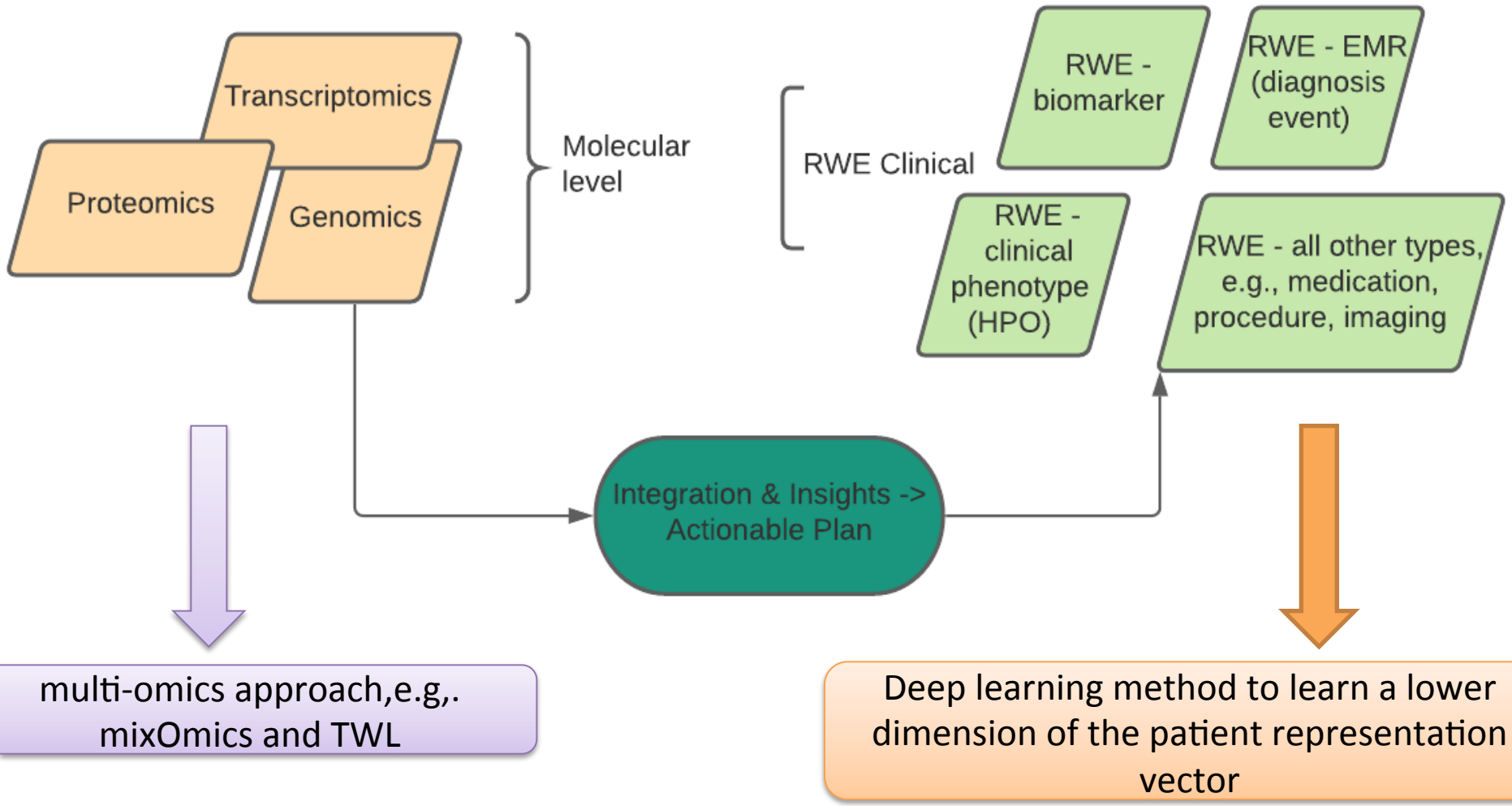
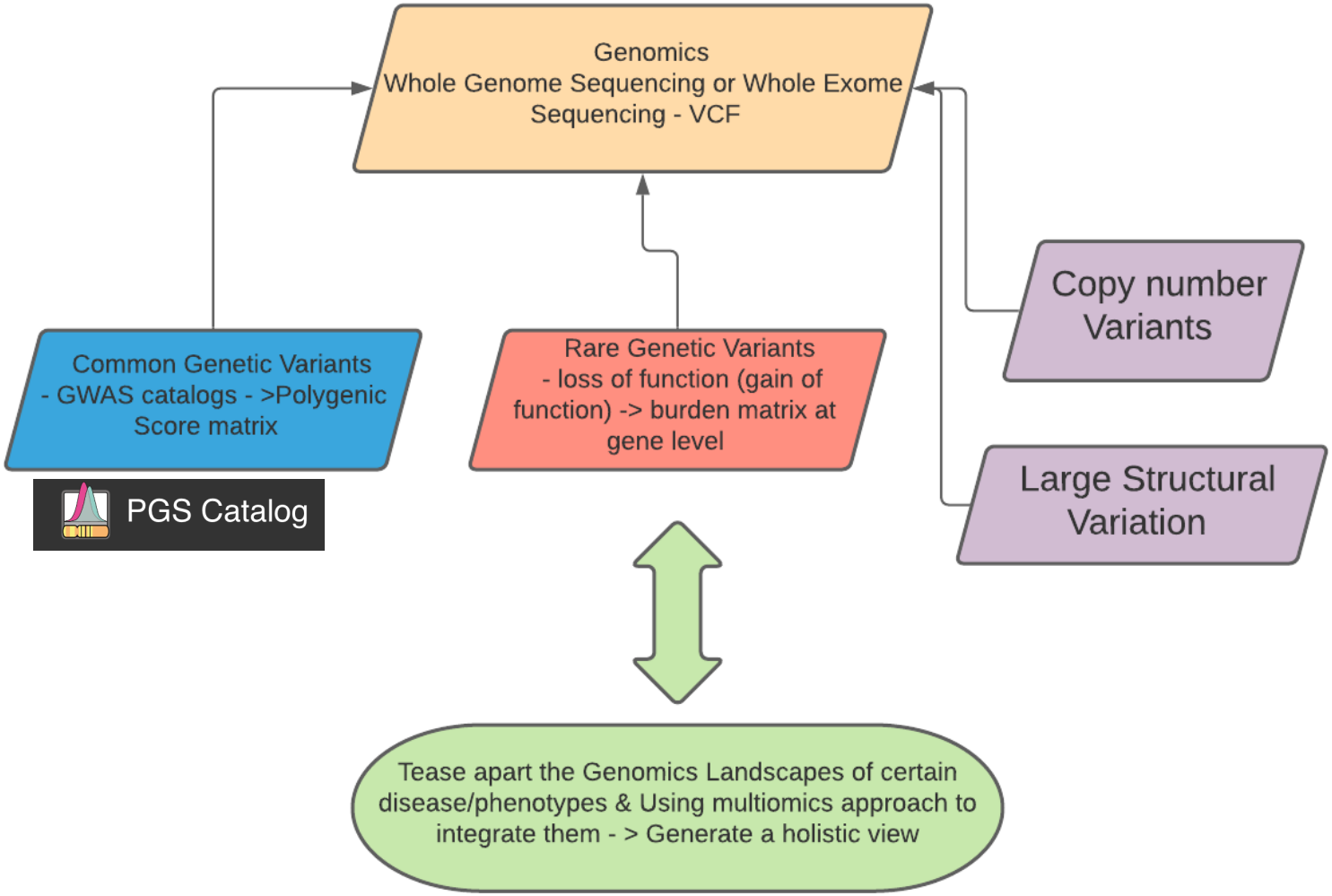
# Using multiomics approaches (RWE + omics) in patient subgroup identification and disease mechanism elucidation

