SUPPORTING RESEARCH ON GENE REGULATORY NETWORK SCIENCE

(12/15/2016)

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

All biological functions are orchestrated by cohorts of molecules that operate together within a biological system through functional interactions. Gene regulatory networks (GRNs) are responsible for the control of gene expression throughout the genome and in every biological process, and consist of interactions between regulatory molecules. Thus in order to achieve a causal explanation for biological functions such as development and physiology, it is essential to understand the control of genome activity at the GRN level.

In our view, understanding the logic of GRNs is the ultimate heuristic interpretation of biology, representing a natural extension of the last 150 years of research on how genotype leads to phenotype. Understanding the logic and function of GRNs is therefore fundamental to biology, comparable in importance to the laws of inheritance, the discovery of the double helical structure of DNA, and the central dogma of molecular biology.

At present, only few GRNs have been experimentally explored, and this field is only in its early beginnings, perhaps comparable to where the field of genetics was after the discovery of the first few genes. Based on the current state of GRN science, we do not have sufficient information to understand how GRNs control the diverse biological functions in various organisms, nor are we in the position to predict the changes in GRN function underlying human disease. It is therefore imperative to expand our current knowledge on the architecture of GRNs that control biological functions. Enhancing GRN science will not only greatly increase our mechanistic understanding of biological systems, but will also provide the foundation to address the origin of complex human diseases.

This White Paper presents a position statement providing a vision for the enhancement of GRN science, including prioritized recommendations on how this vision could be accomplished. We provide background information on the concepts and goals of GRN science, and its applications in many areas of biology. We have identified four major areas that are essential for the success of GRN science. These are 1) enhancing the experimental discovery of GRNs in a variety of biological contexts; 2) developing computational approaches for GRN modeling, 3) building a central website to facilitate sharing of various resources, and 4) strengthening of an interdisciplinary community of theoretical and experimental biologists. Our recommendations show a clear path towards the solution of genomic information systems that will transform the way we approach and comprehend biology as a whole.

2. INTRODUCTION: GOALS OF THIS GRN WHITE PAPER

This White Paper represents at least in part the discussions that were held at a GRN workshop in February 2016. The participants of this workshop [Appendix I and II] consider GRN research to be a highly valuable and exciting new field attracting scientists from various backgrounds. Their consensus is that better orchestration of efforts in this field will greatly facilitate cross-fertilization of knowledge, and strengthen GRN research and its various applications.

The aim of this White Paper is to facilitate the process of enhancing GRN science by providing information on current insights and common goals in this research field, and its application to various other branches of biology. Since GRNs constitute the genomic control systems for a broad range of biological processes, including development and physiology in animals, plants, fungi, and bacteria, this text is meant to be inclusive and accelerate GRN research in various contexts.

2.1. The Need for GRN Science

Since the discovery of the structure of DNA and the uncovering of the genetic code, there has been a push to understand phenotype as a result of genomic sequence information. We come to the realization that knowledge of how genomes encode genes, which code for RNAs and proteins is insufficient to explain complex biological phenomena. This is because genes function in networks and it is their function within these networks that ultimately generate biological outcomes. Thus, identifying the network structure, function and activity is a necessary step toward comprehending the cause of various embryonic and cellular behaviors, as well as predicting phenotypes in disease.

Many biological processes, including invariant developmental processes, are hardwired in the genome, and ultimately understanding the genomic control of biological processes will be key to understanding causality in biological systems. GRNs are networks of regulatory interactions among transcription factors and signaling molecules controlling gene expression. Both the regulatory genes and their regulatory interactions are directly encoded in the genome, and they provide one of the primary control mechanisms of biological processes - transcriptional regulation of gene expression in time and space. GRN science focuses on the mechanisms by which genomic information processing controls gene expression, thereby defining the developmental and effector state transitions in diverse biological contexts.

Understanding how the “algorithm” of the genome controls cellular phenotypes in development requires solving regulatory functions at a genome-wide scale. Ultimately, a GRN model not only includes a molecular parts list and reflects the network architecture, but also explains phenotype. A researcher seeking to understand the mechanistic basis of a particular disease should be encouraged to not only catalog the gene expression differences between diseased and healthy cells, but also understand the GRN underlying these different conditions. Thus, the major goal of GRN science is to discover causal explanations in biology, beyond the acquisition of largely descriptive big data such as cataloging gene expression and mapping mutations in the genome.

