## MOLECULAR BIOLOGY

# The NIH Common Fund/Roadmap Epigenomics Program: Successes of a comprehensive consortium

John S. Satterlee[1,2], Lisa H. Chadwick[2]*, Frederick L. Tyson[3], Kim McAllister[3], Jill Beaver[4,5], Linda Birnbaum[3], Nora D. Volkow[2], Elizabeth L. Wilder[4,5], James M. Anderson[5], Ananda L. Roy[4,5]†

The NIH Roadmap Epigenomics Program was launched to deliver reference epigenomic data from human tissues and cells, develop tools and methods for analyzing the epigenome, discover novel epigenetic marks, develop methods to manipulate the epigenome, and determine epigenetic contributions to diverse human diseases. Here, we comment on the outcomes from this program: the scientific contributions made possible by a consortium approach and the challenges, benefits, and lessons learned from this group science effort.

## WHY DEVELOP AN EPIGENOMICS PROGRAM?

The development of the National Institutes of Health (NIH) Roadmap Epigenomics Program was informed by (i) critical scientific gaps and opportunities, (ii) input from the scientific community, and (iii) analysis of funded NIH research in this area. Epigenomics can be operationally defined as the study of structural and functional DNA and histone modifications that alter the reading and writing of the genome, resulting in the regulation of chromatin architecture, gene activity, and expression without changes to the DNA sequence. Epigenetic marks are those DNA and histone modifications that occur at specific loci within the genome. While analyses of epigenetic regulation of gene expression date back to the 1970s and 1980s, methods to analyze epigenetic modifications at a genome-wide scale were not developed until the early 2000s (1, 2). Advances in DNA sequencing technology, the development of methods such as bisulfite sequencing and chromatin immunoprecipitation sequencing, and generation of highly specific antibodies against posttranslationally modified histones created an opportunity to generate cell-specific epigenomic maps.

Furthermore, a portfolio analysis conducted by the NIH in 2006 on epigenetic/epigenomic research revealed increasing numbers of NIH-funded studies and publications associated with altered gene expression profiles and epigenetic processes between 1998 and 2006. This analysis revealed that limited research was being conducted on specific diseases except for cancer, which constituted more than half of disease epigenetics publications. The majority of these publications defined epigenomes as DNA methylation profiles of tumors, and most studies focused on the area of chromatin regulation and DNA methylation in animal models. Genome-wide analysis of primary human tissues was challenging at the time due to low detection sensitivity and the need for large numbers of cells; hence, there was limited effort to analyze human tissues or to establish human reference epigenomes. Overall, the scientific literature suggested that many conditions could involve altered epigenetic mechanisms, but

testing those hypotheses—particularly by surveying epigenetic marks across the genome in human samples—remained a challenge.

Further difficulties at that time included (i) the potentially limitless definition of the "human epigenome" because epigenetic marks are dynamic and vary from cell to cell (both between individual cells of the same type and between different cell types) and (ii) the need to describe an epigenomic "ground state" in pluripotent cells and how this state changes during differentiation. In addition, grassroots scientific groups had been encouraging a large-scale effort in the area of epigenomics for some time (3). Several scientific meetings were held to identify the best ways to accelerate epigenomic discovery including one held in 2007 to specifically identify what would ultimately become the themes of this program. These gaps and opportunities, coupled with the analysis of scientific areas already addressed by the NIH grant portfolio and input from the scientific community, ultimately informed plans for a trans-NIH Roadmap Program in Epigenomics. (The NIH Roadmap is now known as the NIH Common Fund, but the Roadmap Epigenomics Program retains its name.)

## OVERALL GOALS OF THE ROADMAP EPIGENOMICS PROGRAM

The overarching goal of the program was to establish a set of human reference epigenomes and develop new technologies as fundamental resources for the scientific community to conduct basic and applied research on how epigenetic/epigenomic processes contribute to human development, life span/aging, response to environmental exposures (e.g., physical, chemical, behavioral, and/or social), and disease pathogenesis. The details of the program goals and the funded projects have been previously described (4, 5). The program was designed recognizing that while highly coordinated consortium-based research was required to generate and analyze epigenomic maps across many human cell types and tissues, substantial innovation and discovery through smaller projects, in particular for technology development, were also needed (4). Because of our poor understanding of the epigenomic differences between cell types and between individuals, it was not clear how many samples of a given tissue would need to be analyzed or what epigenomic marks had to be identified to generate a representative epigenomic map. The limited availability of many human cell types and tissues highlighted the need for epigenomic technologies with higher sensitivity. Thus, substantial technology development (including isolating homogeneous cell populations from a given tissue, interrogating extremely limited

