Systems Genetics Approaches to Compare Mechanisms of Smooth Muscle Cell Plasticity in Quiescent and Proliferative States in Coronary Artery Disease

(Technical Paper)

How the Global Footprint Network Promotes Sustainability

(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Diana Albarracin

November 1, 2021

Technical Project Team Members
Noah Perry, Ph.D Candidate

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Diana Albarracin

Mete Civelek, Department of Biomedical Engineering

Hannah Rogers, Department of Engineering and Society

General Research Problem

How can sustainable practices improve people's access to food?

Food inaccessibility is a growing concern around the world (Berlin et al., 2012). Agricultural production continues to increase to meet the growing demand for food (Spiertz, 2009). However, because of their environmental effects, many common agricultural practices are unsustainable. More sustainable agricultural techniques are needed to meet current and future demand for food. Sustainable means of improving food accessibility may also improve cardiovascular health as food insufficiency has been associated with risk factors for several cardiovascular diseases (CVDs) (Liu and Eicher-Miller, 2021; Vercammen et al., 2021). CVDs include coronary artery disease (CAD), the number-one cause of death worldwide (Khera & Kathiresan, 2017).

Systems Genetics Approaches to Compare Mechanisms of Smooth Muscle Cell Plasticity in Quiescent and Proliferative States in Coronary Artery Disease

How can computational approaches be used to design a novel and more comprehensive way of comparing quisicient and proliferative states in coronary artery disease?

Coronary artery disease (CAD) is recognized to be from 40% to 70% heritable, though most of the genes associated with it have unknown mechanisms (Khera and Kathiresan, 2017). These mechanisms are thought to function by regulation of gene expression in the vascular wall where smooth muscle cells (SMCs) play essential roles in the development of atherosclerosis, the underlying cause of CAD (Van der Harst Pim and Verweij Niek, 2018). SMCs possess a high phenotypic plasticity that plays a role in cardiovascular diseases (Frismantiene et al., 2018). Hence, understanding the phenotypic changes in SMCs between the healthy (quiescent) and atherogenesis (proliferative) states in CAD may show important underlying mechanisms.

Different computational approaches are used to identify underlying mechanisms and genes in CAD (Langfelder and Horvath, 2008). The current standard methods used to detect changes in biological pathways between healthy and disease conditions are differential gene expression (DESeq) and gene set enrichment analysis (GSEA) (Anders and Huber, 2010; GSEA, n.d.). GSEA yields insights into different diseases by focusing on groups of genes that share commonalities such as chromosomal location, regulation or biological function (Subramanian et al., 2005). Differential expression analysis works by looking at the level at which specific genes are expressed between different conditions. This allows for understanding at which biological processes may be affected between the proliferative and quisient conditions in CAD (DGEA, n.d.). However, even with these approaches most of the genes associated with CAD still have unidentified mechanisms (Khera and Kathiresan, 2017).

To yield new insights into these unidentified mechanisms associated with CAD, my technical project aims to: (1) Investigate existing pathway analysis methods to compare the disease and healthy state of CAD, (2) determine optimal methods to predict functions of unidentified pathways, and (3) use two different approaches to validate the relevance of the novel pathways found. This is an independent project supervised by Dr. Mete Civelek in the Biomedical Engineering Department and Noah Perry, a Ph.D student, is my collaborator.

Methods for this project include performing genotyping and RNA-seq analysis on aortic smooth muscle cells from 151 donors cultured in quiescent and proliferative conditions. This is followed by performing weighted gene co-expression network analysis (WGCNA) in R (Langfelder and Horvath, 2008). WGCNA is a systems biology method for describing the correlation patterns among genes and cluster them into modules. Then, module preservation statistics will be performed to find modules of highly correlated genes and relating said modules to each other. Module preservation statistics have been shown to be useful for studying differences between the modular structure of networks (Langfelder et al., 2011). Comparing the least preserved modules between the proliferative and quisient state of the SMCs can demonstrate important underlying mechanisms and genes in CAD. These comparisons will be done performing Gene Oncology (GO) Term enrichment using Panther to understand which pathways are enriched in the unpreserved modules (Mi et al., 2019).

Identifying more comprehensive methods of comparing proliferative and quisient conditions in CAD may advance the genetic research in multiple diseases that have been shown to be heritable. Lastly, identifying novel susceptibility mechanisms that function in the vessel wall where the disease develops is expected to be key to finding new therapeutic targets for the disease.

How the Global Footprint Network Promotes Sustainability

How has the Global Footprint Network promoted sustainability?

The Global Footprint Network (GFN), founded in 2003, is a nonprofit that uses United Nations data to calculate countries' Ecological Footprint (EF) (GFN, n.d.c). The EF is a sustainability metric developed in the 1990s by William Rees and Mathis Wackernagel, the CEO of GFN. It quantifies human demand on natural ecosystems against their capacity to regenerate (Wackernagel & Pearce, 2018). The EF can purportedly determine the theoretical date in a given year when, at current rates of consumption, humans will have depleted resources beyond Earth's capacity to regenerate them (GFN, n.d.b). In 2021, this date was July 29th. Though the EF has been applied to guide sustainability policies, some researchers contend that it is an inaccurate measure (Office, 2006; Van den Bergh and Grazi, 2014). Though the EF has been widely used to promote more sustainable practices, it is controversial. GFN's strategies, including its use of the EF, are therefore worth closer examination.

