WHITE PAPER IN SUPPORT OF RESEARCH ON GENE REGULATORY NETWORKS

(draft date: 9/18/2016)

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

Gene regulatory networks (GRNs) constitute the genomic control systems for a broad range of biological processes, including the definition of spatial organization during development, morphogenesis, the differentiation of cell types, and the physiological response to internal and external changes in animals, plants, and prokaryotes. In recent years, the successful solution of several GRNs demonstrated the feasibility of addressing experimentally and computationally the causal control system of a variety of biological processes, opening the gate to a new area of biological sciences. However, despite these important advancements, only a few biological processes have so far been addressed at the GRN level. With this White Paper, we aim to bridge this current gap by illuminating the exciting possibilities of this novel research field, and by providing information and guidelines to facilitate the experimental and computational application of the GRN framework to a variety of biological contexts.

The focus of GRN studies is to understand biological function as a consequence of genomic programs. GRN science addresses the mechanism and causality of development and other biological processes in terms of genomic information processing. Even in the simplest of animals, a layer of intricate regulatory commands is needed for development of the body plan and to control spatial and temporal gene expression. This occurs via a complex set of regulatory components governed by a large set of interaction mechanisms. These mechanisms include, combinatorial processing of regulatory inputs, mediated by transcription factors, signal-sensing mechanisms, as well as modular and evolvable network circuit architectures. GRNs provide a unifying framework for the control principles that underlie seemingly disparate processes in biology. Understanding these processes at the GRN level will allow a predictive understanding of development and other biological processes. Therefore, GRN science is a natural extension of the last 150 years of research on how genotype leads to phenotype. In our view, understanding the logic of GRNs is as fundamental to understanding the biological sciences as are the laws of inheritance, the double helix structure of DNA, and the central dogma of biology.

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 challenges facing GRN research. The participants of this workshop consider GRN research to be a highly valuable and exciting new field attracting scientists from various backgrounds. Their view is that better orchestration of their efforts will greatly facilitate cross-fertilization of knowledge and strengthen GRN research and its various applications. This White Paper presents a position statement emerging from this workshop, illustrating a vision for the future of GRN science, including specific recommendations with explicit prioritizations on how this vision could be accomplished. To facilitate access to GRN science, we also provide a general introduction, discuss the impact of GRN research on different research areas in biology, and provide clarification of some common misconceptions. Our overarching goal is to promote and advance GRN science in order to rapidly improve our understanding of animal development, and evolution, as well as other genomic processes, in health and disease.

2. INTRODUCTION: GOALS OF THIS GRN WHITE PAPER

2.1. The Need for GRN Science

For several years now, biological sciences have driven the acquisition of large data sets, on genomic sequences, RNA and protein expression, and functional genomics. Current advancements in “big data science” have not only produced large amounts of data, but they have also accelerated technological improvements at an unforeseen rate, such that technological limitations are no longer regarded as a major limiting step in biological discovery. Despite the importance of these catalogues it has become clear that identifying all the molecular parts encoded in the genome is only a first step towards understanding how context-specific biological processes are controlled.

The fundamental principle of GRNs stems from the observation that many biological processes, including invariant developmental processes, are hardwired in the genome, and that 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 the regulatory interactions are directly encoded in the genome, and they operate 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 outputs, thereby defining the developmental and effector state transitions in diverse biological contexts.

Transcription factors play a central role in regulating gene expression. Importantly, transcription factors do not function in isolation or in linear pathways, but rather they control gene expression through networks of regulatory interactions (Davidson et al., 2002). Thus, each developmental decision involves the orchestrated transcriptional expression of cell-type-specific gene sets, in response to specific combinations of transcription factors and multiple intercellular signaling pathways, all within the spatial complexities of morphogenesis.

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 will be essential in developing strategies to reprogram cells for regenerative therapeutic purposes. In recent years, mutation of *cis*-regulatory modules (CRMs) that interact with transcription factors has emerged as a pervasive cause of human disease (Maurano et al., 2012). GRN science will provide answers to the genetic origins of developmental diseases. A major outcome of human genetic abnormalities leads to the dysregulation of gene expression, often affecting multiple genes at once, especially in cancer. GRN science will achieve the needed system-level understanding of the underlying biology to find and develop new therapies. Embryogenesis is very robust and highly reproducible, which implies that GRNs provide the mechanisms to canalize quantitatively variable molecular and cellular activities towards discrete phenomenological outcomes. GRN science will explain how these essential properties of developing systems have adapted and evolved.