Deeper knowledge of the gene regulatory programs controlling cell fates *in vivo* will have a broad and significant impact on a number of areas in the life sciences and medicine. GRN science should therefore be considered as a unifying platform that helps bring scientists working in different disciplines, who speak different “languages”, together to communicate and to exchange and develop new ideas under one umbrella. By doing so,

we can address the "whys" and “hows” of biological processes, including cellular differentiation, and apply GRNs with the power to predict specific phenotypic outcomes.

2.2. Importance of Integrating Big Data Science, Modeling and Hypothesis-Based Research

Our overarching goal is to promote and advance GRN science in order to improve our understanding of causality in biological processes including embryonic development and in health and disease. Advances in high-throughput sequencing technologies in the post-genomic era have revolutionized the ways we acquire data in the biological sciences. Recent advances in “big data science” have produced large amounts of DNA, RNA and proteomic data, as well as accelerated technological improvements at an unforeseen rate, such that technological limitations are no longer regarded as a major limiting step in biological discovery. It has become clear however that identifying all the molecular parts encoded in the genome is only a first step towards understanding how context-specific biological processes are controlled. As we shall see below, the knowledge of how these parts are combined and function together is an equally important unsolved question.

However, there is a chasm today between the genomic research communities that use high-throughput methods and those using more traditional experimental approaches. The former often produce large amounts of putative interaction data that is difficult to experimentally verify and may contain many false positives or non-functional interactions. The latter consider small networks and pathways of a few genes with highly-curated interactions at a relatively low throughput, often not fully embedded within the larger regulatory architecture. This separation needs to be bridged by creating a more integrated community that connects these seemingly disparate strategies. It will be essential to have a platform to integrate and share the knowledge generated by diverse approaches and GRN science promises to provide such a platform. All currently available and future data should be readily accessible for GRN study.

Ultimately, no single experimental approach is sufficient to solve GRNs. Thus a combination of experimental perturbation data with large scale descriptive data, such as whole-genome ChIP-seq data and gene expression data, will be extremely valuable to identify functional regulatory interactions. Computational approaches can further enhance the experimental discovery of GRNs. Thus computational and modeling approaches allow the generation of predictions and hypotheses, to augment the efficacy of experimental work. We are well aware that the reliability of regulatory network inference depends on each approach and implementation. It is important to integrate various biological data and studies to evaluate the accuracy of the network and the connections using generally agreed upon standards. Thus, by combining distinct approaches on different scales, e.g., high-throughput genomic with detailed functional interaction studies, and modeling with experimental validation, we will gain the multi-dimensional insights required to access the architecture and function of complex GRNs.

3. THE GRN PARADIGM AND ITS GOALS

GRNs consist of regulatory interactions between sequence-specific DNA binding molecules and regulatory DNA that control the spatial and temporal expression of all genes in the genome. In addition, regulatory processes such as signaling interactions and co-factors that contribute to specific transcriptional regulation are crucial components of GRNs. Thus any process that depends on genomic information, such as animal

development, cell type differentiation, and physiological responses, and many others in eukaryotes and prokaryotes (and viruses), is controlled by GRNs.

When GRN architecture is compared among different species or different biological processes, common mechanisms can be identified that are encoded by similar network circuitries (1-3). Some regulatory circuits are used in diverse biological contexts and "wired" in such a way that they are not easily reorganized. Other types of network linkages are more flexible. Thus, comparative GRN analysis across different species will provide the mechanistic understanding of the similarities and differences between developmental strategies.

The study of GRNs aims at addressing the complex regulatory functions encoded in the genome in order to reveal the causality underlying biological processes. In the following are some of the major aims envisioned by this scientific field:

- *To access genomic information controlling biological processes:* The study of GRNs focuses on mechanisms by which genomic information processing controls gene expression, thereby determining the developmental and effector state transitions in diverse biological contexts. Because of the unidirectional information flow within genomic control systems, GRNs are distinct from other types of networks, such as protein-protein interaction networks or metabolic networks.

