Multi-Omics Data in Environmental Health

David Conti, PhD
Norris Comprehensive Cancer Center
Keck School of Medicine
University of Southern California

USC Center for Translational Research on Environmental Health

Keck School of Medicine of USC
Department of Population and Public Health Sciences

Collaborative Team

Population Health Scientists

Data Scientists

Bench Scientists

Dept. of Population and Public Health Sciences

Leda Chatzi, MD, PhD

David Conti, PhD

Wendy Setiawan, PhD

Max Aung, PhD

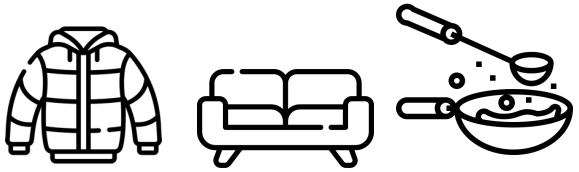
Department of Medicine

Lucy Golden, PhD

Ana Maretti, PhD

How can we advance mechanistic insight linking the human exposome to health across the life-course?

Phthalates and Phenols



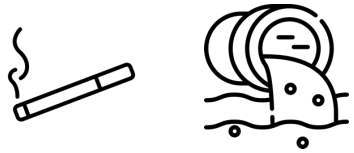
Persistent pollutants (PFAS)

Biological Mechanisms

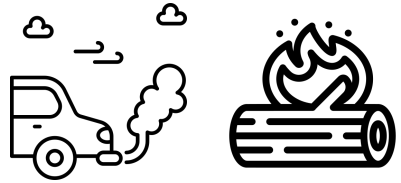
Can we identify the risk factors?

Can we identify the important biological pathways or patterns?

Trace metals



Air pollution



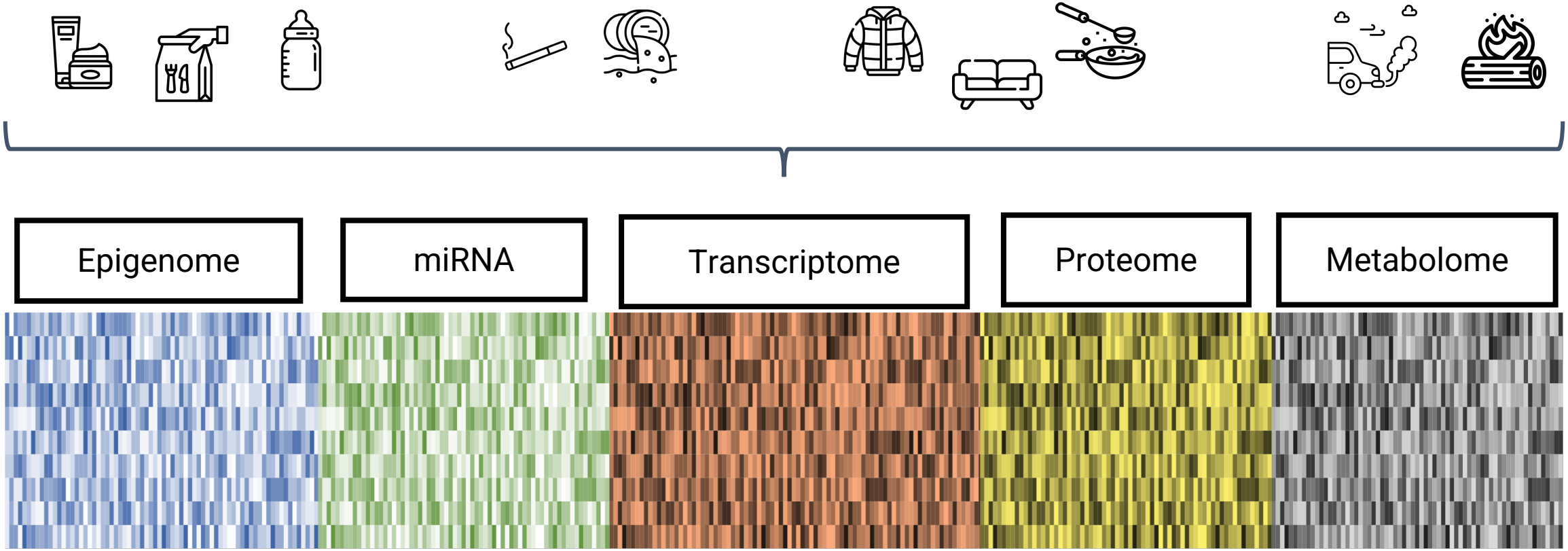
Measuring multiple exposures and omics layers

