Integrative and Comparative Biology

Integrative and Comparative Biology, volume 60, number 2, pp. 261–266
doi:10.1093/icb/icaa073

Society for Integrative and Comparative Biology

SYMPOSIUM INTRODUCTION

Building Bridges from Genome to Phenome: Molecules, Methods and Models—An Introduction to the Symposium

Karen G. Burnett,[1,*] David S. Durica,[†] Donald L. Mykles[‡] and Jonathon H. Stillman[§]

*Grice Marine Laboratory, College of Charleston, 205 Fort Johnson Rd, Charleston, SC 29412, USA; [†]Department of Biology, University of Oklahoma, 730 Van Vleet Oval, Norman, OK 73019, USA; [‡]Department of Biology, Colorado State University, 1878 Campus, Fort Collins, CO 80523, USA; [§]Department of Biology, San Francisco State University, San Francisco, CA 94123, USA

From the symposium "Building Bridges from Genome to Phenome: Molecules, Methods and Models" presented at the annual meeting of the Society for Integrative and Comparative Biology, January 3–7, 2020 at Austin, Texas.

[1]E-mail: burnettk@cofc.edu

Synopsis How stable genotypes interact with their environment to generate phenotypic variation that can be acted upon by evolutionary and ecological forces is a central focus of research across many scientific disciplines represented within SICB. The Building Bridges Symposium brought together scientists using a variety of organisms, methods, and levels of biological organization to study the emergent properties of genomes. Workshops associated with the Symposium aimed to identify the leading edges and major barriers to research in this field, and to recommend future directions that might accelerate the pace of progress. The papers included in this Symposium volume draw attention to the strength of using comparative approaches in non-model organisms to study the many aspects of genotype–environment interaction that drive phenotype variation. These contributions and the concluding white paper also illustrate the need for novel conceptual frameworks that can bridge and accommodate data and conclusions from the broad range of study systems employed by comparative and integrative biologists to address genome-to-phenome questions.

Introduction to the Building Bridges Symposium

How genomes give rise to whole organisms with a variety of complex phenotypes and how these processes are linked to organismal function and performance in response to ecological and evolutionary forces are among the grand challenges articulated by the scientific community (Schwenk et al. 2009) and recognized by the National Science Foundation (NSF). In the last few years, rapid advancements in technologies for genome and transcriptome sequencing, gene editing, epigenomics, proteomics, and metabolomics have made it feasible and affordable to move from studying single genes and their functional importance in a few model organisms to a more complex understanding of how cellular networks maintain the integrity of species identity in non-model organisms, while allowing for plasticity in response to environmental change.

Recognizing that the availability of these technological resources would widely impact the study of organismal science, NSF funded the Animal Genome to Phenome Research Coordination Network (AG2P RCN, https://ag2p.net) in 2015. With support from the AG2P RCN and NSF Symposium funding, we organized a Society-wide Symposium entitled "Tapping the Power of Crustacean Transcriptomics to Address Grand Challenges in Organismal Biology" at the 2016 Annual Meeting of the Society for Integrative and Comparative Biology (SICB). Peer-reviewed papers based on invited presentations were published in *Integrative and Comparative Biology* (Vol. 56) in 2016 along with a white paper based on the 2016 symposium workshop (Mykles et al. 2016). That white paper emphasized the need to standardize and lower the barriers for large-scale analysis of transcriptomic data, so that scientists from many disciplines and research environments

Advance Access publication June 13, 2020

could develop consensus approaches as the "taking off point" for linking genomic with phenotypic diversity.

In the subsequent 4 years, there has been explosive growth in the diversity of models, methods, and computational tools being used to explore the processes that lead from the genotype to the phenotype of individuals, species, or biological communities. Yet, the central questions remain: how can we understand the functional context of such "big" data within the intact organism, and how does it contribute to phenotype?; and, also, what methods and approaches have been most successful in bridging the gap from genome to phenome?

The 2020 SICB Society-wide Symposium "Building Bridges from Genome to Phenome: Molecules, Methods and Models" brought together scientists taking a variety of approaches to address these questions in many different biological systems. The invited talks and the symposium's complementary oral and poster paper sessions delved into recent progress linking phenotype plasticity to changes at the level of the genome, epigenome, and proteome, while exploring the boundaries between variation and speciation. The purpose of this collection of papers from the Symposium is to highlight some of the new approaches, models, and methodological challenges currently faced as investigators try to meld and derive predictive power from different types of datasets (genomic, transcriptomic, epigenetic, proteomic, metabolomic) to understand the framework by which the phenome arises from the genomes of individuals, species, consortia, and communities. As reflected in the breadth of topics and workshops addressed in this resulting symposium volume, participants aimed to assess strengths and weaknesses of current approaches, identify common barriers that inhibit progress and the resources that are needed to overcome those barriers.

