Chapter 1
Systems Hematology: An Introduction

Seth Joel Corey, Marek Kimmel and Joshua N. Leonard

Abstract Hematologists have traditionally studied blood and its components by simplifying it into its components and functions. A variety of new techniques have generated large and complex datasets. Coupled to an appreciation of blood as a dynamic system, a new approach in systems hematology is needed. Systems hematology embraces the multi-scale complexity with a combination of mathematical, engineering, and computational tools for constructing and validating models of biological phenomena. The validity of mathematical modeling in hematopoiesis was established early by the pioneering work of Till and McCulloch. This volume seeks to introduce to the various scientists and physicians to the multi-faceted field of hematology by highlighting recent works in systems biology. Deterministic, stochastic, statistical, and network-based models have been used to better understand a range of topics in hematopoiesis, including blood cell production, the periodicity of cyclical neutropenia, stem cell production in response to cytokine administration, and the emergence of drug resistance. Future advances require technological improvements in computing power, imaging, and proteomics as well as greater collaboration between experimentalists and modelers. Altogether, systems hematology will improve our understanding of normal and abnormal hematopoiesis, better define stem cells and their daughter cells, and potentially lead to more effective therapies.

Keywords Hematology · Models · Reductionist · Systems biology

S. J. Corey (✉)
Department of Pediatrics and Cell & Molecular Biology,
Northwestern University Feinberg School of Medicine, Lurie 5-107,
303 E. Superior St., Chicago, IL 60611, USA
Tel.: 312-503-6694
e-mail: coreylab@yahoo.com

M. Kimmel
Department of Statistics and Bioengineering, Rice University, 2102 Duncan Hall,
6100 Main St., Houston, TX 77005, USA
Tel.: (713) 348-5255
e-mail: kimmel@rice.edu

J. N. Leonard
Technological Institute, Rm E136, 2145 Sheridan Rd., Evanston, IL 60208-3120, USA
Tel.: 847-491-7455
e-mail: j-leonard@northwestern.edu

S. J. Corey et al. (eds.), *A Systems Biology Approach to Blood,*
Advances in Experimental Medicine and Biology 844, DOI 10.1007/978-1-4939-2095-2_1

Blood, pure and eloquent, wrote Max Wintrobe, one of the pioneers of modern hematology. His description was but a reference to lines written by the seventeenth-century English poet John Donne. Hematology, the study of blood and its components, has undergone dramatic changes over the millennia, since man first recognized its power. The first plague visited upon the Egyptians was blood, "I will strike the water of the Nile, and it will be changed into blood. The fish in the Nile will die, and the river will stink and thus the Egyptians will not be able to drink its water." What we consider so vital to human life was viewed as deleterious. Rabbis later writing commentary warned against circumcising a third son after two had died of bleeding. Contemporaneously, Hippocratic writings described blood as one of the four humors. More appreciative of its vital nature, the Greek physicians equated blood with spring and air. The Greek word for blood, *haima*, has been sustained in all things hematologic and hematopoietic.

Like other branches of medicine, hematology has undergone paradigm shifts. From the ancient Jews' and Greeks' attribution of blood to health and disease through the seventeenth century's rationalists who described its circulation through arteries and veins, to the modern physiologists of the past century, our understanding of blood and its components has advanced. The past 50 years have provided us with more intimate knowledge of its components at the subcellular level. This reductionist approach to science has now been superceded by the awareness of complex, large datasets, made possible by proteomic, flow cytometric, microarray, genomic sequencing, and epigenetics. The complexity of blood and its components is also recognized at multiple levels from the subcellular to the macroscopic, such as the environment and its effect on the organism. While physical and chemical laws have been applied to biology, limitations to their applicability and predictability are frequently encountered. Biology is dynamic.

The biomedical discipline that has been called physiology has evolved to a new approach—a modern synthesis of biochemistry, genetics, mathematics, engineering, and machine-based learning. Complex, large datasets of genes, lipids, metabolites, and proteins have made it impossible for one investigator to intuit the whole. This new, integrative field has been called systems biology. In this volume, we seek to introduce physicians and scientists, qualitative and quantitative, to the different facets of systems hematology. Systems hematology embraces this complexity, utilizing engineering principles and computational methods to build and validate models using experimental data. The approach rests on (i) defining all (or the known knowns) of the components, (ii) systematically perturbing and monitoring the components of the system, (iii) reconcile the experimentally observed responses with those predicted by the model, and (iv) designing and performing new experiments to distinguish between multiple or competing models. The goals are to understand how the system works, identify new systems-based properties, and predict outcomes.