Phthalates and Phenols

Trace metals

Persistent pollutants (PFAS)

Air pollution



Application

n= 1200 mother-child pairs

HELIX BUILDING THE EARLY-LIFE EXPOSOME

0 Baseline

2 Birth/Infancy

3 Years

Maternal biomarker-based EDC exposures
(DDE, PCB-153, PFOA, PFOS, BPA, PBDE, Phthalates, heavy metals)

Pregnancy trimester-specific air pollution
(PM, NO2, traffic load)

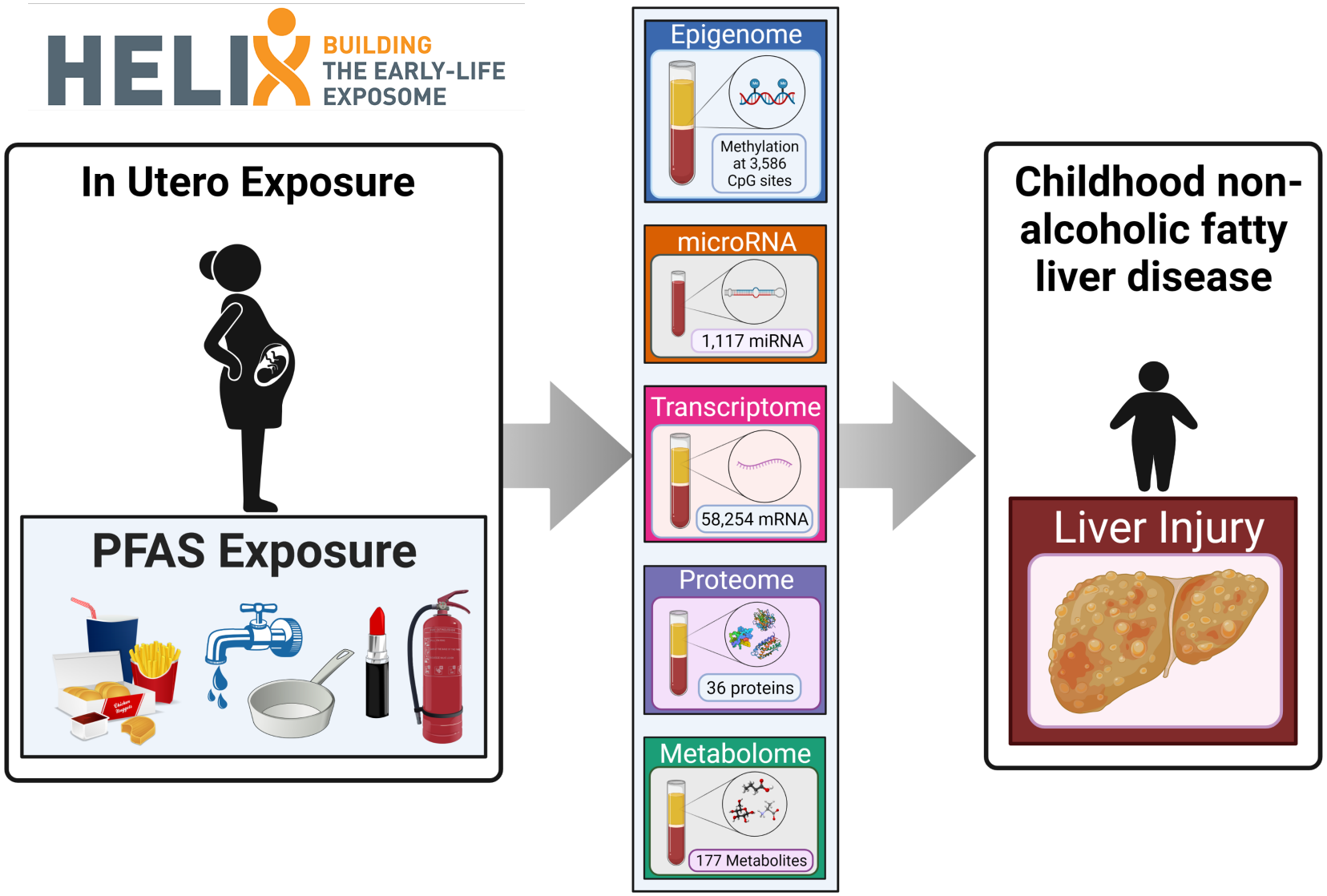
CHILD LIVER INJURY
-Liver Enzymes
-CK-18
-Cytokines and Adipokines