Symposium topics

Even though they came from many disciplines with a wide variety of questions and research approaches, investigators were able to identify common research themes, which were articulated at an organizational workshop prior to the Symposium. Current genome-to-phenome research generally builds on the classic, well-studied framework of quantitative genetics—that genetic variation (G), environmental variation (E), and their interaction (G × E) give rise to phenotypes. The G × E interaction itself is sensitive to ontology (development, aging, diapause, molting), and may be influenced by spatial and temporal variation, giving rise to phenotypes that may be stable or plastic and which can be acted upon by natural selection.

Further, phenotypic variation may not emerge solely as a result of variation at a single genetic locus but can be altered by epigenetic modification. In this Symposium volume, Carneiro and Lyko (2020) review currently understood epigenetic mechanisms that modulate gene expression in animals and how the resulting phenotypic plasticity may facilitate multiple, rapid adaptations to changing environments. In order to minimize the influence of genetic polymorphisms in driving phenotypic changes, these investigators use a model organism—the marbled crayfish, whose mode of parthenogenetic reproduction results in genetically homogeneous, monoclonal populations. Despite their lack of genetic variation, the species is highly invasive across many distinct ecological niches. Carneiro and Lyko (2020) provide evidence linking epigenetic mechanisms such as DNA methylation patterns to adaptability, and thereby invasiveness, in this organism. This contribution illustrates the power of model organisms to uncover mechanisms driving genome-to-phenome phenomena while recognizing specific pitfalls in conducting and interpreting epigenome data.

Phenotypic variation arising from a single genetic locus may also be driven by variation at other genetic loci (epistasis). In the case of mitonuclear epistasis, different phenotypes may arise when the same nuclear genome is placed against different mitochondrial genotypes. The environment can further modify mitonuclear epistasis (G × G × E interactions). Assessing how mitonuclear epistasis contributes to organismal phenotype and, consequently, fitness in natural populations is complicated by the impact of sexual reproduction. In their contribution to this volume, Greimann et al. (2020) take advantage of distinct features of the New Zealand freshwater snail *Potamopyrgus antipodarum* to decouple the effects of sexual reproduction from mitochondrial genome inheritance on mitochondrial and physiological function in naturally occurring sexual and asexual snail populations. The lakes inhabited by *P. antipodarum* span wide environmental gradients, with substantial across-lake genetic structure and mitonuclear discordance. Greimann et al. (2020) integrate cellular, physiological, and behavioral approaches to quantify variation in mitochondrial function across a diverse set of wild *P. antipodarum* lineages, demonstrating the use of this model to dissect phenotypic variation arising from the complex relationships among mitonuclear variation, performance, plasticity, and fitness in natural populations. Their work demonstrates the

importance of considering variation in the mitochondrial genome together with the nuclear genome when building predictive linkages to phenotype, and for careful consideration of phenotypic outcomes across a range of biological scales.

Phenotypic variation may also arise from the G × G × E impacts of multiple species living as endosymbionts, consortia, or assemblages; such relationships might drive local adaptation, evolutionary innovation, and diversification. For example, components of a microbiome, their organization, and interactions could have profound influence on the emergence of the host phenotype and its plasticity to environmental change. In this symposium volume, Hoffman et al. (2020) examined the specific consortia present in the phenotypically discernible layers of the orange crust endemic to particular habitats of the Hawaiian anchialine ecosystem. They found that distinct microbial consortia (i.e., genotypes) appear to have independently assembled multiple times toward a common physical appearance and lamination, with gross similarity in environmental conditions among and within habitats having apparently driven convergent evolution in the functional groups (i.e., phenotypes) recovered per layer between different sites and islands.

Elsewhere in this volume, Kirschman et al. (2020) show that the species composition of microbial gut flora and the genetic background of the host can have strong effects on host physiology and the development of a normal phenotype in the early life of a vertebrate host. These investigators tested the hypothesis that microbial colonization and the genetic background of the host would affect survival, immune gene expression, growth, and development in two populations of three-spined stickleback fish (*Gasterosteus aculeatus*), one anadromous and one freshwater. Their results provide evidence that disruption of microbial flora, coupled with pathogen exposure in early life, can alter host development. The genetic background of the host may also play a role, resulting in idiosyncratic effects on phenotype. These two examples illustrate that emergent phenotypes likely develop from a framework of multiple co-existing genotypes, interacting among themselves and with their environment over time.