[1]Division of Neuroscience and Behavior, National Institutes of Health (NIH), 6001 Executive Blvd., Bethesda, MD 20892, USA. [2]National Institute on Drug Abuse, NIH, 6001 Executive Blvd., Bethesda, MD 20892, USA. [3]National Institute of Environmental Health Sciences, NIH, 111 TW Alexander Dr., Durham, NC 27709, USA. [4]Office of Strategic Coordination, Office of the Director, NIH, 6001 Executive Blvd., Bethesda, MD 20892, USA. [5]Division of Program Coordination, Planning, and Strategic Initiatives, Office of the Director, NIH, 1 Center Dr., Bethesda, MD 20892, USA.
*Present address: Division of Genome Sciences, National Human Genome Research Institute, NIH, 6700B Rockledge Dr., Bethesda, MD 20892, USA.
†Corresponding author. Email: ananda.roy@nih.gov

numbers of cells, and enhancing assay throughput) were necessary to complement the consortium-wide mapping efforts. This combination of integrating large-scale mapping efforts with small-scale discovery-driven approaches and technology development ultimately proved to be critical for the success of this program.

The Roadmap Epigenomics Program ended in 2018. We reflect here on the program achievements to date, applying objective measures while acknowledging that the long-term impact of the program is yet to be measured. Given the vast quantities of data generated in the postgenomic era and the ever-increasing need for integrating multidisciplinary approaches to solve today's biomedical problems, we also evaluate the lessons learned from this large consortium-based approach as well as the many challenges and benefits offered by "group science."

## ROADMAP EPIGENOMICS PROGRAM DELIVERABLES

The Roadmap Epigenomics Program was designed to deliver community resources that would catalyze and expand investigator-initiated epigenomic research across the NIH (4). This effort was distinct from previous efforts in three broad ways: (i) The focus was on human tissues and cells as opposed to animal cells and models; (ii) epigenomic features were studied genome wide rather than at a single locus or a handful of loci; and (iii) an emphasis was placed on discovering novel epigenomic marks. While at this time the ENCODE (Encyclopedia of DNA Elements) Project was cataloging regulatory elements in animal models or human cells grown in culture, the aim of the Roadmap Epigenomics Project was to build on this foundation by analyzing samples taken directly from human tissues and cells (6). The deliverables of the Roadmap Epigenomics Program included (i) reference epigenomic maps, (ii) international epigenomic coordination, (iii) a deeper understanding of the epigenomic basis of disease, (iv) identification of previously unknown epigenetic marks, and (v) improved technologies and methods for monitoring and manipulating epigenomic modifications.

### Reference epigenomic maps

One major focus of the Roadmap Epigenomics Program was to provide reference epigenomic maps for normal human cells and tissues (4). This focus required a goal-driven highly coordinated approach that could not be easily accomplished by individual laboratories (7). Fundamental to achieving this effort was the definition of a "reference epigenome" and a plan for the coordinated production of these datasets by investigator groups using different methods. The consortium approach paved the way to establishing global maps of multiple epigenomic modifications across the genome in different types of human cells and tissues. Simultaneously, showing that disease- and trait-associated genetic variants are enriched in tissue-specific epigenetic signatures was critical in revealing important cell types implicated in specific genetic traits as well as in providing a molecular basis for interpreting human disease types as a community resource (8). Similar to the human genome sequences that set the stage for far-reaching studies of genetic variation association with disease phenotypes (9), a set of 111 human reference epigenomes profiling comprehensive histone modification signatures, DNA accessibility and methylation patterns, as well as RNA expression was described in an integrative analysis paper (10). Furthermore, the data from these experiments have been incorporated into the ENCODE Portal and can now be searched with ENCODE and other related data to reveal human regulatory elements and the cognate transcription factors associated with these elements.

The Human Reference Epigenome Maps generated by the Roadmap Epigenomics Program have provided a deeper and more comprehensive view of our regulatory genome in terms of defining regulatory elements, such as promoters and enhancers, for a given tissue or cell type and have complemented the information obtained from programs like ENCODE and GTEx (Genotype-Tissue Expression). These maps have been used to predict tissue-specific patterns of disease and have informed the functional analysis of numerous genome-wide association study (GWAS) hits relevant to many complex human diseases. The prioritization of GWAS findings to identify causal loci for disease pathways has been a challenge in the human genomics field for some time. The maps generated now allow researchers to prioritize disease-associated variants that overlap GWAS-enriched epigenomic annotations. Additional tools have been developed and defined in recent years that allow researchers to continue to explore genomic regions with tissue- or cell-specific epigenomic features, using Roadmap Epigenomics Program data with greater sophistication (11). Many studies have also used the integration of histone marks, chromatin states, and transcription factor binding site data from Roadmap and ENCODE to inform epigenomic signatures or develop epigenomic biomarkers related to complex human diseases (12).