Omics Data

- Urinary metabolomics
- Serum metabolomics
- Proteomics
- Methylation
- Transcriptomics
- MicroRNA (miRNA)
- Mitochondrial DNA content

Liver Injury Risk in the HELIX cohort

Application



Analysis frameworks for multiple exposures, multiple omics layers, and an outcome

nature COMMUNICATIONS

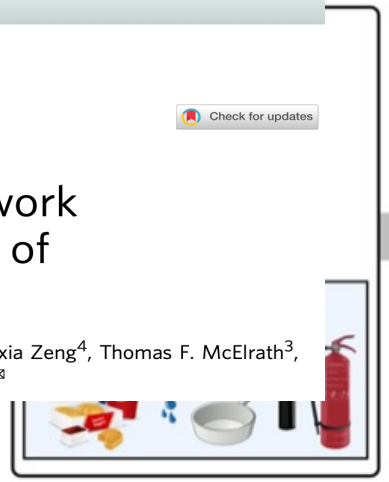
ARTICLE

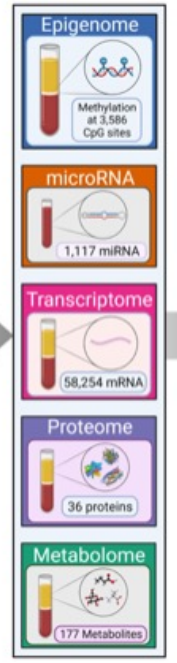
https://doi.org/10.1038/s41467-020-19335-2 OPEN

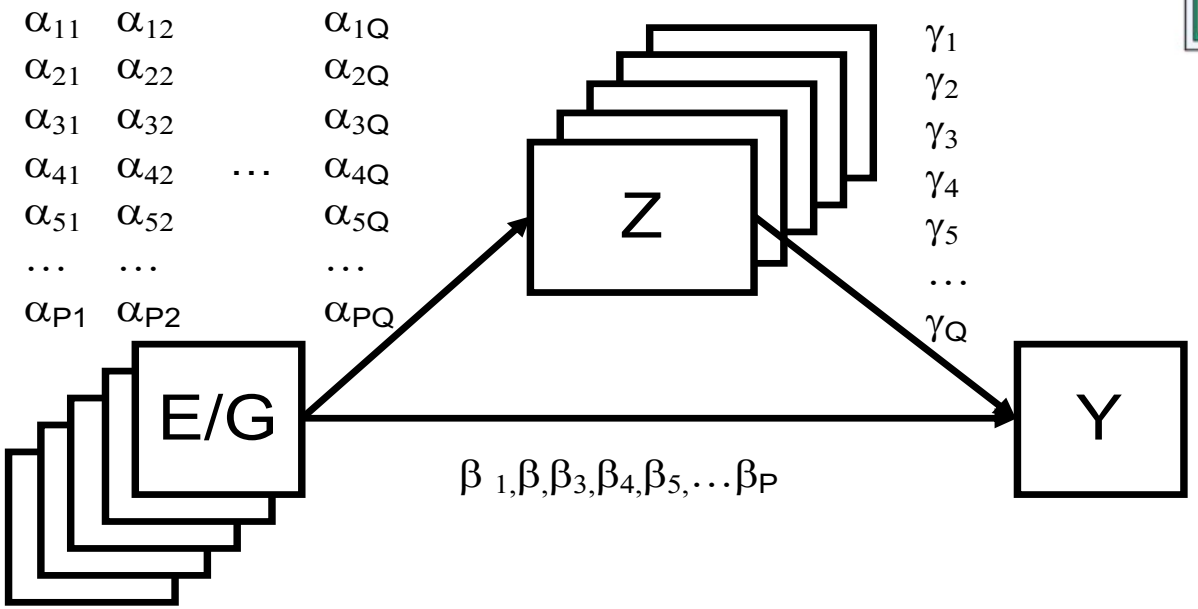
Check for updates

Application of an analytical framework for multivariate mediation analysis of environmental data