The major obstacle to success in systems biology lies in the disciplines practiced by physicians and scientists. Major differences exist in the methods, jargon, and philosophies between quantitative scientists, the theoretical physicists, the mathematicians, the engineers, computer programmers, and experimentalists. Even within the experimentalists, there is diversity and increasing technologization, as evidenced

by cell biology, molecular biology, and proteomics. Until there is a common vernacular, fundamental concepts in the fields of biology, mathematics, engineering, and computation can be understood and transdisciplinary studies can be successful.

Blood as a System

Biological systems operate at multiple levels (or scales): molecular, cellular, tissue, and organismal, and environmental. Stem cells generate differentiated blood cells through a continuous process of asymmetric stem cell division, yielding daughter cells with different capacities for renewal or differentiation. This process occurs in a specialized microenvironment. The blood system consists of highly specialized cells and plasma containing a range of proteins to regulate different processes. Among the blood cells are erythrocytes that shuttle oxygen or its waste product to and from tissues; white blood cells to fight infection and mediate inflammation; and platelets to stop bleeding. Within the compartment of white blood cells, there is variability: neutrophils to engulf foreign agents, lymphocytes to make antibodies and coordinate immunity, and monocytes to process and regulate host defense. Plasma contains more than 1000 proteins [1]. Homeostatic mechanisms insure that the right number of cells is produced, but they are sufficiently dynamic to meet the needs of environmental changes (e.g., hypoxia, infection, or bleeding). While hematologists diagnose and treat patients with anemias, immune deficiencies, leukemias and lymphomas, and hypercoagulability, it is astonishing that such high level of quality control of blood and its elements exists and that blood diseases are not more common.

Systems Properties in Hematopoiesis

Because of the facility in sampling blood or bone marrow repetitively and quantitatively, the blood system is well suited for modeling and validation. Hematopoiesis and the functioning of specialized blood cells involve complex processes that can be examined at the level of genes [2], signal transduction proteins [3], or the population distribution of diverse cell types [4]. Both deterministic and stochastic processes contribute. By viewing hematopoiesis (cell proliferation and differentiation) as a dynamic system and disease as perturbations of the system, one can learn more about both disease and physiological states.

Proliferation and loss are fundamental properties of hematopoietic stem cells and their progeny. Population dynamics offers a quantitative approach in studying them. Asymmetric division results in a stem cell dividing into either another stem cell or a more committed cell, while symmetric division yields either two stem cells or two differentiated daughter cells. These processes can be combined in a series of short steps [5–8]. Models built around these division (a)symmetries usually result in exponential cell growth, but such growth cannot be realistically sustained in vitro

due to spatial and nutrient limitations. Models based on heterogeneous population account for cell proliferation and loss due to death or differentiation.

Differentiation is the other fundamental property of hematopoietic progenitor cells and requires critical processes of cell fate decision making. Decision making occurs as a result of biochemical signaling and gene regulatory networks within the cell [9], [10]. Ultimately, transcription factors determine cellular differentiation and specialization [11]. The relative contributions of instructive and permissive programming in hematopoiesis have long been debated [6, 12–23]. To describe hematopoietic stem cell renewal and differentiation, deterministic and stochastic models have been constructed. James Till, a biophysicist, and Ernest McCulloch, a physician, pioneered the study of hematopoiesis in the early 1960s through their development of a quantitative spleen colony assay, establishment of a hematopoietic stem cell, and data analysis that yielded a stochastic model of hematopoiesis [24], [25]. In their stochastic model [5], cells have two possible fates: (1) differentiate and leave the proliferative compartment or (2) undergo symmetric division forming two colony-forming cells. Each fate was assigned a probability. Drawing random numbers to determine the fate of each cell, Till and McCullouch calculated the diversity of stem cell populations after the course of several generations. Colony generation appears as a well-defined process even though individual cell-fate decisions are random. Regulation acts at the population, not cellular, level and the population of stem cells can be affected by influencing processes that define the effective probabilities of birth and death.