# The Goal and Challenges



# Genomic Data – is it just one layer?




https://www.pgscatalog.org/

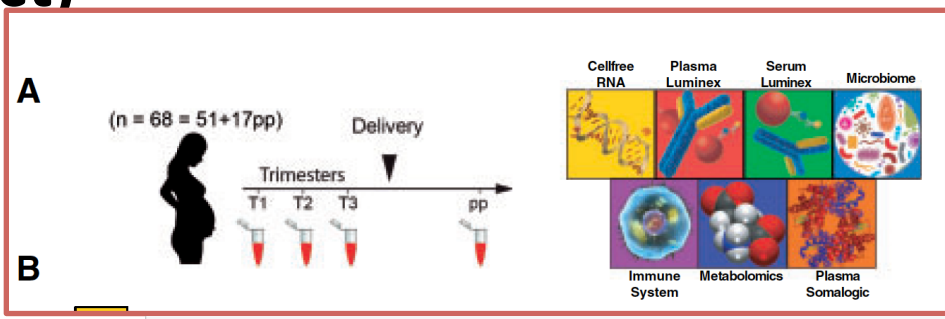
# A Case Study – Multiomics Integration (https://github.com/leafiezyt/multiomics/tree/pet_project)

*Bioinformatics*, 35(1), 2019, 95–103
doi: 10.1093/bioinformatics/bty537
Advance Access Publication Date: 2 July 2018
Original Paper

OXFORD

Systems biology

**Multiomics modeling of the immunome, transcriptome, microbiome, proteome and metabolome adaptations during human pregnancy**




```
df <- as.data.frame(InputData[[2]]) ## plasma luminex
head(df)
```

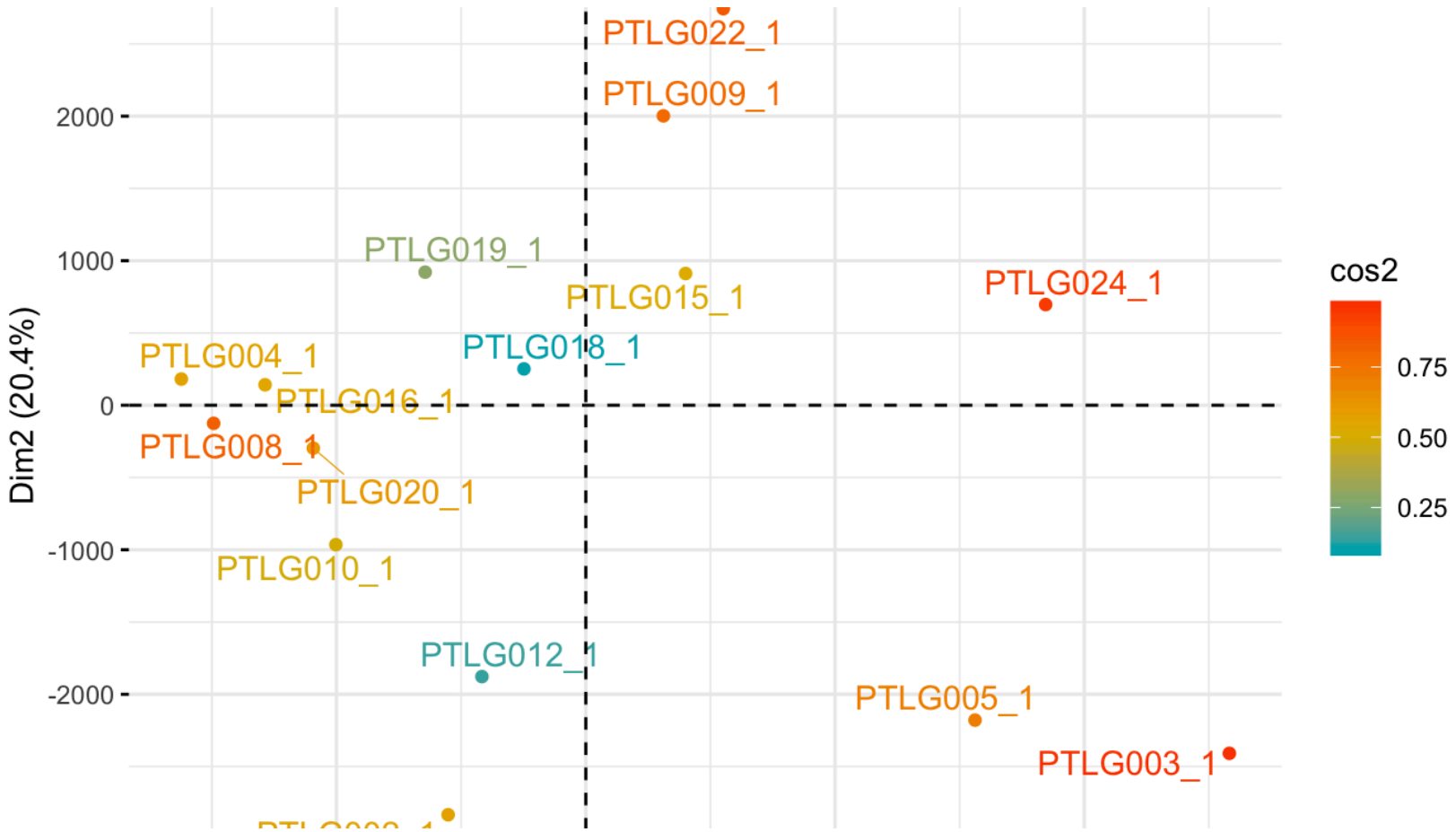
```
##           plasma-IL17F plasma-FASL plasma-TGFA plasma-MIP1A plasma-SDF1A
## PTLG002_BL        59.00       82.00       27.75        46.00       191.25
## PTLG003_BL        87.25      221.50       48.50        71.50       186.25
## PTLG004_BL       686.50     1088.50      453.75       252.25       575.50
## PTLG005_BL        91.25      112.75       31.00        91.00       272.00
## PTLG007_BL       186.75      544.75      194.25       171.25       345.75
## PTLG008_BL        69.50      105.75       97.50        62.00       192.00
```

