

A fundamental challenge in biology is to identify mechanisms that link genotype and phenotype in complex natural environments. **I am most interested in how gene networks and physiological systems react to changes in the environment, and how the structure of these systems may be a constraint on evolutionary adaptations.** My approach to these questions combines experimental quantitative genomics with statistical and computational modeling. Through this work, I hope to advance our understanding of the genetic and mechanistic basis of variation among plant varieties, and to improve our ability to predict the responses of species to environmental disturbances and climate change.

I am trained as a biologist and as a statistician, and have experience in molecular biology, quantitative genetics, genomics, and crop modeling. In my PhD, I studied how gene networks and development shape responses to climate change in sea urchins by applying quantitative genetics tools to gene expression data. At the same time, I completed a MS in statistics and developed methods to study the genetic architecture of high-dimensional genomic traits, and to identify links between these traits and morphology. In my postdoctoral training, funded by the NSF Postdoctoral Fellowship in Biology, I am expanding my work to study the dynamics of gene networks in complex environments. For this, I have switched to a focus on plants, particular *Arabidopsis thaliana*, where I can draw on the long history of research on mechanisms underlying plasticity and responses to the environment.

Looking forward, my future research will continue to build towards a systems perspective of genotype-phenotype relationships in plants under real-world conditions. I am particularly excited to apply these approaches to agriculturally important systems through the University of California Agricultural Experiment Station. My research will focus on the following areas:

- The interactions between physiology, growth and developmental signals in plants
- Statistical inference of genetic variation in gene networks
- Developing models of phenotypic and evolutionary responses to climate change

Below, I briefly describe highlights of my past and current work, and outline plans for my future independent career.

Previous and Current Research

How do gene networks shape plasticity to novel environments? Regulatory genes and signaling molecules operate in gene networks to control development and responses to the environment. However, we know little about how genetic variation operates through gene networks to produce phenotypic variation in natural populations. A major focus of my dissertation research with Greg Wray at Duke University was to test how a developmental gene network that drives early morphogenesis in sea urchins shapes phenotypic variation in different environments. Using a quantitative genetics approach, I measured gene expression variation in this network to address the question: *Do gene networks influence the expression of genetic variation under realistic environmental stresses?* I documented substantial genetic variation and plasticity in the expression of critically important developmental genes in the network, and significant genetic correlations between the expression profiles of directly interacting genes

(Runcie et al *Mol Ecol.* 2012, Garfield, Runcie et al *PloS Biology* 2013). Surprisingly, gene expression responses to high temperatures did not follow the network topology. This demonstrates that the gene network function was effectively buffered across a moderate temperature range. Mechanisms that buffer gene network outputs are therefore critical for understanding gene-trait relationships in different environments.

Which molecular traits are most important for adapting to changing climates? The traits that are the most important for tolerating novel environmental conditions often involve aspects of cell biology or development that are difficult to measure directly. However, knowing if these traits are genetically variable may be an important predictor of whether or not a species can rapidly adapt to climate change. In collaboration with Sam Dupont and Mike Thorndyke at the University of Gothenburg, Sweden, I used RNAseq to identify molecular traits that are involved in the response of sea urchin embryos to a decline in seawater pH, called ocean acidification, and to test if these molecular traits were variable in a Northern European population (Runcie et al *in prep*). I used a combined bioinformatics/statistics approach and identified signatures of variation in various aspects of membrane composition, immune responses, DNA replication and metabolism that were affected by low pH seawater. I found evidence for genetic variation in some, but not all of these traits. This suggests potential constraints, and possible avenues of evolutionary rescue, for this species in response to climate change.

A novel statistical method for evolutionary quantitative genomics: Genomic assays of gene expression offer the possibility of studying evolution in molecular traits such as gene networks or development. However, the immense number of gene expression traits is too large for traditional analytical approaches used in quantitative genetics to be useful. In collaboration with Sayan Mukherjee in the Statistics Department at Duke, I developed a statistical model called *Bayesian Sparse Factor Analysis of Genetic Covariance Matrices*, or BSFG, suitable for estimating genetic correlations at the transcriptome-scale (Runcie and Mukherjee, *Genetics*, 2013). This model is motivated by the idea that gene expression traits are modular, and therefore we can increase our power by looking explicitly for highly correlated groups of genes. Our model produces robust estimates of genetic covariance matrices (G-matrix) for thousands of traits at once. At the same time, our model identifies highly correlated gene sets that may be signatures of variation in important molecular and developmental traits and quantifies their heritability. This model is implemented in freely available software.