A cell uses complex intracellular signaling and gene regulatory networks in order to integrate the multiplicity of cues in its environment and to ultimately make a specific decision. In particular, gene regulatory networks have provided great insights into lineage commitment of hematopoietic progenitors.

Types of Mathematical Models

Different methods of modeling have been developed to describe and predict biological processes. Not all models are accurate, but some are more useful than others. Deterministic models describe the state of a system over time in the absence of random events. These always produce the same output for a given input [26]. In contrast, stochastic models describe the effects of randomness and noise on system output [27]. Statistical models use existing data to estimate a functional relationship between system input and output. Network models graph the direction and magnitude of interactions that exist between the various components in a system [28].

Deterministic models typically consist of one or more differential equations, with each equation describing the change in a system state variable over time, as it depends on other system variables and rates. If the state variable of interest is the number of cells in the population, a differential equation modeling the change in the population over time would consist of the difference between rates of cell production and rates of cell loss:

$$\frac{dN_X}{dt} = (\text{rate at which precursor of X differentiates into X}) - (\text{rate at which X differentiates into next cell lineage}) - (\text{rate at which X dies}) \quad (1.1)$$

where N_X is the number of cells of type X.

Each equation describes the rate of change in the number of cells of given type and maturity in the system by including terms for the rates of cell production, death, and differentiation. Once the equations are established, they are solved either analytically or numerically to determine the population's functional dependence on time. In models describing physiological conditions, the equations tend toward a steady-state solution representing system homeostasis; that is, after sufficient time has elapsed, positive and negative contributions to cell number balance and the population attains a constant level (e.g., $dN_X/dt = 0$ in Eq. 1.1). For disease-state cell populations, other types of behavior such as oscillations or uncontrolled growth are frequently modeled.

Stochastic models are employed to examine the effects of intrinsic and extrinsic randomness on a system. Intrinsic randomness arises from interactions of a finite ("small") number of discrete components, e.g., binding of a given gene's promoters (two copies per diploid genome) by transcription factor's molecules (also a limited number). Extrinsic randomness arises either from variability (genetic and phenotypic) among cells or from environmental fluctuations. The most common type of stochastic model is a Markov process, in which the future state of the system depends only on its current state and is independent of its past states. Monte Carlo simulations are an empirical method to investigate dynamics of a stochastic system, by generating repeated random trajectories and computing frequencies that estimate probability distributions.

Statistical models are sometimes confused with stochastic models. Whereas stochastic models reflect the structure of the biological system, statistical models are data driven. Statistical models can be employed even when no knowledge about system's structure exists and can generate predictions, which may be only statistically validated. However, some statistical models such as Bayesian networks may provide insights concerning the structure. Bayesian network models are built from graphs in which the states of and relationships between network elements are probabilistic. While graph theoretical models can be circular, Bayesian networks have a definite, distinct set of termini. These models have a wide range of uses. For example, a Bayesian network model could be used to predict the probabilities of certain cellular mutations based on abnormalities in protein expression levels (assuming, of course, that there is a relationship between the two). Their structure and necessary constants have to be estimated based on data. Though popular, Bayesian networks suffer from the possible reversal of causality [29].

Network models have recently gained popularity in the social, physical, and biological sciences from the widespread application of graph theory**,** an area of mathematics that investigates the relationships between the objects of a group [30]. Graph theory lends itself to visual representations making it an appealing tool for biologists investigating phenomena ranging from the interactions between populations

in an ecosystem to the interactions between molecular species involved in a signaling pathway. At its simplest, a graph is a map of all known system components or system states and their possible interactions or transitions. Circles (nodes) represent components and states, and lines and arrows (branches or edges) represent relationships between nodes. Graphs help portray topological structures such as loops. Complex dynamics can arise from relatively few interacting components [31], and network maps are widely used to help visualize the interactions. Building upon existing graph theoretical notation, an international group has developed Systems Biology Graphical Notation to standardize the visual representations used to describe biological interaction networks [32].

