AMERICAN SOCIETY OF HEMATOLOGY
2021 L Street, NW, Suite 900, Washington, DC 20036 ph 202.776.0544 fax 202.776.0545 e-mail ASH@hematology.org

March 28, 2013

Re: Request for Information: Research Needs to Facilitate Broad Community Use of the Roadmap Epigenomics Program Data Resource (NOT-RM-13-011).

Submitted electronically to: epigenomics@niehs.nih.gov

To Whom It May Concern:

The American Society of Hematology (ASH) appreciates the opportunity to respond to the Request for Information (RFI): Research Needs to Facilitate Broad Community Use of the Roadmap Epigenomics Program Data Resource (NOT-RM-13-011) issued on February 22, 2012.

ASH represents more than 14,000 clinicians and scientists worldwide committed to the study and treatment of blood and blood-related diseases. These diseases encompass malignant hematologic disorders such as leukemia, lymphoma, and multiple myeloma and non-malignant conditions such as sickle cell anemia, thalassemia, aplastic anemia, venous thromboembolism, and hemophilia. In addition, hematologists have been pioneers in the fields of stem cell biology, regenerative medicine, bone marrow transplantation, transfusion medicine, gene therapy, and the development of many drugs for the prevention and treatment of heart attacks and strokes. ASH membership is comprised of basic scientists, physician scientists, and physicians working in diverse settings, including universities, hospitals and private practices.

Hematologists are currently working on some of the most complex and costly medical problems that affect health care in the United States. Our members are pioneers in development of cellular therapies for chronic and acute hematologic diseases that have been translated into other fields of medicine. Hematologists are developing novel treatments for anemia associated with chronic diseases like cancer and chronic kidney disease. They are also devising alternatives to blood transfusions and enhancements of bone marrow transplantation through the use of umbilical cord blood, stem cells and other technologies. In addition, hematologists have pioneered the use of epigenetics to understand normal blood development and malignant blood diseases, and have shown that somatic mutations in epigenetic regulators and site-specific alterations in epigenetic patterning are characteristic of most leukemias, lymphomas, and other blood malignancies. As such, the scientific community represented by ASH has been very active in elucidating basic mechanisms of epigenetics, discovering how alterations in the epigenome contribute to disease states, and developing specific therapies which target the epigenome to improve outcome for patients with hematologic disorders.

ASH applauds the NIH for its efforts to facilitate broad use of the public data resource developed by the Roadmap Epigenomics Program. The scientific community relies on this important resource to further our understanding on the role of epigenomics in mechanisms of normal development and disease. For example, studies of DNA methylation in normal and diseased states have led to new insights into the pathogenesis of leukemias and lymphomas. The Roadmap Epigenomics Program has allowed hematologists to access a framework for normal DNA methylation patterns as a reference for studies of normal and malignant hematopoiesis. In addition, many of the tissues studied in the Roadmap are of hematologic origin and they represent invaluable resources for the studies of self-renewal, differentiation and transformation. As such these epigenetic maps have greatly empowered studies of blood development and disease states.

ASH strongly supports continuation of the public data resource developed by the Roadmap Epigenomics Program and urges the NIH to expand the resource to include additional tissue types with particular focus on primary specimens, recently identified epigenetic marks, and studies of specific diseases of importance to the hematology community and to patients worldwide. Below are ASH's comments that address the specific questions posed in the RFI, with the Society's recommendations emphasized in *italics*.

Data Presentation and Accessibility:

The major roadblock experienced by hematologists when using reference epigenome data generated by the Roadmap Epigenomics Program is the limited amount of information that is available in purified hematopoietic cells.

A limited number of histone modifications have been studied to date, and generally only one DNA "methylation" assay has been used to assess genome-wide methylation on white blood cells from young (21-43 year old) individuals. Because hematopoiesis is characterized by dynamic changes in stem cell function and in differentiation with increasing age (e.g., increased myelopoiesis), *broadening the age of the samples studied is critical*.

In addition, cells from the white blood cell lineage are currently being studied; there are no data from erythroid or megakaryocytic lineages. The majority of the "stem" cells studied are derived from mobilized CD34+ cells, and only a subset of these represent multipotent hematopoietic stem cells, as it is known that the process of peripheral blood mobilization affects the hematopoietic phenotype and epigenetic state.