Many current investigations are building on the G × E central framework to incorporate other molecular and physiological/developmental data to put more predictive power into the link between genome and phenome. In this Symposium volume, Li and Kültz (2020) argue that proteomes represent critical links that determine how genomes interact with the environment, thereby giving rise to phenotypic variability that can be acted upon by natural selection. In the work presented here, the authors isolate and identify the influence of salinity versus the general stress response in the three-spined stickleback fish *G. aculeatus* by monitoring the gill proteomic response profiles of animals held in mesocosms under controlled conditions with appropriate controls. Their results demonstrate that *G. aculeatus* responds to salinity changes by adjusting osmoregulatory mechanisms such as compatible osmolyte synthesis, transepithelial ion transport, and oxidative energy metabolism, mechanisms that are distinct from transient general stress responses. Furthermore, Li and Kültz (2020) establish salinity as a key factor for causing the regulation of numerous proteins and Kyoto Encyclopedia of Genes and Genomes pathways with established functions in proteostasis, immunity, and tissue remodeling. Their results establish the potential predictive power of proteomics in monitoring phenotypic variation.

Most, if not all, the research presented at the Symposium required that investigators develop or implement software tools and computational methods to merge large, complex datasets across different levels of biological organization. While each dataset captures a distinct layer of biological complexity, such as the genome, epigenome, transcriptome, or proteome, the researchers structured the computational approaches to their data analyses to reproducibly link those layers into networks that track phenotypic change and thereby reveal the mechanisms that may underlie those changes. Exemplifying efforts to design new and better computational approaches for network analysis, Schaefer et al. (2020) demonstrate the application of a new, free, open-source software package, Camoco (Co-analysis of molecular components), to integrate genome-wide association studies with information on gene co-expression networks in the domestic horse. Schaefer et al. (2018) previously implemented Camoco in maize (*Zea mays*), demonstrating that this approach is generalizable to any species with relevant gene expression data, even those with no other data on gene function. In their manuscript included in the present volume, Schaefer et al. (2020) argue for the ongoing need to develop software tools that are reusable to address many biological research questions, can generate reproducible outcomes, and can be made available to users on demand.

Garrett et al. (2020) examine G × E interactive effects on phenotype and their evolutionary potential by using pooled genomic DNA sequencing to look for shifts in allele frequencies in response to

environmental perturbation. These investigators ask whether species like the purple sea urchin, *Strongylocentrotus purpuratus*, has enough genetic variation to yield adaptive phenomes that can survive and grow under ocean acidification conditions that may be associated with global climate change. Garrett et al. (2020) reared *S. purpuratus* larvae under static and variable levels of moderate and extreme pH conditions, then measured survival, growth, and genome-wide allele frequency shifts. They found consistent, directional shifts in genome-wide allele frequencies, as well as distinct patterns and trade-offs in survival and growth under the experimental acidification regimes. These outcomes suggest that *S. purpuratus* populations possess genetic variation for population persistence under extreme pH conditions.

Given the complexity of highly integrated organismal systems, it is not surprising that trade-offs, like those observed by Garrett et al. (2020), emerge in genome-to-phenome investigations across levels of biological organization and sub-disciplines, including physiology, ecology, and evolution. In their contribution to this symposium volume, Mauro and Ghalambor (2020) present a broad overview of the conceptual frameworks used to study and understand the mechanistic basis for trade-offs and how these trade-offs ultimately constrain adaptive evolutionary change. They argue that our understanding of the genetic underpinning of traits depends on environment (ecology) and on selection in response to environmental change (evolution). Advances in molecular biology, omics, and systems/network analysis are making it possible to generate and experimentally evaluate such genome to phenome to fitness maps in biological systems.

Noting the paucity of molecular tools for testing the genetic foundation for many physiological adaptations, especially in marine organisms, Lam et al. (2020) suggest that cellular modeling with primary cells offers a powerful system for uncovering and validating the functional significance of genetic changes in organisms in which transgenesis is impossible, such as federally protected marine mammals. The investigators review how genetic manipulation has advanced mechanistic investigations in other non-traditional mammalian species and outline key considerations for isolating, culturing, and conducting mechanistic experiments with marine mammal cells under conditions that mimic *in vivo* states. Lam et al. (2020) propose that primary tissue culture can be a critical tool for conducting functional studies to provide the missing link between genome- and organismal-level understanding of physiological adaptations in non-model organisms.