- *To achieve genomic circuit based explanations:* A major focus of traditional molecular biology has been the identification of individual molecular functions within a specific biological context. However in the context of gene regulation, it is evident that the same transcription factors may execute a variety of context-specific functions. The regulatory function of transcription factors is therefore not an intrinsic molecular feature, but depends on their environment: the regulatory context and the regulatory circuitry in which they operate. Regulatory circuits execute functions that in essence depend on the constellation of regulatory interactions. Solving the structure and function of genomic network circuitries is therefore key to understanding the genomic information underlying biological processes.

- *To achieve a system-level understanding of biological functions:* Biological processes are executed not by single molecules but by the complex interplay between diverse molecular functions. To access the genomic information underlying control of biological processes we need to think beyond the function of individual genes. A system-level analysis of biological processes requires identification of all molecular players involved in the process, in addition to knowledge of their physical and functional interactions. In the context of GRNs, these components are the regulatory factors, signaling molecules, and *cis*-regulatory sequences that build the core of such networks. The analysis of GRNs represents a truly system-oriented approach where biological function and phenotypes are the emergent property of functional interactions among biological components.

- *To address causality:* Most biological processes are controlled by a hierarchy of regulatory steps. To obtain a causal understanding of biological processes, it is crucial to organize insights on molecular functions in accordance to the sequence in which they occur within a process. Therefore, GRN models are neither bite-size nor are they un-digestible network diagrams with no predictive power -- Rather, they are capable of revealing the regulatory components and interactions that provide the causality for spatial and temporal gene expression, and the resulting biological function. GRN approaches aim at resolving the mechanisms for every single step connecting upstream inputs to the downstream phenotypic outcome, and therefore provide the

organizational framework necessary to obtain a causal understanding of biological processes.

- *To generate provisional and evolving models:* As more and more information accumulates on the molecular components and their functional interactions, it becomes essential to generate models. Such models may either serve to visualize the topology of molecular components and the hierarchy in which they execute their individual functions, or to replicate *in silico* the functional behavior of complex networks. There is a common misconception that a GRN model is the final product to explain a biological process and its value ends there as a map of a developmental process. But as with any model, GRN models are dynamic maps of functional regulatory interactions that change as we gain further knowledge and have predictive power to be harnessed for hypothesis generation.
- *Engineering of GRNs:* A long-term goal of regenerative medicine and synthetic biology is to design artificial network circuits with desired outcomes. A comprehensive understanding of the underlying logic of GRNs should one day enable synthetic biology approaches by which we can create or manipulate regulatory genetic circuits at will. For example, by manipulating key nodes in a GRN we could prevent or reverse a diseased cell state. A detailed understanding of the GRNs controlling cell differentiation will be needed to realize the long sought goal of changing the identity of a mature cell for regenerative medicine, such as the generation of pancreatic β-cells from other terminally differentiated cell types in the body for the treatment of diabetes.
- *To enhance hypothesis-driven science:* The organization of insights on regulatory functions into increasingly complex models of GRNs provides the foundation to generate hypotheses about unresolved biological functions. In the light of GRNs, such hypotheses are often accompanied by specific predictions on the molecular components that might be involved, or a functional outcome that needs explanation. These predictions can then be specifically tested experimentally, thus substantially lowering the number of experiments necessary to either confirm or reject the initial hypothesis.
- *To enhance integrative collaboration within the scientific community:* The analysis of complex GRNs may exceed the resources of smaller research labs. However, with a common framework as well as a common set of evidential standards, it will be possible to integrate individual research results into larger interactive models of GRNs. This will not only increase the value of every individual contribution, but will in turn provide a resource that will enrich and accelerate GRN research in general.

4. IMPACT OF GRN SCIENCE ON VARIOUS RESEARCH FIELDS

GRNs are a central feature of nearly every biological process. They allow a cell to respond to a variety of environmental signals that influence the cell's activities and thereby function within a tissue, organ and organism. GRNs are particularly useful in a number of research areas that are fundamental to life sciences and to understanding human health and disease. The following are some examples:

- *Developmental Biology:* Critical issues in developmental biology include understanding how the embryonic body plan is patterned, how cells acquire specific cell fates, and how morphogenetic movements and interactions lead to the formation of tissues and organs. Embryogenesis is very robust and highly reproducible, which implies that GRNs are involved in canalization of quantitatively variable molecular and

cellular activities to produce discrete phenomenological outcomes. Elucidating the GRNs that regulate these processes will enhance our mechanistic understanding of developmental processes, and will enable powerful comparisons across different organ systems, and across different species (4-12). These insights will also enhance our understanding of the causes of developmental defects.