ASH believes that parallel studies of the epigenome of hematologic tissues from patients with leukemia, lymphoma, myeloma and of specific benign hematologic diseases including sickle cell anemia, thalassemia, and immunodeficiencies would be of importance to the field at large and would represent an invaluable resource for the research community.

ASH has additional recommendations to make reference epigenome data generated by the Roadmap Epigenomics Program useful to the broader research community:

- These data are found on a separate website from ENCODE and TCGA. The ENCODE and TCGA initiatives are also studying samples and epigenetic marks of relevance to the scientific community. *Creating one unified website where all of the data would be easily accessible would be very helpful to the community. Incorporating SNP/allele frequency data from 1000 Genomes would also be useful.*

- For the data to be truly accessible to every researcher, it should be presented in a format such that any researcher, irrespective of whether he or she has any bioinformatics training, can easily explore it. *An effort towards improving data visualization by either incorporating it into existing tools such as the UCSC genome browser or creating a similar independent interface would greatly enhance access to the findings of the Roadmap Epigenomics Program by the greater scientific community.*

- Many of the methods being studied (e.g., bisulfite sequencing and RRBS) cannot distinguish 5-methylC from 5-hydroxymethylC. This limitation is not made explicit. *Incorporation of data using methods that can distinguish between different DNA modifications, which likely have different functionality, is critical.*

Data Analysis:

The most extensive data available in reference epigenome datasets come from categorizing histone marks, and therefore, studies that examine histone marks and the effects of specific mutations on chromatin-modifying enzymes will be able to use these data most directly.

To facilitate analysis of user-generated datasets relative to reference epigenomic data, ASH recommends developing clear guidelines/recommendations for the interpretation of Illumina 450K data that recognize the data's limitations.

Many investigators do not realize that Illumina (or any method that relies on sodium bisulfite-treated DNA) cannot distinguish 5-methylC from 5-hydroxymethylC. Furthermore, most investigators do not realize that about 140,000 probes on this array do not map uniquely to human sequence after bisulfite treatment. The company's choice to place these probes on the array implies that they can be assigned uniquely, but they map to multiple loci and should be removed from analysis. In addition, many of the probes also contain SNPs whose allele frequencies vary across populations, and in some contexts, should be removed for this reason. *Orthogonal platforms are needed to provide the essential methylation data to the hematopoiesis (and broader) field.*

Data Integration:

There remain several roadblocks that prevent integration of reference epigenome data with other kinds of data:

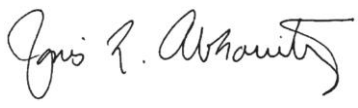
- Generally, publically-available databases (like TCGA) are heavily weighted to samples derived from Caucasian individuals, which makes doing population-based studies or analysis of data from minority groups very challenging. *An effort to expand the number of studies performed on members of minority/under-served groups would be important, particularly given the predisposition for different hematopoietic disorders in different ethnic subsets.*

- Once an investigator uses publically-available data to generate a hypothesis, there is no way to gain access to the actual samples to validate/expand upon a hypothesis. The investigator has to generate her/his own cohort of samples. *ASH urges the NIH to make representative samples studied in the Roadmap be accessible for researchers who wish to functionally interrogate specific aspects of the Roadmap dataset.*

- There is also a need for standardization of analytical methods to integrate different epigenomic data types. Investigators are generally adapting software packages designed for ChIP-Seq data to study other methods of DNA capture. It is not clear whether the assumption that these methods are similar enough to justify the common use of these software packages is valid. *ASH recommends that the NIH involve bioinformatics specialists, hematology researchers, and clinician scientists in discussing how best to analyze, present, and annotate epigenomic data.* ASH membership has experts in each of these areas who would be willing to contribute to building a foundational set of bioinformatics approaches to epigenomic datasets from normal samples and disease states.

The American Society of Hematology looks forward to working with the NIH on this important program, provide further information and be a resource for the NIH. Please contact ASH Senior Manager for Scientific Affairs, Ulyana V. Desiderio, Ph.D., at (202) 776-0544 or udesiderio@hematology.org for any additional information.

Sincerely,



Janis L. Abkowitz, MD
President