# DETECTING AND CHARACTERIZING PLEIOTROPY: NEW METHODS FOR UNCOVERING THE CONNECTION BETWEEN THE COMPLEXITY OF GENOMIC ARCHITECTURE AND MULTIPLE PHENOTYPES

Anna L. Tyler

*The Jackson Laboratory*
*Bar Harbor, ME, USA*
*E-mail: anna.tyler {at} jax.org*

Dana C. Crawford

*Department of Molecular Physiology & Biophysics,*
*Vanderbilt University*
*Nashville, TN 37240, USA*
*E-mail: dana.c.crawford {at} vanderbilt.edu*

Sarah A. Pendergrass

*Department of Biochemistry and Molecular Biology, Pennsylvania State University,*
*Pennsylvania State University*
*University Park, PA 16802, USA*
*E-mail: sap29 {at} psu.edu*

## Introduction

Pleiotropy, the phenomenon in which one gene influences more than one phenotype, was first defined over 100 years ago by Ludwig Plate.[1] Since that time our understanding of pleiotropy has changed and expanded to incorporate knowledge of molecular genetics. With this increase in knowledge has come an increase in an appreciation for the importance of pleiotropy in human health and evolutionary dynamics, but also a corresponding increase in confusion about how pleiotropy should be measured, and how the context in which pleiotropy is measured affects its interpretation. The purpose of this session is to offer a general view of pleiotropy and of the different approaches used to study this phenomenon. During the session, we will be examining a series of questions that are currently being discussed in the literature:

- Which genetic elements can be defined as pleiotropic?
- How are phenotypes defined and counted?
- How prevalent is pleiotropy? Does every gene affect every phenotype, or is pleiotropy more limited?
- How does pleiotropy influence, and how is it influenced by, evolution?
- How does an understanding of pleiotropy improve our understanding of human health?

### *Defining Genetic Elements*

Pleiotropy requires defining a genetic element that affects multiple phenotypes. But which genetic element is appropriate? A gene, a chromosomal segment with high linkage disequilibrium, a mutation? Plate's original definition of pleiotropy came long before the discovery of

DNA, and referred to a "unit of inheritance".[1] A unit of inheritance may refer to a single nucleotide polymorphism (SNP), or a gene, or a large segment of the genome containing multiple genes. Individual mutations may affect a single gene, or multiple genes.[2,3] In this session, the problem is addressed in a variety of ways. **Darabos et al.** use SNPs and genes, while **Philip et al.** use expression quantitative trait loci (eQTL) as genetic elements. The choices made in different experiments clearly have implications for the interpretation of pleiotropy, although the implications of these choices is still being debated.

### Defining Phenotypes

The concept of counting phenotypes is perhaps an even more difficult issue than identifying genetic elements. As Wagner and Zhang point out, a biologist may see two traits, femur length and tail length, where a mathematician familiar with rotation of coordinate systems my see only one: FeTail.[3] Further disagreements may arise as to whether two correlated traits such as femur length and femur width are one trait or two. Finally, there is discussion about whether the relationship between traits represents yet another phenotype that can be affected by genetic manipulation. Relationship QTL, or rQTL, which change the relationships between phenotypes have been identified in mice,[4,5] and are likely present in other organisms as well.

To reduce the number of subjective choices, some studies, such as **Philip et al.** in this session, use mRNA expression levels as phenotypes. mRNA expression is relatively easy to measure comprehensively and at scale. However, problems such as high-dimensional data with relatively few samples and correlation between phenotypes arise in these studies. One class of methods used to address these problems is dimensionality reduction, and in this session **Philip et al.** discuss one such dimensionality reduction approach.

Measurement of physiological traits in humans is addressed by **Hall et al.**. This paper examines the concept of standardized, high-througput phenotyping in a range of medically relevant areas from physiological measurements available in electronic medical records to environmental exposures.