### International epigenomics coordination

The Roadmap Epigenomics Program was also a founding member of the International Human Epigenome Consortium (IHEC) created in 2010. The primary goal of IHEC was to pursue an international effort in coordinating the production of reference maps of human epigenomes for key cellular states relevant to human health and diseases. A critical component of IHEC goals was also to coordinate the development of common bioinformatics standards, data models, and analytical tools to organize, integrate, and display the epigenomic data generated by the broad international community. An important step toward this goal was the first installment of 41 coordinated papers showcasing the achievements and scientific progress made by IHEC in core areas of current epigenomic investigation (13). Both Roadmap and IHEC were designed to improve compatibility and interoperability of diverse datasets while ensuring that data generation efforts would not be duplicative. Accordingly, the Roadmap Epigenomics Mapping Consortium developed a common set of metadata standards and data standards that were, in part, based on standards developed by the ENCODE program. These standards have been adopted by IHEC and further refined in collaboration with other stakeholder groups. Roadmap Epigenomics Program participation in IHEC also helped spur significant additional international investment in epigenomic mapping and disease research. As of September 2018, IHEC members, including Roadmap and ENCODE, have generated 8870 epigenomic datasets (http://epigenomesportal.ca/ihec/).

### Epigenomic basis of diseases and novel epigenomic marks

It is important to point out that in addition to coordinated mapping efforts, the Roadmap Epigenomics Program included discovery-driven investigations into the role of epigenomic changes in a wide array of diseases. These projects generated foundational knowledge necessary for translation of epigenomic information into improved preventive, diagnostic, and therapeutic strategies. Here, we provide

## Epigenomics Program publications 2007–2017 mean and median RCR by year (*n* = 652)
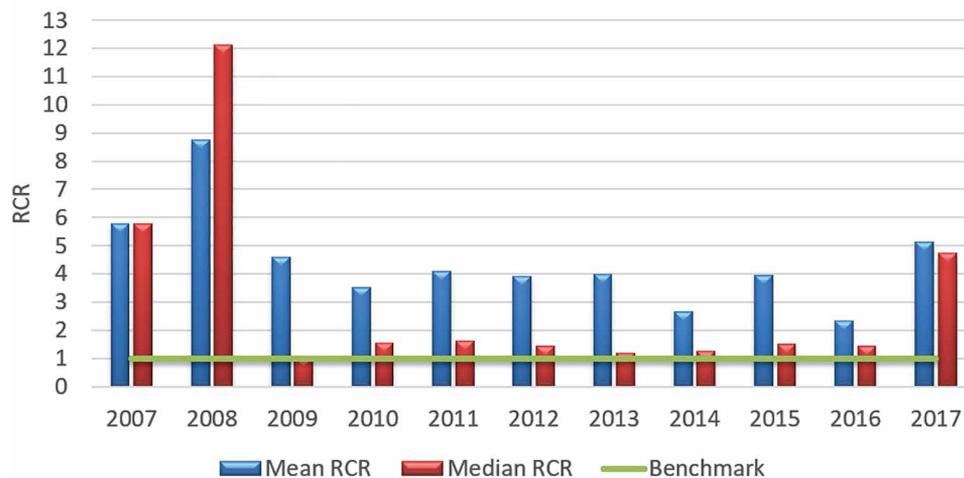


**Fig. 1. Mean and median RCR (Relative Citation Ratio) of Roadmap Epigenomics Program research articles for each year.** The RCR benchmark of 1.0 is based on the median RCR for all NIH-funded publications, including reviews. This excludes 2017 articles that did not have an assigned RCR value at the time of analysis.