Further supporting the need to investigate environmental effects on functional adaptations across levels of biological organization, Iverson et al. (2020) present a meta-analysis of thermal performance responses across multiple scales, including enzyme activities, respiration of isolated mitochondria, and whole-animal metabolic rate. Thermal responses differed drastically across levels of biological organization, sometimes showing completely opposite patterns among levels for the same temperature ranges. Details across studies, including the taxa, enzymes, and tissues examined, generally did not affect the differences in thermal responses among biological levels. The authors suggest that physiological responses may differ across levels due to differing selection pressures or differing effects of biochemical laws. In any case, these results further emphasize the complexity in the G × E relationship: the environment can affect even similar phenotypes differently.

Finally, Gust et al. (2020) demonstrate an approach for bridging the genome-to-phenome gap, guided by the adverse outcome pathway (AOP), which is widely accepted and employed by toxicologists. In their example, the investigators aimed to identify the mechanisms underlying toxicological phenotypes of lethargy and weight loss in response to exposure to nitroaromatic munitions, such as 2,4,6-trinitrotoluene (TNT). Gust et al. (2020) tested the hypothesis that inhibition of peroxisome proliferator-activated receptor (PPAR)α signaling in TNT exposures was a molecular initiating event pathway that impacted lipid metabolic processes, thus affecting systemic energy budgets, ultimately resulting in body weight loss. Results from a series of transcriptomic, proteomic, lipidomic, *in vitro* PPARα nuclear signaling, and PPARα knock-out investigations ultimately supported the hypothesis. Given that PPARα antagonism represented a critical response in the AOP, a phylogenetic analysis was conducted which indicated that PPARα amino acid relatedness generally tracked species relatedness, providing context for extrapolating this genome-to-phenome relationship between species. Gust et al. (2020) suggest that such an overarching investigative framework guided by AOP principles might prove useful for assessing the drivers of phenotypic variation in other biological sciences.

Most recently, the grand challenge of linking genotype to phenotype is just one question being addressed within the NSF-sponsored initiative of Reintegrating Biology (https://reintegratingbiology.

org/)—an effort to fully understand biological systems and realize their potential. In 2019, a series of virtual town halls, jumpstart meetings, and workshops were conducted to involve the broader biological community in identifying new research questions that could be addressed by combining approaches and perspectives from different subdisciplines of biology, as well as to identify the key challenges, scientific gaps, physical infrastructure, and workforce training that are needed to answer these questions. As a consequence of one such jumpstart meeting in December 2019 in Austin, TX, Westerman et al. (2020) present their vision of challenges and solutions for utilizing big data to address genome-to-phenome questions. Among their recommendations are initiatives to establish minimum best practices for experimental design and data collection, tool development and sharing, to develop and maintain shareable data repositories, to promote training in data science, and to fund research that concurrently addresses both basic and applied questions, like climate change and preservation of natural resources.

The Building Bridges Symposium series concludes with a white paper co-authored by the symposium organizers (Burnett et al. 2020) with the input of more than 40 scientists who participated in a workshop on the last day of the SICB meeting. In small group discussions, participants evaluated the progress of current work, identified some of the most promising approaches and models presented or discussed in all of the symposium sessions, then singled out the leading edges as well as the key barriers to this research, and articulated resources that would be vital to driving progress toward elucidating genome-to-phenome phenomena. Major themes of the discussion were the continuing need for an arching theoretical framework for conducting genome-to-phenome research and stronger networks for identifying interdisciplinary collaborators and tools to approach such large-scale theoretical concepts.

Future directions

The Building Bridges Symposium brought together scientists from a variety of biological disciplines, each framing genotype to phenotype questions from a distinctive perspective and each using study organisms and techniques uniquely suited to address that working question. Informal and formal workshop discussions revealed a common recognition of

Fig. 1 Organizers and Speakers in the 2020 Building Bridges from Genome to Phenome Symposium. Left to right: Kurt Gust, Don Mykles, David Durica, Alex Mauro, Johnathon Li, Melissa Pespeni, Rob Schaefer, Karen Burnett, Justin Havird, Cameron Ghalambor, Scott Santos, Kat Milligan-Myhre, Joel Sharbrough. Not pictured: Dan Hahn, Dietmar Kültz, Jonathon Stillman.

the scientific strength that derives from such a broad range of comparative approaches. Participants also recognized common barriers to their research progress—for example, access to specific "omics" techniques and instrumentation, data analysis, handling, and storage. Yet, participants most commonly called for a renewed emphasis on the development of conceptual and integrative frameworks for understanding the principles underlying the emergence of phenotypic variation, with an eye toward being able to accommodate observations from the broad range of comparative study systems currently in use. Going forward, scientists and funding agencies should prioritize the development of collaborative, interdisciplinary working teams aimed at the theory and testing of simple "building blocks" for the emergence of phenotypic variation. Large multidisciplinary conferences, like SICB, would provide particularly good venues for such teams to meet and develop the kinds of novel frameworks that might serve as a bridges from genomes to phenomes.