- *Cell Biology:* GRNs control cell biology by determining the spatial and temporal deployment of the basic and specific cellular functions. They ultimately provide the information necessary for the differentiation of cells specialized for various functions. “Cell biology GRNs” will facilitate our understanding of ciliogenesis, autophagy, cell migration, epithelial-to-mesenchymal transitions and tubulogenesis and all other functions of cells (13-15). It is also important to link cell biological GRN processes to developmental GRNs because regulated effector genes feed on multiple cellular processes.
- *Immunology and Infectious Diseases:* Recent work has elucidated the importance of signaling and transcriptional pathways as well as epigenetic signatures that underlie the development of distinct immune cell lineage selection (16-18). GRN study will shed light on the mechanisms controlling the differentiation and activation of immune cells after sensing pathogens, as well as the mechanism underlying the aberrant response of lymphocytes to self-antigens that ultimately leads to autoimmune responses, which occurs among 3-5% of the human population. Additionally, understanding how a virus or infectious cellular agent interacts with host cells at the GRN level to control host cell behavior will be valuable for controlling infection and preventing disease.
- *Evolutionary Biology:* Comparative GRN analysis provides a most powerful way to understand how diversity in structure and function of biological systems occurred during evolution as a result of genomic sequence change, and offers an experimental approach for engineering evolutionary change. Study of GRNs in evolution have proven useful in explaining phenotypic changes, including the origins of morphogenetic novelties, variation/noise, and developmental systems drift, in terms of network rewiring and co-option of genes or modules in evolution (19-23).
- *Genome Biology:* GRN research provides a way to assess the function of individual regulatory elements in the larger context of a biological process, which is one of the goals of the field of functional genomics (24-27). GRN science will reveal how the CRM structures in the genome are organized and maintained, or have diverged across different species, thereby shedding light onto the evolution of the regulatory genome during animal evolution (28).
- *Systems and Synthetic Biology:* GRN analysis at a systems level also offers promising approaches for engineered outcomes by designing new GRNs and cellular entities. These are informed by inputs from synthetic biology, theoretical systems analysis, and comparative analyses of developmental networks. These approaches are expected to provide an unparalleled opportunity to develop gene circuits designed to perform specific functions (29, 30) and to test developmental, evolutionary and cell biological hypotheses in a uniquely controlled paradigm.
- *Plant Biology:* Plants are easily affected by the environment due to their sessile mode of life, and have a series of mechanisms for responding to environmental changes and insults, leading to adaptation of their gene expression states. GRN study will uncover not only how plants control development, but also how plants respond to various stresses, including abiotic stress such as drought, excessive watering, and extreme

temperatures. Understanding the GRN involved in stress response pathways will have significant impact to secure food supply.

- *Applications to Human Health and Disease:* The experimental solution of GRNs underlying development and cell type specification will be invaluable for predicting the influence of genetic modifications in birth defects, gametogenesis/fertilization, and pediatric disease, as well as adult onset diseases arising from tissue-level defects in cellular differentiation such as skin diseases and cancers. In recent years, mutation of CRMs that interact with transcription factors has emerged as a pervasive cause of human disease (31). GRN analysis will therefore be necessary to identify the causes of birth defects and other diseases including those with multi-genic origins such as cancer (32). By providing a system-level understanding of the underlying biology, GRN science will therefore help find and develop new therapies. GRN science will also be essential in developing strategies to reprogram cells for regenerative therapeutic purposes. Thus GRN science will generate more informed hypotheses about regulatory DNA associated with disease, and accelerate the use of induced pluripotent cells (iPSCs) and other cell types for analyzing and repairing human developmental processes and human diseases.

5. WHAT IS NEEDED FOR GRN SCIENCE

Due to the inherent complexity of GRNs, this emerging research field requires support mechanisms that might differ from those required for traditional biology approaches as well as recent big data science projects. Knowledge of the particular biological system, and the experimental approaches applicable to it, are of crucial importance to GRN science, building directly upon more traditional single lab efforts. On the other hand, individual labs may not have the manpower to address large scale GRNs experimentally. Thus the advancement of GRN science will require leveraging both single lab and big data efforts and in addition foster novel schemes of collaborative interactions.

