN_2	E	D	D	DE	DE
	CN_3	-	-	-	DE	-

Conclusions and Continuing Work

- DE lipids share common, related patterns of expression across diets
  - Decrease in PUFA, increase in CARB
  - Fast or Slow changes within each diet
- DE behavior of lipids can be near-completely accounted for by macronutrient intake
- Lipids form distinct subnetworks of highly correlated lipids under PUFA and CARB
- Lipid subnetworks show significant enrichment and depletion for DE labels in both diets
- Lipid subnetworks show significant network level DE
- Subnetwork level DE algorithm successfully captures many small changes throughout network
  - Subnetworks without enrichment for active lipids have network level DE (PN_2, and CN_3).
  - 99.5% of individual lipids have no DE under Model 2, but several lipid subnetworks (PN_1, and CN_2) have network level DE under Model 2
- Continuing Work
  - Create a set of "lipid" biomarkers to identify diet of individual based on lipidomic profile
  - Determine additional physical/chemical properties unifying lipids with common DE labels or subnetwork assignment.

This work was supported by NIH T32 Proteome Informatics of Cancer Training Program at the University of Michigan (T32-CA-140044) and by the Michigan Regional Comprehensive Metabolomics Resource Core (U24 DK097153). Special thanks to Tanu Soni and Thekkelnaycke Rajendiran.
Contact information: tealg@umich.edu