Current Status of Systems Biology

The success of systems analysis of hematopoiesis will depend upon technological breakthroughs and collaborations between the biological and physical sciences that yield accurate predictions and emergent properties. With each discipline using a different language, this is easier said than done. Changes in undergraduate, graduate, and medical curricula must be implemented to train a new generation of biomedical researchers fluent in quantitative or engineering disciplines [33–35]. Systems biology requires a balance between models sufficiently complex to describe a system and yet simple enough to be clinically useful. Understanding large quantities of data well enough to validate a model is especially challenging. The development of Systems Biology Markup Language (SBML) has made it easier to develop biology-oriented software packages, such as COPASI, Simmune, MetaCore, and Cytoscape, which aid model building and data analysis [32, 36–39]. Since 2001, the number of such packages developed for systems biology has grown from 5 to over 170. With computational power becoming ever greater and cheaper, the number and diversity of such software packages will only increase, bringing within their scope models that may not be impossible to validate with current technology. At present, most models of hematopoiesis are built at a single scale, e.g., cellular or molecular. The future lies in building models that span multiple scales, incorporating more of the connections that exist between them and thereby being able to account for some of the complexity that arises from the connections. Among the fundamental questions in normal and leukemic hematopoiesis that systems biology will address are: integration of signaling pathways, circuits, and networks that determine cell fate, multi-scale modeling of stem cell plasticity, synthesis of genetic and epigenetic data, global analysis of phosphoproteins, dynamics of hematopoiesis in the bone marrow microenvironment presented in three-dimensional imaging, and cellular engineering to expand selective blood cell compartments for therapy. The complexity or density of experimental data will demand a systems approach. More in-depth coverage may be found in the few textbooks of systems biology and bioinformatics that have appeared, none solely devoted to hematologic topics [40–43].

References

1. Qian WJ, Monroe ME, Liu T, et al. Quantitative proteome analysis of human plasma following in vivo lipopolysaccharide administration using 16O/18O labeling and the accurate mass and time tag approach. Mol Cell Proteomics. 2005;4(5):700–9.
2. Laslo P, Pongubala JM, Lancki DW, Singh H. Gene regulatory networks directing myeloid and lymphoid cell fates within the immune system. Semin Immunol. 2008;20(4):228–35.
3. Hlavacek WS, Faeder JR, Blinov ML, Posner RG, Hucka M, Fontana W. Rules for modeling signal-transduction systems. Sci STKE. 2006;17(344):re6.
4. Roeder I, Loeffler M. A novel dynamic model of hematopoietic stem cell organization based on the concept of within-tissue plasticity. Exp Hematol. 2002;30(8):853–61.
5. Till JE, McCulloch EA, Siminovitch L. A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. Proc Natl Acad Sci U S A. 1964;51:29–36.
6. Ogawa M. Hemopoietic stem cells: stochastic differentiation and humoral control of proliferation. Environ Health Perspect. 1989;80:199–207.
7. Ogawa M, Pharr PN, Suda T. Stochastic nature of stem cell functions in culture. Prog Clin Biol Res. 1985;184:11–9.
8. Vogel H, Niewisch H, Matioli G. The self renewal probability of hemopoietic stem cells. J Cell Physiol. 1968;72(3):221–8.
9. Kestler HA, Wawra C, Kracher B, Kuhl M. Network modeling of signal transduction: establishing the global view. Bioessays. 2008;30(11–12):1110–25.
10. Kirouac DC, Madlambayan GJ, Yu M, Sykes EA, Ito C, Zandstra PW. Cell-cell interaction networks regulate blood stem and progenitor cell fate. Mol Syst Biol. 2009;5:293.
11. Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. Cell. 2008;132(4):631–44.
12. McCulloch EA. Stem cells in normal and leukemic hemopoiesis (Henry Stratton Lecture, 1982). Blood. 1983;62(1):1–13.
13. McCulloch EA. Stem cell renewal and determination during clonal expansion in normal and leukaemic haemopoiesis. Cell Prolif. 1993;26(5):399–425.
14. Mehr R, Agur Z. Bone marrow regeneration under cytotoxic drug regimens: behaviour ranging from homeostasis to unpredictability in a model for hemopoietic differentiation. Biosystems. 1992;26(4):231–7.
15. Morrison SJ, Weissman IL. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. Immunity. 1994;1(8):661–73.
16. Novak JP, Stewart CC. Stochastic versus deterministic in haemopoiesis: what is what? Br J Haematol. 1991;78(2):149–54.
17. Ogawa M. Stochastic model revisited. Int J Hematol. 1999;69(1):2–5.
18. Quesenberry P, Abedi M, Dooner M, et al. The marrow cell continuum: stochastic determinism. Folia Histochem Cytobiol. 2005;43(4):187–90.
19. Abkowitz JL, Catlin SN, Guttorp P. Evidence that hematopoiesis may be a stochastic process in vivo. Nat Med. 1996;2(2):190–7.
20. Roeder I, Glauche I. Towards an understanding of lineage specification in hematopoietic stem cells: a mathematical model for the interaction of transcription factors GATA-1 and PU.1. J Theor Biol. 2006;241(4):852–65.
21. Palani S, Sarkar C. Integrating extrinsic and intrinsic cues into a minimal model of lineage commitment for hematopoietic progenitors. PLoS Comput Biol. 2009;5:e1000518.
22. Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature. 2008;453(7194):544–7.
23. Huang S, Guo YP, May G, Enver T. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. Dev Biol. 2007;305(2):695–713.
24. Becker AJ, Mc CE, Till JE. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. Nature. 1963;197:452–4.