The experimental solution of GRNs is a high effort, low throughput, and high reward research endeavor. This is why so few GRNs have been experimentally solved at a relatively deep level. However, where information on GRN structure and function has become available (4), it has not only illuminated the genomic control of a given process, but also served as rich foundation for other research projects with diverse applications. GRN science also employs other approaches, such as the visualization and dynamic modeling of GRNs, since any circuit beyond a few nodes is no longer intuitively accessible. The following is a list of areas crucial for GRN science. Enhancement of these areas will directly contribute to advancing GRN research.

- *Experimental identification of GRNs:* The goal is to obtain experimental evidence for GRNs in different biological contexts. While available experimental approaches are not applicable in every model system, it is nevertheless possible to enunciate a unifying standard of evidence for solving GRNs. Ultimately, a high quality analysis of GRNs includes evidence from each of the following categories of data, which all contribute in different ways to the understanding of GRN structure and function:
    - *Experimental identification of molecular players:* This includes experimental analysis of expressed regulatory genes in different biological contexts. Technologies currently applied include qPCR, RNA-seq, and Nanostring, for quantitative expression data, and *in situ* hybridization approaches for spatial gene expression data. New technologies such as single cell transcriptomics and

quantitative *in situ* hybridization technologies will allow identification of regulatory states at even higher resolution.

- *Experimental identification of physical sequence-specific interactions between regulatory factors and DNA:* this includes experimental approaches identifying the direct interaction between transcription factors and regulatory DNA sequences, such as chromatin immunoprecipitation (ChIP), and to some extent data obtained from ATAC-seq (Assay for Transposase-Accessible Chromatin sequencing) and Hi-C (a high-throughput sequencing chromatin conformation capture assay) data.
- *Experimental identification of functional interactions between regulatory factors and DNA:* a combination of *cis*- and *trans*-perturbation approaches is usually required for unambiguously identifying the function of a particular transcription factor in the control of a given *cis*-regulatory sequence and its associated gene. This information is provided by experimental data analyzing the consequence of transcription factor perturbation or the perturbation of signaling pathways. Evidence for functional interactions can also be generated by *cis*-regulatory experiments, by mutation of predicted binding sites for given transcription factors.
- *Experimental analysis of cis-regulatory function*: regulatory DNA sequence is a key component of GRNs, and any information on the functional and structural organization of cis-regulatory modules directly illuminates a piece of a GRN.

- *GRN modeling*: Predicting the behavior of network circuits consisting of anything but a very small number of genes becomes impossible without the help of computational approaches. The goal is to produce computational GRN models of development and other biological processes that can predict the phenotypic outcome of functional GRN interactions under various conditions. Thus advanced modeling approaches are required to reveal the dynamic behavior and system-level output of complex GRNs, allowing the *in silico* interrogation of how biological systems will change in response to genetic and/or environmental changes. While a complete GRN model is an ultimate goal, it is one that cannot be achieved in one step, but instead gradually with the acquisition of new knowledge. Thus it is useful to construct a GRN model even for incomplete GRN data, as a means to identify the gaps and predict what is missing. Such models are useful to generate hypotheses to direct and accelerate future experiments designed to identify missing components of the GRN. GRN models should thus not necessarily be regarded as ultimate answers but rather as models that facilitate the gradual approach to a complete picture by providing hypotheses to test, validate, and improve our knowledge of the biological system.
- *Databases and Resources:* GRN science is an emerging area of research with a distinct focus on multi-disciplinary approaches using a combination of genomic, molecular and bioinformatics tools. Thus it is essential to develop a platform that many users coming from different backgrounds can use to share, integrate, and analyze different types of data sets. Thus, progress in GRN science would be greatly facilitated by building a central website providing access to essential GRN resources available. This way, knowledge of the particular biological system and the experimental approaches applicable to it, can be shared among all laboratories. Examples of needed resources are as follows.
  - *GRN visualization tools:* The graphical representation of GRN architecture is necessary for visualization and communication of GRN information. Currently there are only few software programs that are used to represent GRNs, such as BioTapestry, and additional computational tools will hopefully be developed in

future. For GRN research to be broadly accessible, platforms need to be easy to use by researchers constructing a network model and by researchers wanting to use network models to predict biological outcomes. Ideally, we want to visualize a GRN at multiple levels – single nodes to subcircuits to an entire biological process – so that we can understand the flow of information through the network as genes are activated or repressed. From this graphic representation the investigator can gain a sense of how the network functions and identify areas where information is missing.