some notable examples of program activities driven by individual efforts. For instance, De Jager *et al.* (*14*) took advantage of two prospective studies of aging. They identified 11 regions of the genome in which brain frontal cortex methylation status was associated with Alzheimer's disease pathology in both a discovery cohort and a smaller replication cohort. This study also identified eight nearby genes that were differentially expressed in Alzheimer's disease brains relative to unaffected controls. In a second study, Reynolds *et al.* (*15*) identified age-associated methylation changes in monocytes and T cells, which correlated with both local gene expression changes and clinical measures of vascular aging. These epigenomic changes were enriched in regions predicted to have roles in regulating gene expression programs, such as enhancers. A third example was provided by Yang *et al.* (*16*) who compared DNA methylation profiles of peripheral blood mononuclear cells from inner city children with asthma to a healthy control population and identified 11 genes with methylation and gene expression changes specifically associated with asthma, all of which were validated in an independent cohort. Many of these genes are related to immune function, specifically T cell maturation and $T_H2$ (T helper 2) immunity, raising the possibility that in the future, epigenetic therapies could be effective for reversing immune dysfunction in children with asthma. In 2010, the Ren laboratory published a highly cited paper comparing the epigenomes of pluripotent and lineage-committed human embryonic stem cells (*17*), one of the many advances made by the Roadmap Consortium in stem cell epigenomics.

In addition, a search for novel epigenomic marks was fundamental to the design of the overall program to help more fully understand the diversity of epigenomic modifications that might need to be considered as part of a reference epigenome and provide a critical balance between large and small projects contributing to a collective goal. Studies funded by this program identified at least 70 additional novel histone modifications, including lysine crotonylation (Kcr), which was shown to be evolutionarily conserved (*18*). In addition, lysine β-hydroxybutyrylation (Kbhb) was discovered, which was markedly induced in response to elevated

β-hydroxybutyrate levels in cultured cells and in livers from mice subjected to prolonged fasting or streptozotocin-induced diabetic ketoacidosis (*19*). The extent to which the novel marks identified are associated with human maladies is a current area of further exploration.

### Epigenomic technology development

The technology development projects supported by this program were focused on enabling novel or markedly improved epigenomic monitoring or manipulation. Among the innovative epigenomic monitoring technologies developed were the MethylC-seq assay, chromatin affinity purification with mass spectrometry, and the development of ligands for in vivo imaging of epigenetic enzymes in humans (*20–22*). Methods for locus- or mark-specific manipulation of the epigenome include a light-inducible CRISPR-Cas9 system for control of endogenous gene activation (*23*), a CRISPR-Cas9 acetyltransferase approach to perform epigenome editing (*24*), and epigenome editing to silence distal regulatory elements (*25*). These studies are important contributions to the epigenome editing tool box, which will aid in modern genome- and epigenome-based therapies and which, in part, helped spawn a new Common Fund program on Somatic Cell Genome Editing (https://commonfund.nih.gov/editing).

### MEASURING SUCCESS

It can be extremely difficult to measure the ultimate success of a scientific program because it is difficult to disentangle the effects of the program versus other nonprogrammatic effects. In addition, scientific advances often take many years to achieve their ultimate impact on human health and disease. In our evaluation of the Roadmap Epigenomics Program, we focused on the following quantitative measures of success: (i) use of data, publications, and resources; (ii) new discoveries of epigenomic mechanisms in human disease; and (iii) new high-value technologies. On the basis of a bibliometric analysis, publications from the Roadmap Epigenomics Program

were found to be highly influential. A total of 857 articles published between 2008 and 30 November 2017 were identified using an internal NIH database linking grants to publications. All 857 articles were analyzed for their influence using iCite (26), which provides a Relative Citation Ratio (RCR) for each publication. Of these publications, 699 were categorized as research articles and received 44,245 citations. The mean RCR value of the Epigenomics Program's research publications is 3.60 and the median RCR is 1.44 compared to an NIH-wide benchmark median RCR value of 1.0. This indicates that publications from the Epigenomics Program tended to be more influential (Fig. 1). Both the number of "Epigenomics" publications in PubMed as well as the number of NIH-funded grants increased significantly since the program started in 2007 (Fig. 2, A and B). However, it is unclear whether this is a direct consequence of the start of the program or whether this increase would have happened independently. Likewise, a substantial increase in epigenomics research in specific disease areas, in addition to cancer, was also evident since 2007 (Fig. 2C), which is an indication of how this program catalyzed new discoveries related to epigenomic mechanisms of human diseases. Last, we analyzed patent applications as an indicator of commercialization potential. As of December 2017, there were 27 granted patents out of 55 patent applications from the Epigenomics Program.