25. Siminovitch L, McCulloch EA, Till JE. The distribution of colony-forming cells among spleen colonies. J Cell Physiol. 1963;62:327–36.
26. Fall C, Marland E, Wagner J, Tyson J. Computational cell biology. Vol. 20. New York: Springer; 2005.
27. Wilkinson D. Stochastic modelling for systems biology. New York: Chapman & Hall/CRC; 2006.
28. Palsson B. Systems biology: properties of reconstructed networks. Cambridge: Cambridge University Press; 2006.
29. Pearl J. Models, reasoning and inference. New York: Cambridge University Press; 2000.
30. Barabasi AL. Scale-free networks: a decade and beyond. Science. 2009;325(5939):412–3.
31. Amaral LA, Diaz-Guilera A, Moreira AA, Goldberger AL, Lipsitz LA. Emergence of complex dynamics in a simple model of signaling networks. Proc Natl Acad Sci U S A. 2004;101(44):15551–5.
32. Le Novere NH, Mi H, et al. The systems biology graphical notation. Nat Biotechnol. 2009;27(8):735–41.
33. HHMI/AAMC. Scientific foundations for the future physicians. 2009. https://www.aamc.org/download/271072/data/scientificfoundationsforfuturephysicians.pdf. Accessed 12 Oct 2014.
34. Council NR. BIO2010: transforming undergraduate education of future research biologists. Washington, DC: National Academies Press; 2003.
35. Wingreen N, Botstein D. Back to the future: education for systems-level biologists. Nat Rev Mol Cell Biol. 2006;7(11):829–32.
36. Killcoyne S, Carter GW, Smith J, Boyle J. Cytoscape: a community-based framework for network modeling. Methods Mol Biol. 2009;563:219–39.
37. Meier-Schellersheim M, Xu X, Angermann B, Kunkel EJ, Jin T, Germain RN. Key role of local regulation in chemosensing revealed by a new molecular interaction-based modeling method. PLoS Comput Biol. 2006;2(7):e82.
38. Mendes P, Hoops S, Sahle S, Gauges R, Dada J, Kummer U. Computational modeling of biochemical networks using COPASI. Methods Mol Biol. 2009;500:17–59.
39. Moore JH. Bioinformatics. J Cell Physiol. 2007;213(2):365–9.
40. Alon U. An introduction to systems biology: design principles of biological circuits. Boca Raton: Chapman & Hall/CRC; 2007.
41. Palsson BO. Systems biology, properties of reconstructed networks. New York: Cambridge University Press; 2006.
42. Kitano H. Foundations of systems biology. Cambridge: MIT Press; 2001.
43. Polanski A, Kimmel M. Bioinformatics. New York: Springer; 2007.

http://www.springer.com/978-1-4939-2094-5

A Systems Biology Approach to Blood
Corey, S.J.; Kimmel, M.; Leonard, J.N. (Eds.)
2014, IX, 403 p. 73 illus., 52 illus. in color., Hardcover
ISBN: 978-1-4939-2094-5