### The Prevalence of Pleiotropy

In addition to specific relationships between individual genes and phenotypes, many studies of pleiotropy are concerned with quantifying pleiotropy itself. Ronald Fisher promoted the idea of "universal pleiotropy" in which every gene affects every phenotype to some extent either directly or indirectly.[6] This idea was implicit in his geometric model of adaptation.[6,7] However, since the 1930's molecular genetics experiments have revealed a more modular model of pleiotropy.[8] In modular pleiotropy, gene actions are limited to a specific set of processes or phenotypes and are relatively independent from other phenotypes[3,9]

Modular pleiotropy is supported in the literature. Wagner and Zhang[3] review the results from experiments in yeast, nematode and mouse, using a variety of methods of counting genetic elements and phenotypes. In each of these experiments, the vast majority of genetic elements affect very few phenotypes, while only a few elements affect a large number of phenotypes. These distributions are surprisingly consistent across the different experiments. This limited scope of the majority of genes supports the hypothesis that gene action is relatively limited to

phenotypes modules. The paper by **Darabos et al.** presented in this session shows additional evidence in support of a modular view of pleiotropy.

### Pleiotropy in an Evolutionary Context

Whether pleiotropy is universal or modular has an impact on how pleiotropic genes are influenced by selection. Complex organisms have vastly more cell types than prokaryotes, but only about four-fold more genes.[10] The necessary increase in pleiotropy per gene that this statistic suggests could limit the evolvability in complex organisms due to potentially wide-spread effects of single mutations. Des Marais and Rausher[11] have proposed that gene duplication may provide an escape from these evolvability limitations, as each gene copy can take over a subset of the original gene's functions. Other studies have addressed molecular mechanisms by which genes evolve to be more pleiotropic. This process may preferentially recruit genes to new biological processes rather than adding new biological functions.[3] For example, new processes might include changes in tissue expression, subcellular localization, interacting partners and context-sensitive transcription.[3]

In this session, the relationship between pleiotropy and evolution will be addressed by the keynote speaker, **James Cheverud** in his talk titled "Genetic Variation and Evolution of Pleiotropy."

### Pleiotropy and Human Health

The importance of pleiotropy in human health is undeniable. Pleiotropy coupled with dynamic networks that exist between the genetic architecture, signaling pathways, intermediate phenotypes, and outcome traits can be an important part of health and disease and may become important for network-based medicine.[12] Phenomics, phenome scans, and phenome-wide association studies may provide a high-throughput way for exploring both pleiotropy and the diseasome.[13–19] Identifying genetic variation that confers both protection for some traits/outcomes but risk for others may both highlight important genetic regions, and also show important features of larger biological networks. Knowing which genes influence which phenotypes may aid in drug repurposing for genetically related diseases, as well as predicting off-target effects of targeted therapies. All papers in this session address human health either directly or indirectly. **Philip et al.** investigate QTL that interact to affect kidney health in a mouse model of kidney disease. **Darabos et al.** explore the relationships between SNPs, genes and pathways, and phenotypes to show novel molecular relationships between human diseases. And finally **Hall et al.** discuss methods of standardized, high-throughput phenotype measurement in patients with type 2 diabetes (T2D).

### Session Contributions

The keynote lecture for this session will be given by **James Cheverud** who has worked extensively on pleiotropy and the evolution of pleiotropy in mammals. He will speak on the "Genetic Variation and Evolution of Pleiotropy."

**Philip et al.** investigate epistasis and pleiotropy at the transcript level in an F2 mouse cross designed to examine kidney function. This paper presents a method called the Com-

bined Analysis of Pleiotropy and Epistasis (CAPE) which combines information across multiple phenotypes to infer directional interactions between genetic variants. This method has previously been used to examine pleiotropy related to physiological traits and now focuses on pleiotropy at the level of transcription. The authors found loci on eight chromosomes that interact to influence three expression modules. This method was further able to distinguish between which markers are truly pleiotropic and affect more than one module, and which are indirectly pleiotropic, affecting multiple modules through interactions with other genetic loci. This paper directly addresses several open issues in pleiotropy research including dimension reduction for high-dimensional phenotype spaces and distinguishing direct pleiotropy from indirect pleiotropy.