## KEY CHALLENGES AND LESSONS LEARNED FROM THE EPIGENOMICS ROADMAP CONSORTIUM

When the NIH Roadmap Epigenomics Program was being developed and implemented, there were several issues that needed to be addressed: (i) the overall structure of the program (e.g., the balance between smaller discovery and technology development projects versus larger goal-directed mapping efforts), (ii) how best to coordinate the activities of the consortium as a whole and address the challenges of working in a group, and (iii) how to establish and best achieve the goals of the consortium. A "consortium approach" can bring to mind large projects with NIH-mandated goals that offer little or no opportunity for investigator-initiated innovation. Concerns about this type of approach were voiced at the outset of the Epigenomics Program (27). However, the design of the Epigenomics Program included both large data-generating projects as well as smaller innovation- and discovery-focused projects. Common Fund programs are intended to change paradigms, develop innovative tools and technologies, and/or provide fundamental foundations for research that can be used by the global biomedical research community. These programs often involve a combination of highly coordinated research to develop large datasets or tools, innovative technology development to further enhance the capabilities of the consortium, and demonstration projects to road test the value of the data/tools. As stated earlier, the combination of large-scale data-generating projects and smaller, nimble technology-driven methods for which the primary goals are innovation and discovery allows the work of the consortium as a whole to evolve while maintaining an overall focus on delivering large amounts of high-quality data to end users. However, the value of coordinated efforts to deliver community resources that could not be provided by individual groups versus a need for more focused, hypothesis-testing investigator-initiated approaches was also debated (27). At the root of the debate was whether the field was ready for a large mapping effort; mapping epigenetic modifications was more daunting than sequencing the
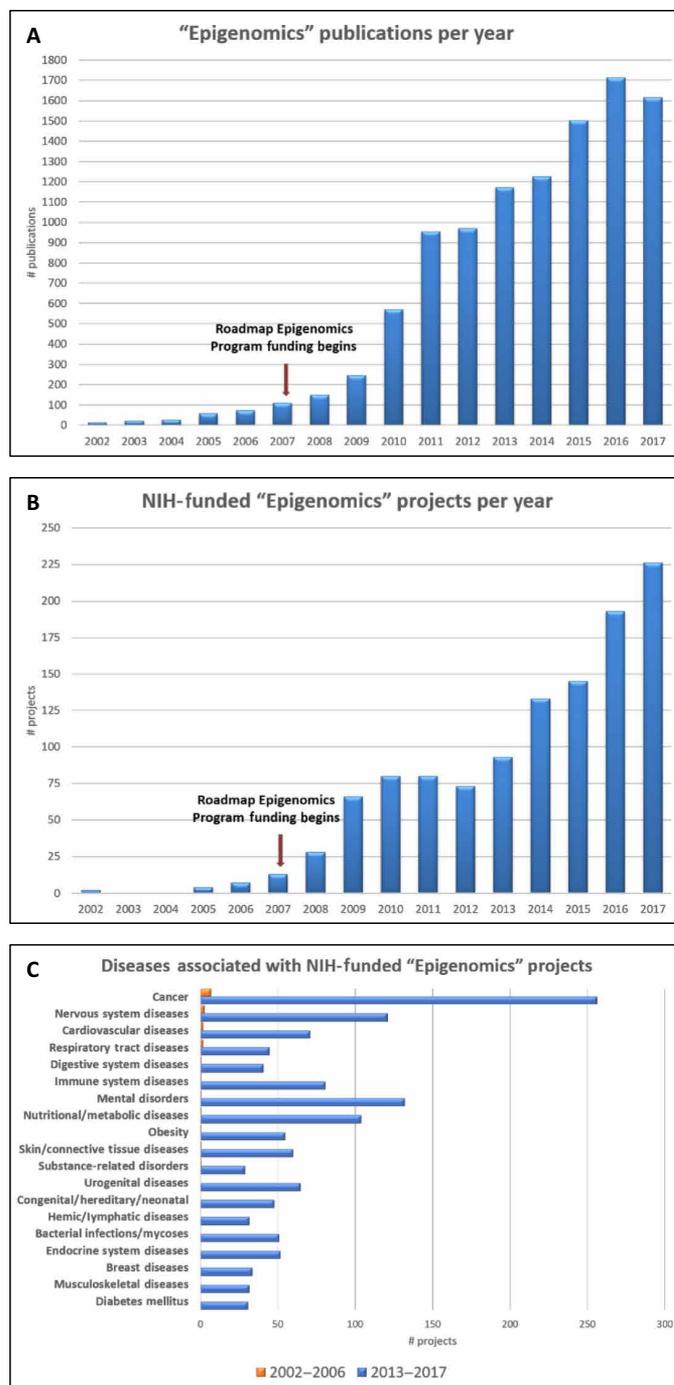
**Fig. 2. Influence of the Roadmap Epigenomics Program on the field of epigenomics research.** (**A**) Epigenomics publications per year identified by searching PubMed for "epigenom*." There were 10,439 epigenomics publications identified. (**B**) NIH-funded Epigenomics projects per year identified by searching an NIH grant database for new awards with epigenom* in the title or specific aims. Excludes subprojects and intramural projects. (**C**) Number of projects per year for top 10 conditions of NIH-funded Epigenomics projects. Projects were identified by searching an NIH grant database for new awards with epigenom* in the title or specific aims. Excludes subprojects and NIH intramural research projects.