Understanding how the “algorithm” of the genome controls cellular phenotypes in development requires the integration of transcription factor-CRM interaction data, often at a genome-wide scale, together with molecular details on regulatory function, often at the single gene level, for every regulatory gene in a living embryo. Ultimately, a GRN model not only includes a molecular parts list and reflects the network architecture, but also

provides phenotypic insights. Thus, the major goal of GRN science is to arrive at causal explanations in biology, beyond the acquisition of largely descriptive big data such as cataloging gene expression and mapping mutations in the genome. GRN science should therefore be considered as a unifying platform that helps bringing 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 developmental processes and apply GRNs with the power to predict specific phenotypic outcomes.

2.2 State of the Field

Since the discovery of the structure of DNA and the uncovering of the genetic code, there has been a push to understand how the next level of the “genetic code” leads to phenotype. Genes make RNAs that mostly make proteins, but how do phenotypes arise from such linear regulatory events? We now understand that genes function in networks and it is their behavior 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 and predicting phenotypes when the system goes awry.

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

There is a chasm today between the genomic research communities that use high-throughput methods and traditional modelers who only consider small networks containing a few genes. The former often produce large amounts of putative interaction data that is difficult to experimentally verify and may contain many false positives. The latter consider small networks and pathways of a few genes with highly-curated interactions but these aren’t fully embedding 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 be such a platform. All currently available and future data should be readily accessible for GRN study.

Computational and modeling approaches will generate hypotheses and augment the limitation of experimental approaches. GRN science will provide insight into the overall structure of transcription factor-mediated gene regulation, can predict transcription factors that play important regulatory roles and elucidate how they work together with other transcription factors. This type of integrative approach will provide many leads for follow-up and experimental validation studies. We are well aware that the reliability of regulatory network inference depends on each approach and implementation. Thus, it is important to integrate various biological data and studies to examine the accuracy of the network and the connections using generally agreed upon standards. Thus, by combing distinct approaches on different scales, e.g., high-throughput genomic and detailed interaction studies, we get the best of both approaches.

2.4. Prioritization of Needed Resources

GRN approaches aim to resolve the mechanisms for every single step connecting upstream inputs to the downstream phenotypic outcomes, and therefore provide the

organizational framework necessary to understand the causality underlying biological processes. In recent years, the technology of DNA sequencing has improved, generating more data at lower expense, and providing increased sensitivity capable of obtaining data from single cells. With such changes, we envision that the progress of GRN science will be greatly influenced by the availability of research funding and resources in the following areas.

- ***Experimentation:*** Functional interactions between transcription factors and regulatory DNA are the building blocks of GRNs, and a large number of high-confidence data sets are needed. Since experimental approaches are essential in GRN studies, it will be important to develop medium to high-throughput assays to enhance the discovery of network structures, and to support the experimental solution of biological processes at the GRN level. It will also be invaluable to establish data quality assessment and standardization across the GRN field.
- ***Visualization:*** While GRN science will require integration of sophisticated approaches, the GRN network structures and outputs need to be presented to researchers via broadly accessible platforms.
- ***Modeling:*** As more information accumulates on the molecular components and their functional interactions, it becomes essential to generate models. Modeling approaches also help identify gaps in current understanding of GRNs, and stimulate further experimental exploration. Models should therefore not be regarded as the endpoint of molecular analyses, but as an important tool in the stepwise discovery of GRNs and hypothesis building.
- ***Training:*** GRN studies require a unique interdisciplinary mix of theoretical and experimental biologists working in different disciplines, including 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.
- ***Databases/Resources:*** It will be indispensable to have a central one-stop information center where essential GRN resources are available. Examples of needed resources include visualization tools, literature and evidence based databases, comparison tools to identify GRN structures, sophisticating modeling tools, and transcription factor binding databases that are tailored for different model organisms.

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 that change the activity of transcription factors are crucial components of GRNs. What distinguishes GRNs from other biological networks is that both the regulatory components as well as the physical and functional regulatory interactions are directly encoded in the genome. Thus any process that depends on genomic information, such as animal development, cell type differentiation, and physiological response in eukaryotes and prokaryotes, and many others, is controlled by GRNs.