**Darabos et al.** also combine analysis of epistasis and pleiotropy. This paper constructs a bipartite network of genetic elements and phenotypes reported in GWAS data and other public repositories in a method similar to that used to construct the human diseasome.[20] However, unlike the diseasome, the network constructed by **Darabos et al.** includes non-disease phenotypes, such as hair color, as well as risk-associated SNPs that fall outside of coding regions. The authors constructed networks at three different levels of resolution: SNPs, genes, and pathways. These networks show that most genes have limited pleiotropic effects, supporting a model of modular pleiotropy. The pathway-base network also proves to be particularly informative and shows well established links between glaucoma and blood pressure, as well as glaucoma and type 2 diabetes. The network also shows a novel relationship between glaucoma and Alzheimer's disease, a connection that has only recently begun to be investigated. This paper shows the powerful predictions in human health that can be made by taking into account both epistasis and pleiotropy.

**Hall et al.** conduct and environment-wide association study (EWAS) to investigate contributions of environmental exposures and lifestyle choices to type 2 diabetes (T2D) in a high-throughput manner. The study employs a combination of resources, including electronic medical health records, the PhenX toolkit for standardized exposure measurement, and the Diet History Questionnaire. The authors find that moderate alcohol use is associated with decreased risk of T2D, and that low amounts of activity during leisure time, as well as smoking are positively associated with T2D. These relationships replicated in two independent populations and are supported by previous literature. This paper demonstrates the importance and practicality of standardized, high-throughput measurements of human phenotypes and environmental exposures, a field that is critical to further study of pleiotropy in humans.

### Acknowledgments

### References

1. F. W. Stearns, *Genetics* **186**, 767 (2010).
2. J. Hodgkin, *The International Journal of Developmental Biology* **42**, 501 (1998).
3. G. P. Wagner and J. Zhang, *Nature Reviews Genetics* **12**, 204 (2011).

4. J. M. Cheverud, T. H. Ehrich, T. T. Vaughn, S. F. Koreishi, R. B. Linsey and L. S. Pletscher, *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution* **302**, 424 (2004).
5. M. Pavlicev, J. P. Kenney-Hunt, E. A. Norgard, C. C. Roseman, J. B. Wolf and J. M. Cheverud, *Evolution* **62**, 199 (2008).
6. R. A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon, 1930).
7. H. A. Orr, *Evolution* **54**, 13 (2000).
8. J. J. Welch and D. Waxman, *Evolution* **57**, 1723 (2003).
9. P. Mitteroecker, *Evolutionary Biology* **36**, 377 (2009).
10. N. Lane and W. Martin, *Nature* **467**, 929 (2010).
11. D. L. Des Marais and M. D. Rausher, *Nature* **454**, 762 (2008).
12. A.-L. Barabási, N. Gulbahce and J. Loscalzo, *Nature Reviews Genetics* **12**, 56 (2011).
13. R. M. Bilder, F. Sabb, T. Cannon, E. London, J. Jentsch, D. S. Parker, R. Poldrack, C. Evans and N. Freimer, *Neuroscience* **164**, 30 (2009).
14. M. B. Lanktree, R. G. Hassell, P. Lahiry and R. A. Hegele, *Journal of Investigative Medicine* **58**, 700 (2010).
15. D. Houle, D. R. Govindaraju and S. Omholt, *Nature Publishing Group* **11**, 855 (2010).
16. A. Rzhetsky, D. Wajngurt, N. Park and T. Zheng, *Proceedings of the National Academy of Sciences* **104**, 11694 (2007).
17. N. Ghebranious, C. A. McCarty and R. A. Wilke, *Personalized Medicine* **4**, 175 (2007).
18. S. A. Pendergrass, K. Brown-Gentry, S. Dudek, A. Frase, E. S. Torstenson, R. Goodloe, J. L. Ambite, C. L. Avery, S. Buyske, P. Bůzkŏvá *et al.*, *PLoS Genetics* **9**, p. e1003087 (2013).
19. J. C. Denny, M. D. Ritchie, M. A. Basford, J. M. Pulley, L. Bastarache, K. Brown-Gentry, D. Wang, D. R. Masys, D. M. Roden and D. C. Crawford, *Bioinformatics* **26**, 1205 (2010).
20. K. I. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal and A. L. Barabasi, *Proceedings of the National Academy of Sciences* **104**, 8685 (2007).