- The case study questions
  - just for exemplary for approaches
  - not enough sample sizes, no covariates to adjust for confounders

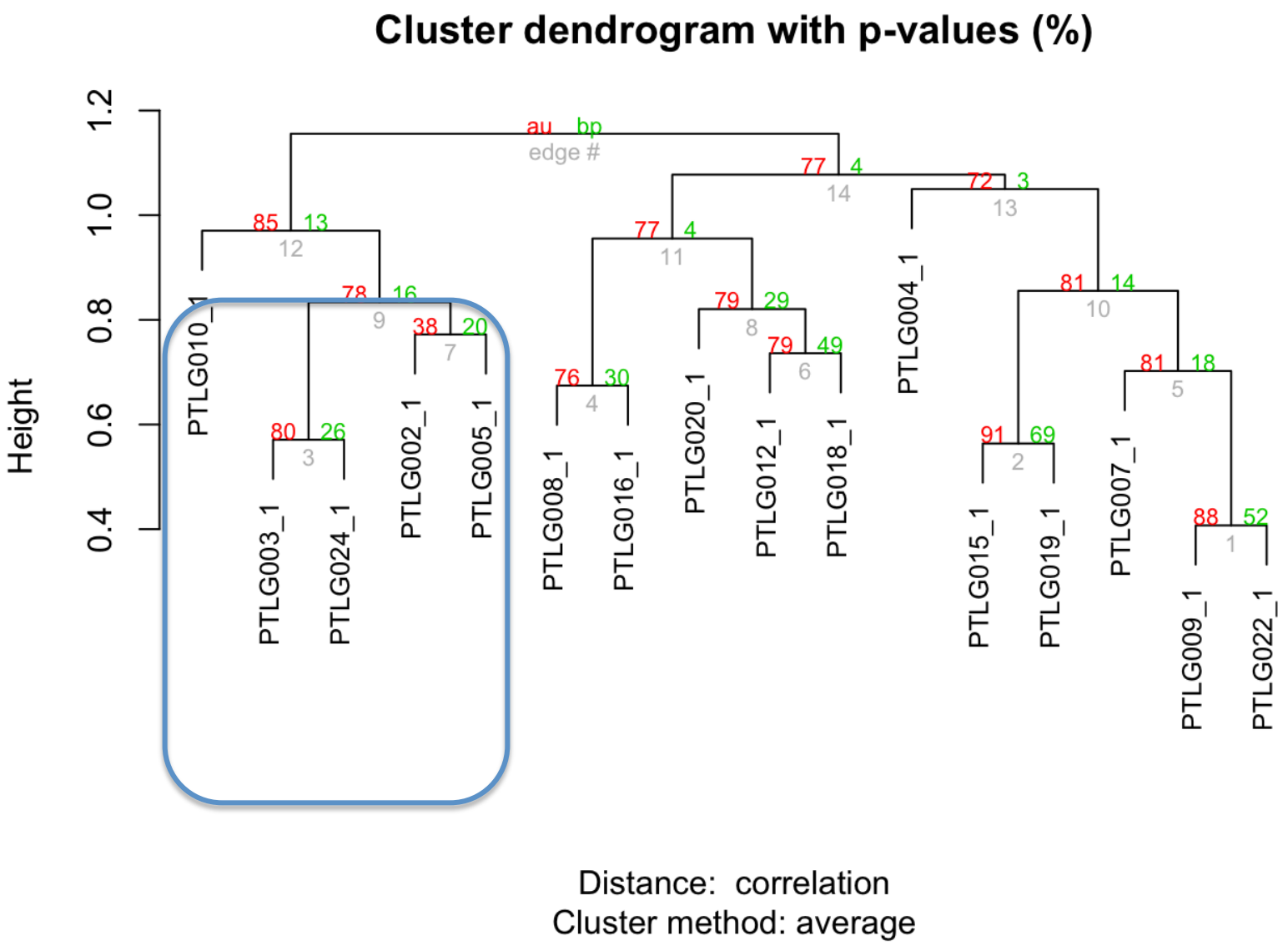
- Single layer of the data – can be done at each omics level
  - Plasma-luminex
  - Subgroup discovery if there's any
  - characteristics of the subgroups
    - using ML & model building to understand the biology behind