human genome because there was no clear definition of a single reference epigenome. Nevertheless, the program was launched in 2008 with the expectation that coordinated mapping efforts would ultimately lead to the creation of reference epigenomes and that these would be "durable goods" for the benefit of the community in the long run. International data standards, benchmarking across laboratories as methods were developed, development of computational tools, and data management that addressed the needs of different types of users were worthwhile goals that would not have been achieved without a coordinated effort. Last, community outreach and engagement via presentations at meetings and workshops was strongly emphasized by external program consultants who represented the interests of the broad scientific community, because community resources are only valuable if the community knows about them.

A second major issue was how best to coordinate the activities of the consortium as a whole and address the challenges of working in a group. As rewarding as consortium science can be, the formation of a consortium also presents several challenges. Investigators are asked to adopt a community-centric focus rather than concentrating on their own individual research interests. This represents a shift in thinking as well as decision-making for themselves and their laboratories. Consistent with this perspective, instead of maximizing their own benefit, the Roadmap Epigenomics Consortium adopted a data- and resource-sharing policy that emphasized early sharing of tools and methods. For instance, MethylC-seq for whole-genome bisulfite sequencing of mammalian cells (20) was readily adopted by the consortium after the steering committee and NIH staff decided that all mapping centers should use it for their reference epigenomes. While this initially caused some feasibility concerns, rapid sharing of data and tools facilitated this shift without diminishing the recognition of the data/tool developers.

A third important issue was how best to establish and achieve the goals of the consortium. Newly funded members of large consortia may be uncomfortable with the goals established by the NIH. Moving from the high-level objectives established via funding opportunities to the articulation of specific consortium goals, development of plans for benchmarking between laboratories, and establishment of integrated analysis plans requires iterative discussion so that consensus can be established. The Epigenomics Consortium was governed by a steering committee, consisting of consortium members and NIH staff. Substantial input was also received from the external program consultants regarding consortium goals and expectations. Although it took extensive time and effort, the investigators ultimately collectively coordinated among themselves by focusing on high-achieving goals leading to prominent publications that everyone recognized could only be accomplished through a considerable consortium effort that was greater than the sum of which all the laboratories could complete individually.

## MOVING FORWARD

In addition to the obvious synergistic relation between Roadmap Epigenomics, IHEC, and ENCODE, several other Common Fund programs have benefited from the lessons learned by the Epigenomics Program. In particular, having established a successful blueprint for mapping efforts, programs like 4D Nucleome (4DN) also ventured into this space in attempting to create, among other things, three-dimensional (3D) maps of genomic interactions as a long-term community tool. Although creating 3D maps of genomic interactions is a far more complex task, direct comparison of Epigenomics and ENCODE datasets with 3D maps generated by the 4DN program and ENCODE would ultimately provide the scientific community with a powerful framework to integrate structural and functional genomic, epigenomic, and transcription factor binding datasets to better understand human health and disease. Likewise, the newly minted Common Fund Human Biomolecular Atlas Program, focusing on the ambitious goal of providing the map of the human body at cellular resolution, may also benefit from the Epigenomics Program data. Our understanding of the human genome, sequenced in the early 1990s, has been further untangled by epigenomic and nucleomic research and may culminate in an unprecedented view of the human body at subcellular resolution. Scientific progress will continue to require the judicious use of large goal-driven programs that produce technologies and tools that can accelerate discovery from individual laboratory science. The Common Fund is committed to continued evaluation and refining of best practices for design and management of consortia to maximize their scientific and health impact.

## Highlights

1) Critical analysis of current funding reveals gaps where Common Fund support can have the greatest impact;
2) Scientific gaps often occur around issues that cannot be addressed by a single researcher and that require a multidisciplinary and coordinated effort to achieve;
3) Development of community resources requires consistent input from scientists who represent the user community;
4) Community resources need to be developed in the context of the international community—this requires dedicated time and effort;
5) Principal investigators funded to generate a community resource need to commit to consortium goals and community outreach;
6) Early sharing of large datasets and tools does not impede the work of data/tool generators—it enhances the impact of the data/tools
7) The combination of consortium-driven community resource development in addition to discovery and technology development projects can result in rapid advances for the field as a whole.