- *Literature and evidence based database:* Thousands of published gene perturbation experiments are already available in the scientific literature. Improved data mining approaches to identify and incorporate published experimental results would be immensely valuable and effective for maximizing the impact of previous NIH-funded research. One approach to achieve this would be to leverage the highly annotated large scale datasets housed in model organism databases including MGI, Xenbase, ZFIN, Flybase, Wormbase and Echinobase.
- *Communication and standardized language*: Given the diverse backgrounds of scientists in this field, a standardized language will be important to facilitate interactions. For example, even the term "Gene Regulatory Network" is commonly used but its meaning varies in different research communities. Formalizing term usage is beneficial in all fields of inquiry to clarify meanings and facilitate communication, and should be adopted in the GRN field. A common framework as well as common evidential standards will promote the integration of individual research results into larger models of GRNs. This will not only increase the value of individual contributions, but will in turn provide a resource that will enrich and accelerate research in all laboratories.
- *Training:* GRN studies require a unique interdisciplinary mix of theoretical and experimental biologists with training in computer science, mathematics, genomics, molecular biology and developmental biology. Cross-disciplinary training of GRN scientists is essential and urgently needed. Training courses should be developed and supported for all levels of scientists, providing additional training for current researchers and cross-disciplinary training for new students

6. MECHANISMS TO ACHIEVE GOALS

The following resources are essential for the success of GRN science.

- *Funding mechanisms*: It is critical to advocate for grant funding in emerging areas of GRN research and training activities involving graduate and postdoctoral fellows, especially for those who wish to be cross-trained in interdisciplinary fields. The funding structures may be multi PI or multi-disciplinary NIH and NSF partnerships, as GRN science requires a highly integrative approach. Because the field of GRN science is very diverse, broad and integrative, grant reviewers should be recruited from a spectrum of scientific disciplines, and the review process should be consolidated under a special emphasis panel.
- *Web portal:* Data sharing is an integral and essential part of GRN science and being able to readily access different data sets and various analyses with common scientific standards is required. Highly interactive online GRN models, generated using BioTapestry or other formats, will allow researchers to visualize complex networks, and to share access to information on various small or large network models. A web

portal should be developed that includes a GRN blog, organized discussion group, course postings, literature, meeting announcements, contacts, and educational components.

- *Training initiatives*: It is important to train advanced graduate students, postdoctoral scholars, and professional scientists in a comprehensive theory of GRN structure, the use of BioTapestry and other formats, and the various computational platforms for representation of GRNs and models. A course on GRN science is currently offered at the Marine Biology Laboratory (Woods Hole, MA). This type of course should continue to be offered and perhaps expanded in future.
- *Conferences and Symposia*: National and international GRN meetings should be held in alternate years to foster communication and collaborations. We encourage promotion of a dedicated scientific section at various national and international societies to demonstrate the importance of GRN science and publication of special journal issues devoted to GRN research.

7. CONCLUDING STATEMENT

We stand in an unprecedented time in biomedical research at which rapidly advancing technologies are resulting in an avalanche of big data on the nature, structure and regulation of genome processes. At the same time decades of R01-funded research has provided a deep understanding of how some individual genes and pathways function to control biological processes. We believe that major advances await by integrating big data with detailed molecular mechanisms. GRN research provides a unified conceptual framework for such data integration. Elucidating the structure and functional logic of GRNs will provide a mechanistic system-level understanding of how information encoded in the genome is executed to control cell fate and cell function in development, stem cells, homeostasis and disease.

GRN science provides a unifying framework for broad areas in biology that can be applied to obtain causal understanding, thus opening unique possibilities for an integrative approach to modern biological sciences. Since GRN research lies at the intersection of such traditionally separate research fields as developmental biology, molecular biology, systems biology, biological engineering, modeling, evolution, and functional genomics, application of an integrative framework will accelerate advancements in all of these research areas by facilitating the transfer of insights and by cross-fertilization.

8. APPENDICES

I. Authors of the GRN white paper (Steering Committee)

Co-Chairs:

Ken Cho (UC Irvine, CA)

Isabelle Peter (Caltech, CA)

Members:

Bill Longabaugh (Inst. Systems Biol, WA)

Dave McClay (Duke U, NC)

Sally Moody (GWU, D.C.)