Participants include the GFN, the agenda of which is to change "how the world manages its natural resources" (GFN, n.d.a). It promotes sustainable practices that are consistent with ecological limits as measured by the EF metric. Other participants include nonprofits such as the San Francisco Planning and Urban Research Association (SPUR), which uses GFN footprint data to guide more sustainable policy (SPUR, 2011). Some advocacies partner with the GFN to promote sustainable practices. For example, the Slow Food organization collaborates with GFN to promote sustainable food systems (SFI, n.d.). Scientists who criticize the EF have alleged it has eight shortcomings as a sustainability metric (Van den Bergh & Grazi, 2014). In response to such criticisms, some have proposed changes to the EF (Mancini et al., 2016).

Researchers have evaluated the EF as a measure of humans' demands on natural ecosystems; some have proposed revisions. For example, Jóhannesson et al. (2019) recommend improvements to EF's marine component. Giampietro and Saltelli (2014) fault the EF Carbon Footprint equation for the unclear figures for its variables. While Li et al. (2022) find the EF useful, they propose accounting for complexities such as mixed land uses.

Further study of the participants and their relationships will indicate the GFN's impact on sustainability. The effects of the EF's limitations on GFN's sustainability efforts also merit attention. The effect of the eight proposed limitations of the EF on sustainability measures will be assessed. The response of Wackernagel, the codeveloper of the EF, to its alleged faults will also be studied.

References

Anders and Huber (2010). Differential expression analysis for sequence count data. *Genome Biology*, *11*(10). Springer Nature.

Berlin, L., Schattman, R., and Hamilton, J. (2012, 10 Dec.). Working Toward the Common Table: The Policy and Program Implications of Vermont's Unified Approach to Social Justice, Food Insecurity, and Local Food. *Journal of Hunger and Environmental Nutrition*, *7*(4), 426–435. Web of Science.

DGEA. (n.d.). Differential Gene Expression Analysis. Training-Modules. https://hbctraining.github.io/Training-modules/planning_successful_rnaseq/lessons/sample_level_QC.html

Frismantiene et al., (2018). Smooth muscle cell-driven vascular diseases and molecular mechanisms of VSMC plasticity. *Cellular Signaling, 52,* 48-64. Science Direct.

GFN (n.d.a). Global Footprint Network. About Us. https://www.footprintnetwork.org/about-us/

GFN (n.d.b). Global Footprint Network. Earth Overshoot Day. https://www.footprintnetwork.org/our-work/earth-overshoot-day/

GFN (n.d.c). Global Footprint Network. Our History. https://www.footprintnetwork.org/about-us/our-history/

GSEA. (n.d.). https://www.gsea-msigdb.org/gsea/index.jsp

Khera and Kathiresan (2017). Genetics of coronary artery disease: Discovery, biology and clinical translation. *Nature Reviews Genetics*, *18*(6), 331–344. Web of Science.

Langfelder and Horvath (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559. Web of Science.

Langfelder and Horvath (2011). Is my network module preserved and reproducible? *PLoS Computational Biology*, *7*(1), e1001057. PubMed.

Liu and Eicher-Miller (2021). Food Insecurity and Cardiovascular Disease Risk. *Current Atherosclerosis Reports*, *23*(6), 24. NCBI.

Mancini et al. (2016). Ecological Footprint: Refining the carbon Footprint calculation. *Ecological Indicators*, *61*, 390–403. Science Direct.

Mi et al. (2019). Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system. *Nature Protocols, 14*, 703-721. NCBI.

Office (2006). *Switzerland's ecological footprint—A contribution to the sustainability debate | Publication*. Federal Statistical Office. https://www.bfs.admin.ch/bfs/en/home/statistics/catalogues-databases/publications.assetdetail.343230.html

SFI (n.d). Slow Food International. Foodnected—Funded Projects. https://www.slowfood.com/what-we-do/funded-projects-slow-food/foodnected/

Spiertz (2009). Nitrogen, Sustainable Agriculture and Food Security: A Review. *Sustainable Agriculture*, 635-651. SpringerLink.

SPUR (2011). San Francisco Planning and Urban Research Association. *Measuring San Francisco's Ecological Footprint*. (2011, July 5). SPUR. https://www.spur.org/news/2011-07-05/measuring-san-franciscos-ecological-footprint

Subramanian et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, *102*(43), 15545–15550. PubMed.

Van Den Bergh and Grazi (2014). Ecological Footprint Policy? Land Use as an Environmental Indicator. *Journal of Industrial Ecology*, *18*(1), 10–19. Science Direct.

Van der Harst Pim and Verweij Niek (2018). Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circulation Research*, *122*(3), 433–443. PubMed.

Vercammen et al. (2021). Longitudinal Analysis of Food Insufficiency and Cardiovascular Disease Risk Factors in the Coronary Artery Risk Development in Young Adults Study. *American Journal of Preventive Medicine*. PubMed.

Wackernagel and Pearce (2018). Day of reckoning. *New Scientist*, *239*(3189), 20–21. Science Direct.

Giampietro and Saltelli (2014). Footprints to nowhere. *Ecological Indicators*, *46*, 610–621. Science Direct.

Jóhannesson et al. (2019). Increasing the accuracy of marine footprint calculations. *Ecological Indicators 99* (Apr.), 153-158. Science Direct.