## REFERENCES AND NOTES

1. A. Bird, Perceptions of epigenetics. *Nature* **447**, 396–398 (2007).
2. T. Kouzarides, Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).
3. S. Beck, A. Olek, J. Walter, From genomics to epigenomics: A loftier view of life. *Nat. Biotechnol.* **17**, 1144 (1999).
4. B. E. Bernstein, J. A. Stamatoyannopoulos, J. F. Costello, B. Ren, A. Milosavljevic, A. Meissner, M. Kellis, M. A. Marra, A. L. Beaudet, J. R. Ecker, P. J. Farnham, M. Hirst, E. S. Lander, T. S. Mikkelsen, J. A. Thomson, The NIH Roadmap Epigenomics Mapping Consortium. *Nat. Biotechnol.* **28**, 1045–1048 (2010).
5. J. S. Satterlee, A. Beckel-Mitchener, K. McAllister, D. C. Procaccini, J. L. Rutter, F. L. Tyson, L. H. Chadwick, Community resources and technologies developed through the NIH Roadmap Epigenomics Program. *Methods Mol. Biol.* **1238**, 27–49 (2015).
6. C. E. Romanoski, C. K. Glass, H. G. Stunnenberg, L. Wilson, G. Almouzni, Roadmap for regulation. *Nature* **518**, 314–316 (2015).
7. J. S. Satterlee, D. Schübeler, H. H. Ng, Tackling the epigenome: Challenges and opportunities for collaboration. *Nat. Biotechnol.* **28**, 1039–1044 (2010).
8. M. T. Maurano, R. Humbert, E. Rynes, R. E. Thurman, E. Haugen, H. Wang, A. P. Reynolds, R. Sandstrom, H. Qu, J. Brody, A. Shafer, F. Neri, K. Lee, T. Kutyavin, S. Stehling-Sun, A. K. Johnson, T. K. Canfield, E. Giste, M. Diegel, D. Bates, R. S. Hansen, S. Neph, P. J. Sabo, S. Heimfeld, A. Raubitschek, S. Ziegler, C. Cotsapas, N. Sotoodehnia, I. Glass, S. R. Sunyaev, R. Kaul, J. A. Stamatoyannopoulos, Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).
9. E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov,

C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, Y. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blöcker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen, J. Szustakowki; International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

10. Roadmap Epigenomics Consortium, A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, V. Amin, J. W. Whitaker, M. D. Schultz, L. D. Ward, A. Sarkar, G. Quon, R. S. Sandstrom, M. L. Eaton, Y. C. Wu, A. R. Pfenning, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R. A. Harris, N. Shoresh, C. B. Epstein, E. Gjoneska, D. Leung, W. Xie, R. D. Hawkins, R. Lister, C. Hong, P. Gascard, A. J. Mungall, R. Moore, E. Chuah, A. Tam, T. K. Canfield, R. S. Hansen, R. Kaul, P. J. Sabo, M. S. Bansal, A. Carles, J. R. Dixon, K. H. Farh, S. Feizi, R. Karlic, A. R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. Elliott, T. R. Mercer, S. J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R. C. Sallari, K. T. Siebenthall, N. A. Sinnott-Armstrong, M. Stevens, R. E. Thurman, J. Wu, B. Zhang, X. Zhou, A. E. Beaudet, L. A. Boyer, P. L. De Jager, P. J. Farnham, S. J. Fisher, D. Haussler, S. J. Jones, W. Li, M. A. Marra, M. T. McManus, S. Sunyaev, J. A. Thomson, T. D. Tlsty, L. H. Tsai, W. Wang, R. A. Waterland, M. Q. Zhang, L. H. Chadwick, B. E. Bernstein, J. F. Costello, J. R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J. A. Stamatoyannopoulos, T. Wang, M. Kellis, Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).

11. Y. He, T. Wang, EpiCompare: An online tool to define and explore genomic regions with tissue or cell type–specific epigenomic features. *Bioinformatics* **33**, 3268–3275 (2017).

12. R. F. Pérez, J. R. Tejedor, G. F. Bayón, A. F. Fernández, M. F. Fraga, Distinct chromatin signatures of DNA hypomethylation in aging and cancer. *Aging Cell* **17**, e12744 (2018).

13. H. G. Stunnenberg; International Human Epigenome Consortium, M. Hirst, The International Human Epigenome Consortium: A blueprint for scientific collaboration and discovery. *Cell* **167**, 1145–1149 (2016).