Max T. Aung[1], Yanyi Song[1], Kelly K. Ferguson[2], David E. Cantonwine[3], Lixia Zeng[4], Thomas F. McElrath[3], Subramaniam Pennathur[4,5,6], John D. Meeker[7] & Bhramar Mukherjee[1,8]✉







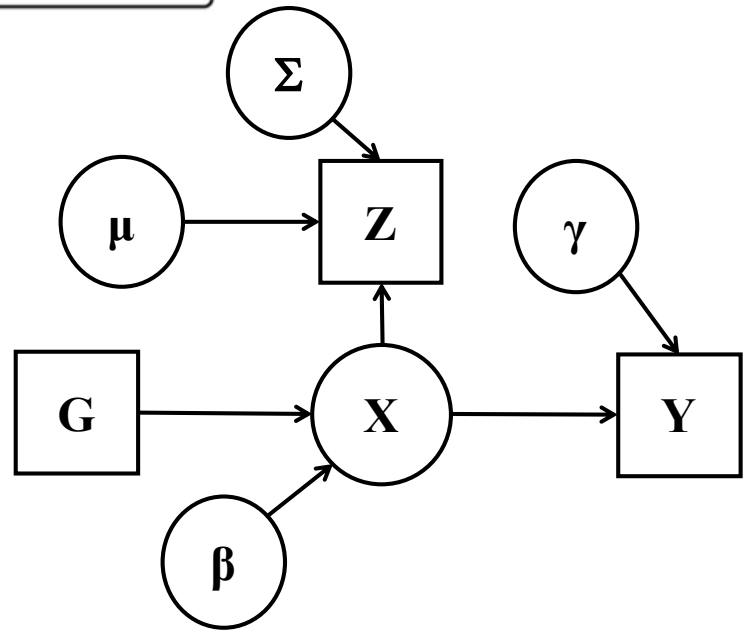
Bioinformatics, 36(3), 2020, 842–850
doi: 10.1093/bioinformatics/btz667
Advance Access Publication Date: 29 August 2019
Original Paper