- Multi-omics integration
  - Plasma luminx and immune cell profiling

# PCA on the Plasma Luminx Data



# Two Main Subgroups Identified among the 17 Patients (1st trimester)



# Which features/plasma biomarkers associates with the subgroups & maybe potential biomarkers' hypotheses emerged?

```
## [1] "plasma-BDNF"
##
##  Pearson's product-moment correlation
##
## data:  df2[, i] and df2$clus
## t = -4.5001, df = 14, p-value = 0.0004991
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.9156315 -0.4415154
## sample estimates:
##       cor
## -0.7689313
##
## [1] "plasma-VCAM1"
##
##  Pearson's product-moment correlation
##
## data:  df2[, i] and df2$clus
## t = 3.0593, df = 14, p-value = 0.008492
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.2000388 0.8591175
## sample estimates:
##       cor
## 0.6329779
```

BDNF

RESEARCH ARTICLE

## Blood heavy metals and brain-derived neurotrophic factor in the first trimester of pregnancy among migrant workers

**Ye Htet Zaw[☯], Nutta Taneepanichskul[☯*]**

College of Public Health Sciences, Chulalongkorn University, Bangkok, Thailand

☯ These authors contributed equally to this work.
* nutta.taneepanichskul@gmail.com

Abstract

**Background**

Lead, mercury, cadmium and arsenic are the priority heavy metals of major public health concern in industrialized countries. Exposure to them can cause cognitive impairment and depressive disorders through an effect on Brain-derived neurotrophic factor (BDNF) which is an important biomarker of pregnancy. Despite a number of prior studies on heavy metals pollution, there is few of studies on the effect of heavy metals on BDNF during early pregnancy. This study aims to examine the association between maternal blood heavy metals concentrations and BDNF during the first trimester pregnancy among Myanmar migrants in Thailand.

VCAM1

› Acta Obstet Gynecol Scand. 2002 Aug;81(8):713-9.