For example, GRNs encoded by animal genomes control the activation of critical genes to regulate their amplitude, timing and location of gene expression during development. Recent conceptual and technological advances demonstrate that development is controlled by GRNs, and that GRNs explain the causality in development by showing how the regulation of gene expression is progressively generated by dynamic gene interactions. Thus, the regulatory logic controlling an organism's development is hardwired in the genome (Peter and Davidson, 2015).

When GRN architecture is compared among different species or different biological processes, common mechanisms can be identified that are encoded by similar network circuitry (Briscoe and Small, 2015; Davidson and Levine, 2008). Some regulatory circuits are used in many 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 discovery that particular developmental functions even in very different organisms and biological contexts are encoded by similar network architectures illuminates the importance of accessing the GRN code.

The study of GRNs aims at making accessible 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:*** Many biological processes are directly determined by genomic information, including cell fate specification, response to intercellular signaling interactions, development of multicellular organisms, animal physiology and many more. 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. Thus GRNs are distinctly and uniquely different from other types of networks, such as protein-protein interaction networks or metabolic networks, because of the unidirectional information flow within genomic control systems.
- ***To achieve genomic circuit based explanation:*** 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 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 processes 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 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 the biological processes and its value ends there as a map of a developmental process. On the contrary, GRN models should be viewed as dynamic maps of functional regulatory interactions that change as we gain further knowledge and have predictive power to be harnessed for hypothesis generation.
- ***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 directly addressed in specifically designed experiments, 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 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. 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. These GRNs will be invaluable for predicting the influence of the primary genes and their modifiers that are mutated in birth defects, gametogenesis/fertilization, and pediatric disease, as well as adult onset diseases arising from defects in tissue-level defects in cellular differentiation such as skin diseases and cancers.
- ***Cell Biology:*** GRNs control cell biology by determining the spatial and temporal deployment of the basic activities of cellular function. They ultimately provide the information necessary for the differentiation of cells specialized for various functions. "Cell biology GRNs" will facilitate our understanding of innumerable processes that affect disease states, such as ciliopathies and aberrant autophagies, and abnormalities in processes required for normal organ formation, such as cell migration, epithelial-to-mesenchymal transitions, signaling pathways, and tubulogenesis. It is important to link cell biological GRN processes to developmental GRNs because regulated effector genes feed on multiple cellular processes, and modulate the GRN-cell biology interface.
- ***Evolutionary Biology:*** 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 approach for engineering evolutionary change using an experimental basis. GRNs will be useful in explaining phenotypic changes, variation/noise, and developmental systems drift, in terms of network rewiring and co-option of genes or modules in evolution.
- ***Genomics:*** It is only in recent years that the functions of regulatory regions in the genome are being revealed. 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 in the field of functional genomics. GRN science will also reveal how the CRM structures in the genome are maintained or have diverged among different species, thereby shedding light onto the evolution of GRN structures during animal development.
- ***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 test developmental, evolutionary and cell biological hypotheses in a uniquely controlled paradigm.
- ***Application for Human Health and Disease****:* Advances in elucidating and validating the GRNs in all of the processes mentioned above will have significant impact on our understanding of adult homeostasis and disease states. They will predict underlying causes of birth defects and disease, and therefore be of great value in screening and in designing therapeutics. In addition, they will generate more informed hypotheses about regulatory DNA associated with disease, and accelerate the use of induced pluripotent cells (iPSCs) for analyzing human developmental processes and human disease. Cancer formation can be considered as reactivation of early developmental programs when the cells go awry. By understanding the GRN structures required to

maintain the cancerous cellular state, GRN science will shed light into the causality of cancer and lead to better strategies for treatment at the systems level.

5. WHAT IS NEEDED FOR GRN SCIENCE

Due to the inherent complexity of GRNs, this emerging research field requires support mechanisms that might be different from both 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, high reward research endeavor. This is why so few GRNs have been experimentally solved at a fairly complete level. However, where information on GRN structure and function has become available, 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 given 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 providing information on expressed regulatory genes, such as qPCR, RNA-seq, and Nanostring, for quantitative expression data, and *in situ* hybridization approaches for spatial expression data. Novel technologies are currently being developed with the potential to combine both spatial and quantitative gene expression data, such as single cell transcriptomics and quantitative in situ hybridization technologies.
    - *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) data.
    - *Experimental identification of functional interactions between regulatory factors and DNA:* 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. 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.