Dynamic model of an Arabidopsis developmental network in complex environments: Plants in the field face multiple simultaneous environmental signals that can occur at unpredictable times during development. Two complementary approaches are used to model the developmental response of plants to the environment: physiology-based “crop models”, and genetic models of gene interactions. As part of my postdoc, I aim to bridge these two approaches by testing how the expression of central integrator genes linking flowering time to temperature and light signals in *Arabidopsis* correlate with crop model parameters. I am extending an existing physiology-based model of flowering time control in *Arabidopsis* and using RNAseq to measure gene expression signatures of these genes in response to controlled environmental perturbations. Based on this model, I am testing how: a) the timing of winter during development, and b) diurnal variations in temperature, impact gene networks that controlling the timing of flowering.

Future directions

The research focus of my lab will continue to draw on statistical, modeling, and genomics approaches to study phenotypic variation in complex environments. Extending my postdoctoral research, the central model system for my lab will be the physiology and phenology of *Arabidopsis*. But, the modeling and genomic tools I develop will be applicable to crop systems as well. Below, I describe four examples of lines of research I envision my lab pursuing.

The interactions between physiology, growth and developmental signals in plants: Plant development (the creation of new structures) is often modeled independently from plant growth (the expansion of existing structures). However, molecular mechanisms that regulate development can alter patterns of growth, and the output of growth may feed back on developmental pathways. In this way, the timing of developmental transitions such as flowering is carefully regulated in response to both external and internal signals. Across natural accessions of *Arabidopsis*, there are strong genetic correlations between the timing of flowering and physiological parameters of growth. However, little is known about how the genetic pathways known to regulate flowering interact with the physiological state and size of a plant. I will build a spatially explicit physiological model of plant growth that defines how various plant tissues interact. For example, leaf growth may indirectly impact the timing of flowering: more, or larger, leaves may increase the total concentration of leaf-produced signals that are integrated by the meristem to initiate flowering. I will parameterize this model by measurements of plant morphology, photosynthesis and gene expression in leaves and meristem tissue under simulated seasonal environments in which temperature, daylength, and water availability are simultaneously varied. I expect that such a model will greatly improve predictions of phenology across a wider range of genotypes and environments than any model solely focusing on physiology or gene networks alone. I am submitting a pre-proposal as a PI to the NSF IOS division in January to fund this work.

Statistical inference of genetic variation in gene networks: *Arabidopsis* provides a powerful system to link natural genetic variation and molecular mechanisms of environmental response. Drawing on the dynamic gene networks models I am developing in my postdoc, I will investigate the genetic architecture of the activities of flowering development pathways. I will use genome-wide associations (GWAS) with existing field data of flowering phenology to identify genetic variants associated with the parameters of my gene network models. This project will test the hypothesis that natural variation in the response of flowering to complex environments results from genetic effects on genetic pathway activities.

Developing models of phenotypic and evolutionary responses to climate change: By explicitly modeling the link between genetic variation and the environment, process-based models like I am developing for *Arabidopsis* are powerful tools to predict phenotypes of plants under the novel environments produced by climate change. Using the GWAS framework described above, I will use genomic prediction techniques to model the interactions of the developmental pathways in different populations across environmental gradients and thus predict flowering phenology under the projected climatic regimes.

Applications to crop phenology: I will use my appointment at the Agricultural Research Station to extend my work to developing genomics tools for crop improvement. The phenology of fruit production is a critical factor determining value and yield in many crops. Homologs of many genes that regulate flowering phenology in *Arabidopsis* are known to operate in similar pathways in crop plants. Drawing on my experience translating genomic tools to non-model sea urchin species as part of my dissertation research, I will use gene expression profiling of mapping populations of crop plants planted in different seasonal environments to identify developmental and physiological mechanisms that can be targeted for selection and crop improvement.

Summary

My work through my graduate career and into my current postdoc has developed a framework for understanding links between genetics, environmental change, and phenotypic variation. The key to this link is developing genomic and statistical tools to assay and model how developmental systems integrate environmental signals. My current and future work with *Arabidopsis* aims to build dynamic models of genetic pathways that can link natural genetic variation and flowering phenology in the field. This work will provide new insights into differences in performance among plant varieties in both current and novel environments.