OXFORD

Genetics and population analysis

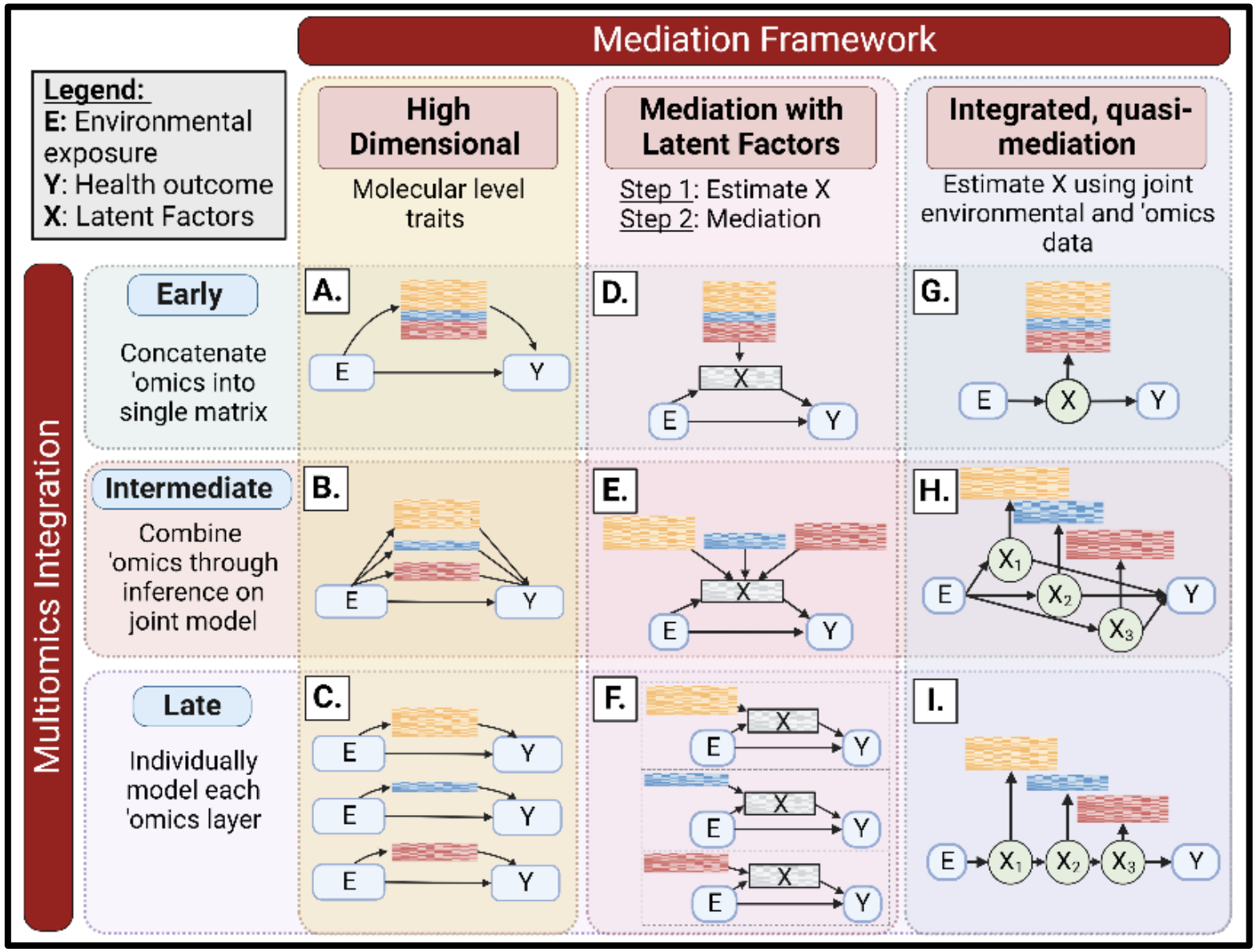
A latent unknown clustering integrating multi-omics data (LUCID) with phenotypic traits

Cheng Peng[1], Jun Wang[1], Isaac Asante[2], Stan Louie[2], Ran Jin[1], Lida Chatzi[1], Graham Casey[3], Duncan C. Thomas[1] and David V. Conti[1,*]