## Soluble tumor necrosis factor receptor II and soluble cell adhesion molecule 1 as markers of tumor necrosis factor-alpha release in preeclampsia

Wil Visser [1], Ilse Beckmann, Marco A H Knook, Henk C S Wallenburg

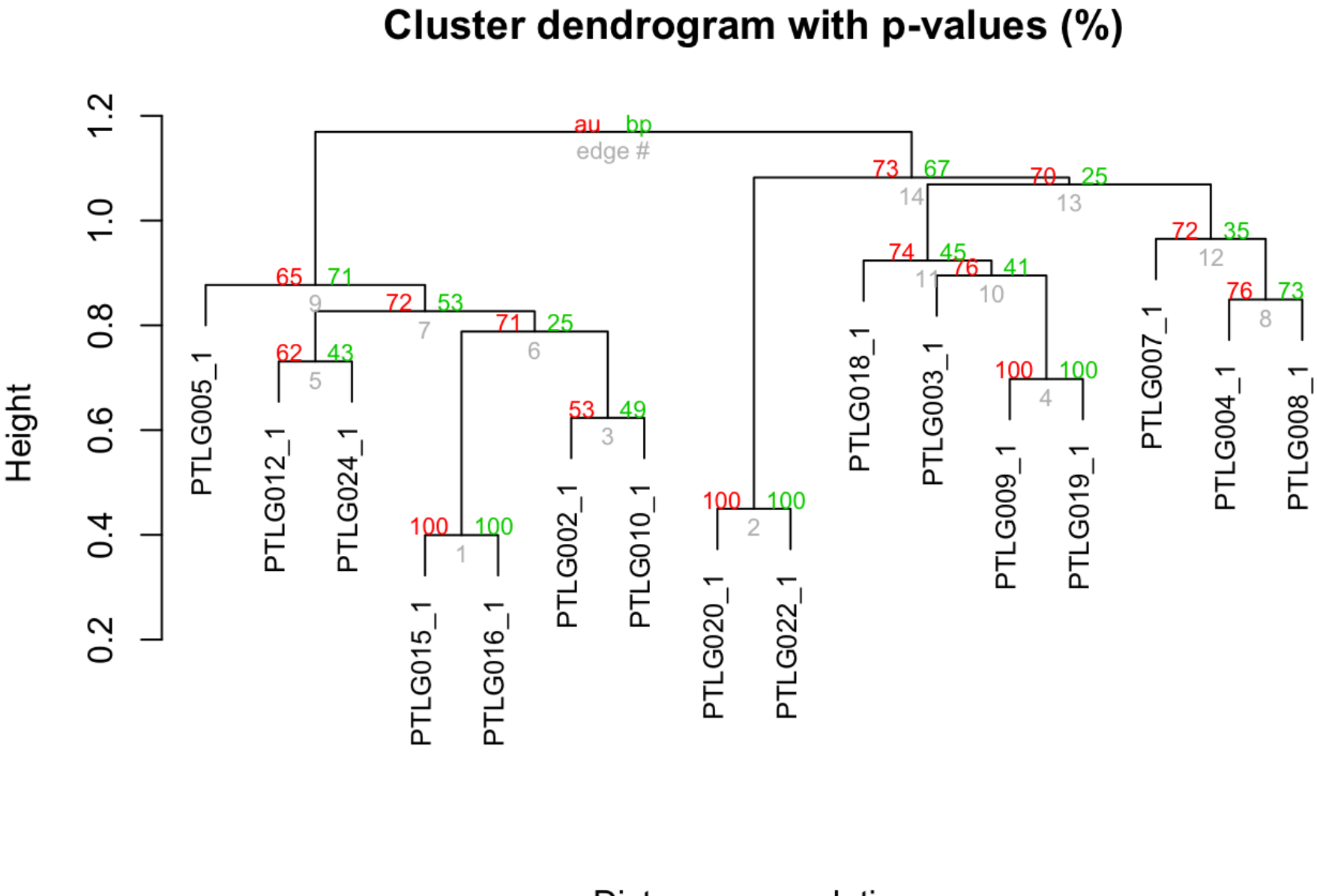
Affiliations + expand

PMID: 12174154

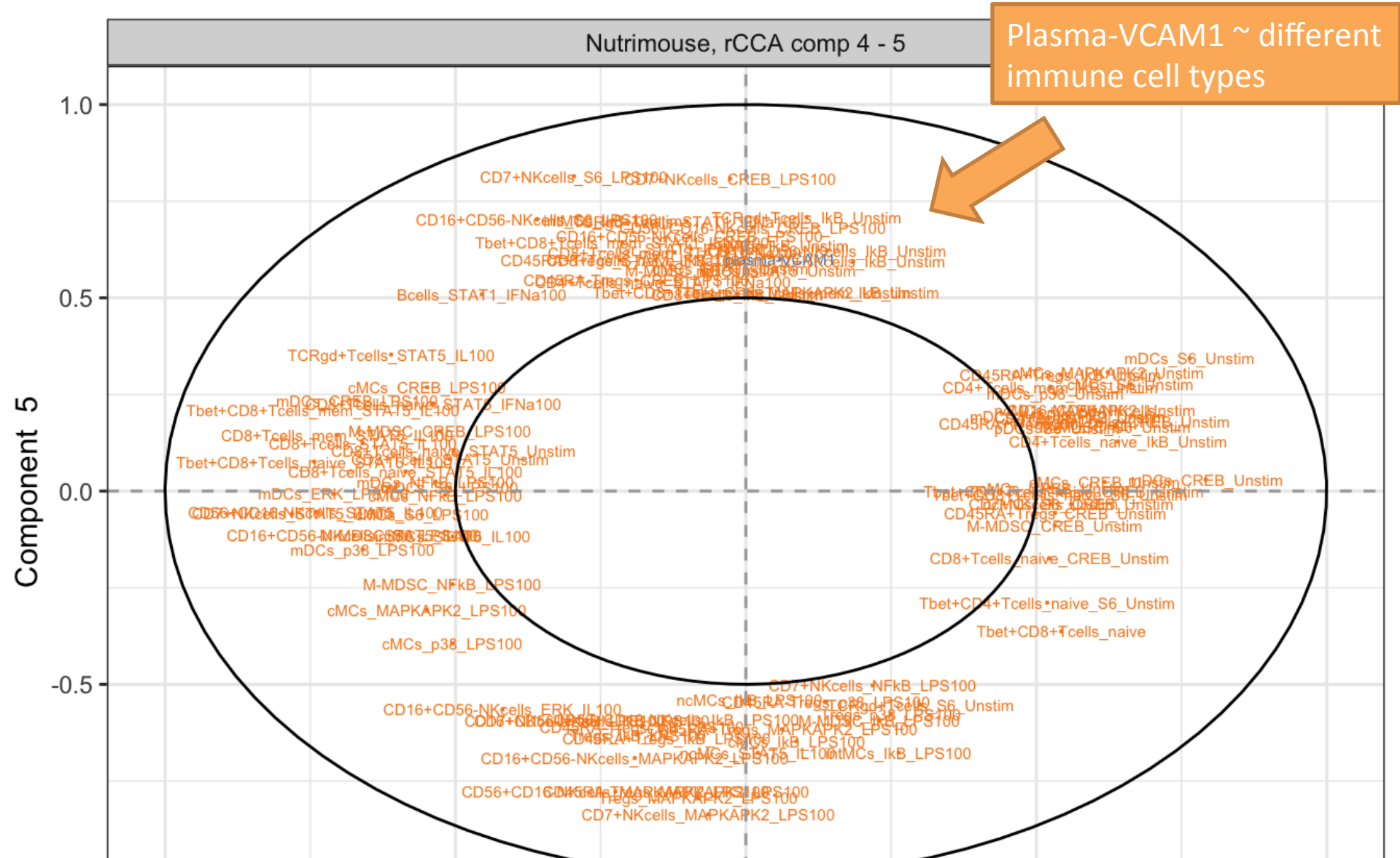
Abstract

**Background:** The purpose of this case-controlled study was to investigate whether plasma concentrations of TNF-receptors I and II and tumor necrosis factor-alpha-induced cell adhesion molecule 1 VCAM-1 could serve as more sensitive markers of tumor necrosis factor-alpha release in preeclamptic women than a direct measurement of circulating tumor necrosis factor-alpha.