- *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 visualization:*** The graphical representation of GRN architecture is necessary for visualization and communication of GRNs. Currently there are only few software programs that are used to represent GRNs, such as BioTapestry. 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 wants to visualize a GRN at multiple levels – single nodes to subcircuits to 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.

- ***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. It is important to realize that even incomplete GRNs (rudimentary topological models) are still extremely useful and contain information to predict biological outcomes, discover new network structures, address evolutionary questions, and find the reason behind various disease states.

- ***Continuously improving and updating GRN models:*** While a complete GRN model is an ultimate goal, it is one that cannot be achieved in one strike. Thus it is useful to obtain a model for even largely incomplete GRN data, as a means to identify the gaps and predict what is missing. Such models direct and accelerate future experiments designed to identify missing components of the GRN. Models should thus not necessarily be regarded as ultimate answers but rather as approaches to test, validate, and improve our knowledge of GRNs.

- ***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 where 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 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.

- ***Literature searches:*** Tens of thousands of published gene perturbation experiments are already available in the scientific literature, and 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 every individual contribution, but will in turn provide a resource that will enrich and accelerate research in all laboratories.

6. MECHANISMS TO ACHIEVE GOAL

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. 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 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 would be important to train advanced graduate students, postdoctoral scholars, and professional scientists in a comprehensive theory of developmental GRN structure, the use of BioTapestry, 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***: Both national and international GRN meetings should be held in alternate years to foster communications and collaborations among GRN scientists. We encourage promotion of a dedicated scientific section at various national and international societies to demonstrate the importance of GRN science and to publish special issues on GRN research in journals.

7. CONCLUSION STATEMENT

We stand in an unprecedented time in biomedical research where technological advances 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. What is needed now is the integration of this big data with detailed molecular mechanisms. GRN research provides a means to bridge and integrate this wealth of data into a unified conceptual framework. 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. In particular it will be critical to determine the regulatory logic and GRN circuits operating a variety of biological processes, and to elucidate the molecular basis underlying given regulatory functions such as AND, OR, and NOT logic gates within regulatory networks.

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 diverse traditionally separate research fields, including developmental biology, molecular biology, systems biology, biological engineering, modeling, evolution, and functional genomics, application of a unifying principle and language will accelerate advancements in all of these research areas by facilitating the transfer of insights and the cross-fertilization among these different biological fields.

8. APPENDICES

Authors of the GRN white paper

Ken Cho
Dave McClay
Sally Moody
Bill Longabaugh
Isabelle Peter
Harinder Singh
Aaron Zorn

GRN Conferences

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

GRN Courses

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

References

Briscoe, J., Small, S., 2015. Morphogen rules: design principles of gradient-mediated embryo patterning. Development (Cambridge, England) 142, 3996-4009.

Davidson, E.H., Levine, M.S., 2008. Properties of developmental gene regulatory networks. Proc. Natl Acad. Sc.i USA 105, 20063-20066.

Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim, O., Brown, C.T., Livi, C.B., Lee, P.Y., Revilla, R., Rust, A.G., Pan, Z., Schilstra, M.J., Clarke, P.J., Arnone, M.I., Rowen, L., Cameron, R.A., McClay, D.R., Hood, L., Bolouri, H., 2002. A genomic regulatory network for development. Science (New York, N.Y 295, 1669-1678.

Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., Shafer, A., Neri, F., Lee, K., Kutyavin, T., Stehling-Sun, S., Johnson, A.K., Canfield, T.K., Giste, E., Diegel, M., Bates, D., Hansen, R.S., Neph, S., Sabo, P.J., Heimfeld, S., Raubitschek, A., Ziegler, S., Cotsapas, C., Sotoodehnia, N., Glass, I., Sunyaev, S.R., Kaul, R., Stamatoyannopoulos, J.A., 2012. Systematic localization of common disease-associated variation in regulatory DNA. Science (New York, N.Y 337, 1190-1195.

Peter, I.S., Davidson, E.H., 2015. Genomic Control Process, Development and Evolution. Academic Press/Elsevier.