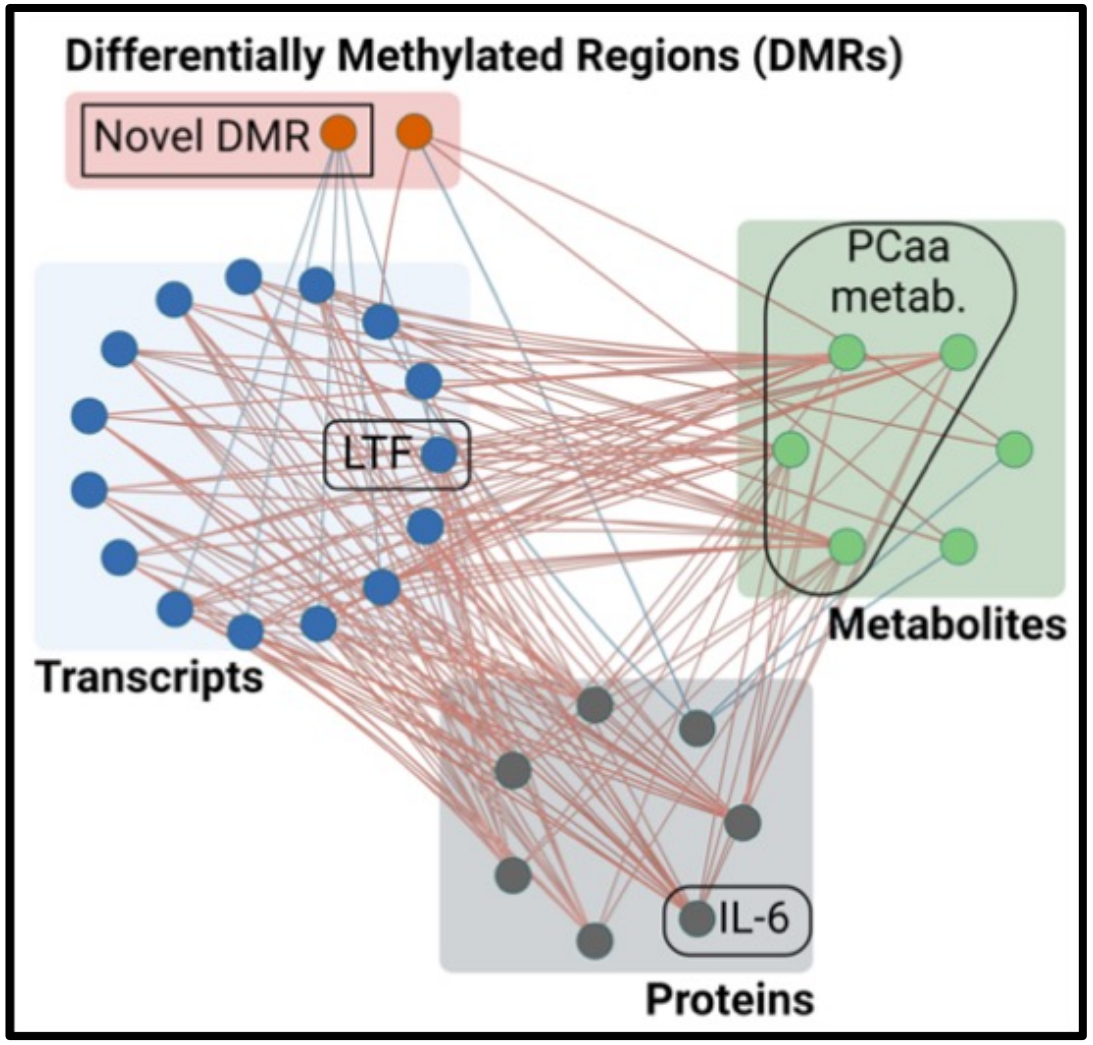
Multi-omics Integration

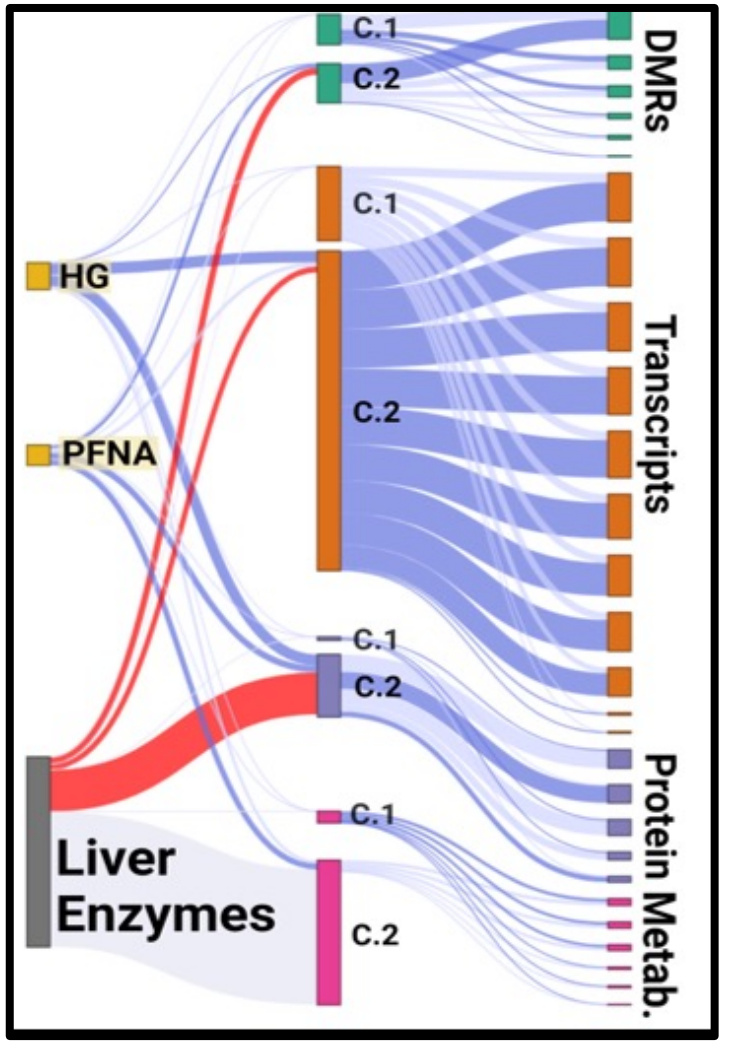
Analytical Framework



Liver Injury Risk in the HELIX cohort

Application





Future Directions

Analytic Considerations

- **Omic features:**
  - High dimensional features within each omic layer.
    - Currently use machine learning for feature selection.
    - Omic features often highly correlated.
  - Balance estimation and inference within and across omic layers.