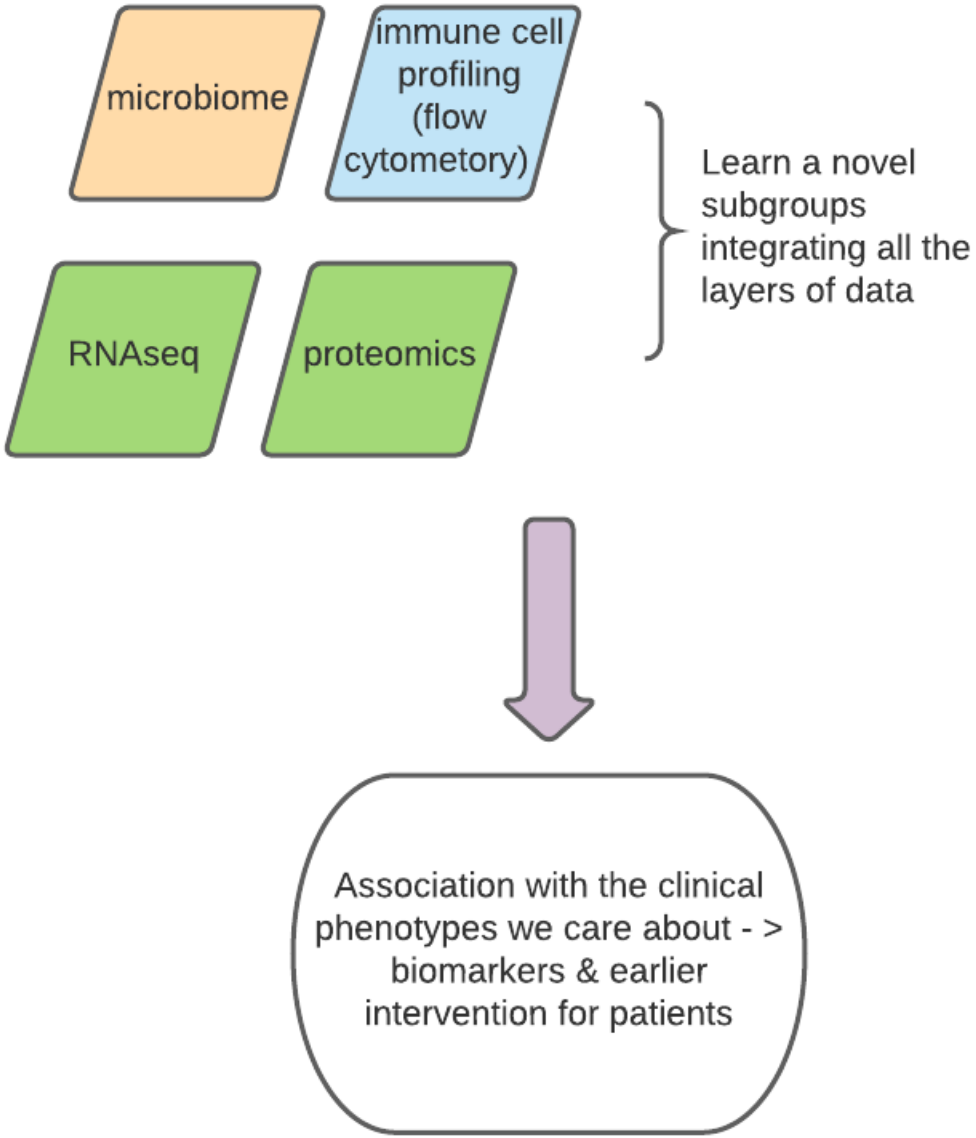
# Immune Cells Profiling



# rCCA – Looking for Correlated Features among Two Omics Datasets



# More Future Directions/ Brainstorming Ideas




Lots of directions to go. We still need to have the patients covariates to address for confounders.

# Conclusion

- Tease apart each layer of the data & investigate
  - needs a lot of domain knowledge
  - thorough and sometimes a lot of manual work, but very important and rewarding
- Multi-omics approach for high dimensional omics datasets integration
- RWE + multi-omics

Machine learning & unsupervised/supervised approaches

SCIENTIFIC REPORTS

OPEN

## Learning from Longitudinal Data in Electronic Health Record and Genetic Data to Improve Cardiovascular Event Prediction

Received: 14 August 2018
Accepted: 23 November 2018
Published online: 24 January 2019

Juan Zhao[1], QiPing Feng[2], Patrick Wu[1,3], Roxana A. Lupu[4], Russell A. Wilke[4], Quinn S. Wells[5], Joshua C. Denny[1,5] & Wei-Qi Wei[1]