Acknowledgments

The authors thank the invited speakers (Fig. 1), as well as speakers in the complementary oral and poster presentations for their participation in the Building Bridges Symposium and Workshops. The authors also extend their sincere appreciation to Program Officer Susan Williams and the Burk and Associates staff for their assistance in organizing the symposium.

Funding

This work was supported by the National Science Foundation (Symposium Award IOS-1927470) and the Animal Genome to Phenome Research Coordination Network (Award IOS-1456942). We gratefully acknowledge the support of the Society for Integrative and Comparative Biology, the SICB Divisions of Comparative Physiology and Biochemistry, Comparative Endocrinology, Evolutionary Developmental Biology, Ecoimmunology and Disease Ecology, Ecology and Evolution, Invertebrate Zoology, Phylogenetics and Comparative Biology as well as The Crustacean Society and American Microscopical Society.

References

Burnett KG, Durica DS, Mykles DL, Stillman JH, Schmidt CJ. 2020. Recommendations for advancing genome to phenome research in non-model organisms. Integr Comp Biol 60 (doi: 10.1093/icb/icaa059).

Carneiro VC, Lyko F. 2020. Rapid epigenetic adaptation in animals and its role in invasiveness. Integr Comp Biol 60 (doi: 10.1093/icb/icaa023).

Garrett AD, Brennan RS, Steinhart AL, Pelletier AM, Pespeni MH. Forthcoming 2020. Impacts of static, variable, and extreme pH on the genetic and phenotypic responses of purple sea urchin larvae. Integr Comp Biol 60.

Greimann ES, Ward SF, Woodell JD, Hennessey S, Kline MR, Moreno JA, Peters M, Montooth KL, Neiman M, Sharbrough J. 2020. Phenotypic variation in mitochondrial function across New Zealand snail populations. Integr Comp Biol 60 (doi: 10.1093/icb/icaa066).

Gust K, Ji Q, Luo X. 2020. Example of adverse outcome pathway (AOP) concept enabling genome-to-phenome discovery in toxicology. Integr Comp Biol 60 (doi: 10.1093/icb/icaa064).

Hoffman SK, Seitz KWS, Harvard JC, Weese DA, Santos SR. Forthcoming 2020. Phenotypic comparability arising from genotypic variability amongst physically structured microbial consortia. Integr Comp Biol 60.

Iverson E, Nix R, Abebe A, Havird JC. Forthcoming 2020. Thermal responses in biological rates differ across levels of organization. Integr Comp Biol 60.

Kirschman LJ, Khadjinova A, Ireland K, Milligan-Myhre K. Forthcoming 2020. Early life exposure to an enteric pathogen affects organ development in threespine stickleback raised in germ-free conditions. Integr Comp Biol 60.

Lam EK, Allen KN, Torres-Velarde JM, Vazquez-Medina JP. Forthcoming 2020. Primary tissue culture provides a system for functional genome-to-phenome investigations in marine mammals. Integr Comp Biol 60.

Li J, Kültz D. 2020. Proteomics of osmoregulatory responses in threespine stickleback gills. Integr Comp Biol 60 (doi: 10.1093/icb/icaa042).

Mauro AA, Ghalambor CK. 2020. Trade-offs, pleiotropy, and shared molecular pathways: how genetic and physiological integration can constrain adaptation. Integr Comp Biol 60.

Mykles DL, Burnett KG, Durica DS, Joyce BL, McCarthy FM, Schmidt CJ, Stillman JH. 2016. Resources and recommendations for using transcriptomics to address grand challenges in comparative biology. Integr Comp Biol 56:1183–91.

Schaefer RJ, Cullen J, Manfredi J, McCue ME. Forthcoming 2020. Functional contexts of adipose and gluteal muscle tissue co-expression networks in the domestic horse. Integr Comp Biol 60.

Schaefer RJ, Michno J-M, Jeffers J, Hoekenga O, Dilkes B, Baxter I, Myers CL. 2018. Integrating coexpression networks with GWAS to prioritize causal genes in maize. Plant Cell 30:2922–42.

Schwenk K, Padilla DK, Bakken GS, Full RJ. 2009. Grand challenges in organismal biology. Integr Comp Biol 49:7–14.

Westerman E, Bowman SEJ, Davidson B, Davis MC, Larson ER, Sanford C. 2020. Deploying big data to crack the genotype to phenotype code. Integr Comp Biol 60 (doi: 10.1093/icb/icaa055).