Harinder Singh (CCHMC, OH)

Aaron Zorn (CCHMC, OH)

II. GRN Conferences

2016 Organizational GRN workshop, February 2-3, Caltech, Pasadena

In order to facilitate the progress in GRN research, 28 leading scientists in the field, with backgrounds in developmental biology, genetics, mathematics and computational biology gathered on February 2-3, 2016 at Caltech, Pasadena, to discuss major opportunities and challenges facing GRN research.

Attendees of the conferences who endorsed the GRN white paper are:

Scott Barolo, Associate Professor, Cell & Developmental Biology, University of Michigan

Marianne Bronner, Professor, Division of Biology and Biological Engineering, California Institute of Technology

Benoit Bruneau, Professor, Department of Pediatrics, UC San Francisco

Robert A. Cameron, Senior Research Associate, Division of Biology and Biological Engineering, California Institute of Technology

Aravinda Chakravarti, Director, Center for Complex Disease Genomics & Professor, Medicine Center for Complex Disease Genomics, Johns Hopkins

Ken W.Y. Cho, Professor, Developmental and Cell Biology, UC Irvine.

Lionel Christiaen, Associate Professor, Center for Developmental Genetics & Department of Biology, New York University

James N. Coulombe, Program Director, Developmental Genetics and Systems Developmental Biology, Developmental Biology & Structural Variation Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development

Douglas H. Erwin, Curator, Paleozoic Invertebrates, Smithsonian Institution

Roberto Feuda, Division of Biology and Biological Engineering, California Institute of Technology

Veronica Hinman, Associate Professor, Department of Biological Sciences Carnegie Mellon University

Bill Longabaugh, Senior Software Engineer, Institute of Systems Biology

Morris Maduro, Professor, Department of Biology, UC Riverside

Richard Mann, Professor, Department of Biochemistry & Molecular Biophysics, Columbia University

Dave McClay, Professor of Biology, Duke University

Anne-Helene Monsoro-Burg, Professor, Paris Sud University

Sally A. Moody, Professor and Chair, Anatomy & Regenerative Biology, George Washington University

Ali Mortazavi, Assistant Professor, Developmental and Cell Biology, UC Irvine

Leon Peshkin, Systems Biology, Harvard Medical School

Isabelle Peter, Research Professor, Division of Biology and Biological Engineering, California Institute of Technology

John Reinitz, Professor, Departments of Statistics; Ecology and Evolution; Molecular Genetics & Cell Biology, University of Chicago

Ellen Rothenberg, Professor, Division of Biology and Biological Engineering, California Institute of Technology

Maike Sander, Professor, Pediatrics & Cellular & Molecular Medicine, UC San Diego

Tatjana Sauka-Spengler, Associate Professor, Department of Medicine, University of Oxford

Harinder Singh, Director, Division of Immunobiology & Professor, Pediatrics, Cincinnati Children's Hospital.

Saurabh Sinha, Professor, Dept. of Computer Science, University of Illinois

Angela Stathopoulos, Professor, Division of Biology and Biological Engineering, California Institute of Technology

Andrea Streit, Professor, Developmental Neurobiology, King's College, London

Aaron M. Zorn, Associate Director, Digestive Health Center, Department of Pediatrics, Cincinnati Children's Hospital.

III. GRN Courses

Annual Course on Gene Regulatory Networks for Development, Marine Biological Laboratory, Woods Hole