- **Need to adjust for study design covariates.**
- **Temporal or biological relation to data:**
  - Exposures -> Omics -> Outcome
- **Potential to incorporate external biological info**
  - From experiments, ontologies, etc.

Overall goals:

- Identify causal features.
- Identify relevant biological patterns.
- Predict outcomes.

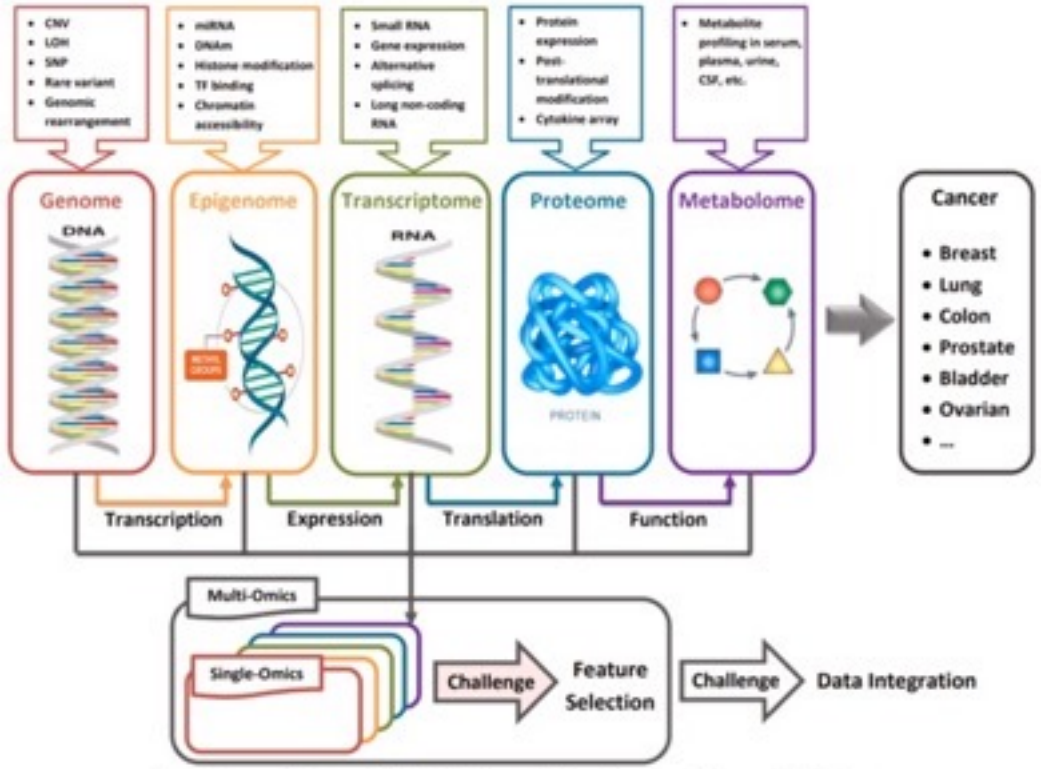


Fig. 1. The overall diagram of relationship between single and multi omics data analysis challenges.