IV. References

1. Davidson EH & Levine MS (2008) Properties of developmental gene regulatory networks. *Proc. Natl Acad. Sc.i USA* 105(51):20063-20066.
2. Briscoe J & Small S (2015) Morphogen rules: design principles of gradient-mediated embryo patterning. *Development (Cambridge, England)* 142(23):3996-4009.
3. Peter IS & Davidson EH (2015) *Genomic Control Process, Development and Evolution* (Academic Press/Elsevier).
4. Peter IS & Davidson EH (2011) A gene regulatory network controlling the embryonic specification of endoderm. *Nature* 474(7353):635-639.
5. Oliveri P, Tu Q, & Davidson EH (2008) Global regulatory logic for specification of an embryonic cell lineage. *Proc. Natl. Acad. Sci. USA* 105(16):5955-5962.
6. Aguilar-Hidalgo D*, et al.* (2013) A Hh-driven gene network controls specification, pattern and size of the Drosophila simple eyes. *Development (Cambridge, England)* 140(1):82-92.
7. Swiers G, Chen YH, Johnson AD, & Loose M (2010) A conserved mechanism for vertebrate mesoderm specification in urodele amphibians and mammals. *Developmental biology* 343(1-2):138-152.
8. Simoes-Costa M & Bronner ME (2015) Establishing neural crest identity: a gene regulatory recipe. *Development (Cambridge, England)* 142(2):242-257.
9. Maduro MF*, et al.* (2015) MED GATA factors promote robust development of the C. elegans endoderm. *Dev Biol* 404(1):66-79.
10. Tumpel S, Wiedemann LM, & Krumlauf R (2009) Hox genes and segmentation of the vertebrate hindbrain. *Curr. Top. Dev. Biol.* 88:103-137.
11. Chan TM*, et al.* (2009) Developmental gene regulatory networks in the zebrafish embryo. *Biochim Biophys Acta* 1789(4):279-298.
12. Herrmann F, Groß A, Zhou D, Kestler HA, & Kühl M (2012) A boolean model of the cardiac gene regulatory network determining first and second heart field identity. *PLoS One* 7(10):e46798.
13. Christiaen L*, et al.* (2008) The transcription/migration interface in heart precursors of Ciona intestinalis. *Science (New York, N.Y* 320(5881):1349-1352.
14. Rafiq K, Shashikant T, McManus CJ, & Ettensohn CA (2014) Genome-wide analysis of the skeletogenic gene regulatory network of sea urchins. *Development (Cambridge, England)* 141(12):2542.
15. Saunders LR & McClay DR (2014) Sub-circuits of a gene regulatory network control a developmental epithelial-mesenchymal transition. *Development (Cambridge, England)* 141(7):1503-1513.
16. Amit I*, et al.* (2009) Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science (New York, N.Y* 326(5950):257-263.
17. Laslo P*, et al.* (2006) Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell* 126(4):755-766.

18. Yui MA & Rothenberg EV (2014) Developmental gene networks: a triathlon on the course to T cell identity. *Nature reviews. Immunology* 14(8):529-545.
19. Erwin DH & Davidson EH (2009) The evolution of hierarchical gene regulatory networks. *Nat Rev Genet* 10(2):141-148.
20. Peter IS & Davidson EH (2011) Evolution of gene regulatory networks controlling body plan development. *Cell* 144(6):970-985.
21. Frankel N*, et al.* (2011) Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* 474(7353):598-603.
22. Prud'homme B*, et al.* (2006) Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* 440(7087):1050-1053.
23. Schneider I*, et al.* (2011) Appendage expression driven by the Hoxd Global Control Region is an ancient gnathostome feature. *Proc Natl Acad Sci U S A* 108(31):12782-12786.
24. Dickel DE*, et al.* (2016) Genome-wide compendium and functional assessment of in vivo heart enhancers. *Nat Commun* 7:12923.
25. Betancur P, Bronner-Fraser M, & Sauka-Spengler T (2010) Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest. *Proc Natl Acad Sci U S A* 107(8):3570-3575.
26. Swanson CI, Evans NC, & Barolo S (2010) Structural rules and complex regulatory circuitry constrain expression of a Notch- and EGFR-regulated eye enhancer. *Dev. Cell.* 18(3):359-370.
27. Ben-Tabou de-Leon SB & Davidson EH (2010) Information processing at the foxa node of the sea urchin endomesoderm specification network. *Proc. Natl Acad. Sci. USA* 107(22):10103-10108.
28. Chan YF*, et al.* (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science (New York, N.Y* 327(5963):302-305.
29. Elowitz MB & Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403(6767):335-338.
30. Stricker J*, et al.* (2008) A fast, robust and tunable synthetic gene oscillator. *Nature* 456(7221):516-519.
31. Maurano MT*, et al.* (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science (New York, N.Y* 337(6099):1190-1195.
32. Chatterjee S*, et al.* (2016) Enhancer Variants Synergistically Drive Dysfunction of a Gene Regulatory Network In Hirschsprung Disease. *Cell* 167(2):355-368.e310.