14. P. L. De Jager, G. Srivastava, K. Lunnon, J. Burgess, L. C. Schalkwyk, L. Yu, M. L. Eaton, B. T. Keenan, J. Ernst, C. McCabe, A. Tang, T. Raj, J. Replogle, W. Brodeur, S. Gabriel, H. S. Chai, C. Younkin, S. G. Younkin, F. Zou, M. Szyf, C. B. Epstein, J. A. Schneider, B. E. Bernstein, A. Meissner, N. Ertekin-Taner, L. B. Chibnik, M. Kellis, J. Mill, D. A. Bennett, Alzheimer's disease: Early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat. Neurosci.* **17**, 1156–1163 (2014).

15. L. M. Reynolds, J. R. Taylor, J. Ding, K. Lohman, C. Johnson, D. Siscovick, G. Burke, W. Post, S. Shea, D. R. Jacobs Jr., H. Stunnenberg, S. B. Kritchevsky, I. Hoeschele, C. E. McCall,

D. Herrington, R. P. Tracy, Y. Liu, Age-related variations in the methylome associated with gene expression in human monocytes and T cells. *Nat. Commun.* **5**, 5366 (2014).

16. I. V. Yang, B. S. Pedersen, A. Liu, G. T. O'Connor, S. J. Teach, M. Kattan, R. T. Misiak, R. Gruchalla, S. F. Steinbach, S. J. Szefler, M. A. Gill, A. Calatroni, G. David, C. E. Hennessy, E. J. Davidson, W. Zhang, P. Gergen, A. Togias, W. W. Busse, D. A. Schwartz, DNA methylation and childhood asthma in the inner city. *J. Allergy Clin. Immunol.* **136**, 69–80 (2015).

17. R. D. Hawkins, G. C. Hon, L. K. Lee, Q. Ngo, R. Lister, M. Pelizzola, L. E. Edsall, S. Kuan, Y. Luu, S. Klugman, J. Antosiewicz-Bourget, Z. Ye, C. Espinoza, S. Agarwahl, L. Shen, V. Ruotti, W. Wang, R. Stewart, J. A. Thomson, J. R. Esker, B. Ren, Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6**, 479–491 (2010).

18. M. Tan, H. Luo, S. Lee, F. Jin, J. S. Yang, E. Montellier, T. Buchou, Z. Cheng, S. Rousseaux, N. Rajagopal, Z. Lu, Z. Ye, Q. Zhu, J. Wysocka, Y. Ye, S. Khochbin, B. Ren, Y. Zhao, Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**, 1016–1028 (2011).

19. Z. Xie, D. Zhang, D. Chung, Z. Tang, H. Huang, L. Dai, S. Qi, J. Li, G. Colak, Y. Chen, C. Xia, C. Peng, H. Ruan, M. Kirkey, D. Wang, L. M. Jensen, O. K. Kwon, S. Lee, S. D. Pletcher, M. Tan, D. B. Lombard, K. P. White, H. Zhao, J. Li, R. G. Roeder, X. Yang, Y. Zhao, Metabolic regulation of gene expression by histone lysine β-hydroxybutyrylation. *Mol. Cell* **62**, 194–206 (2016).

20. R. Lister, M. Pelizzola, R. H. Dowen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q. M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren, J. R. Ecker, Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).

21. S. D. Byrum, A. Raman, S. D. Taverna, A. J. Tackett, ChAP-MS: A method for identification of proteins and histone posttranslational modifications at a single genomic locus. *Cell Rep.* **2**, 198–205 (2012).

22. H. Y. Wey, T. M. Gilbert, N. R. Zürcher, A. She, A. Bhanot, B. D. Taillon, F. A. Schroeder, C. Wang, S. J. Haggarty, J. M. Hooker, Insights into neuroepigenetics through human histone deacetylase PET imaging. *Sci. Transl. Med.* **8**, 351ra106 (2016).

23. H. Yumerefendi, A. M. Lerner, S. P. Zimmerman, K. Hahn, J. E. Bear, B. D. Strahl, B. Kuhlman, Light-induced nuclear export reveals rapid dynamics of epigenetic modifications. *Nat. Chem. Biol.* **12**, 399–401 (2016).

24. T. S. Klann, J. B. Black, M. Chellappan, A. Safi, L. Song, I. B. Hilton, G. E. Crawford, T. E. Reddy, C. A. Gersbach, CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nat. Biotechnol.* **35**, 561–568 (2017).

25. P. I. Thakore, A. M. D'Ippolito, L. Song, A. Safi, N. K. Shivakumar, A. M. Kabadi, T. E. Reddy, G. E. Crawford, C. A. Gersbach, Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* **12**, 1143–1149 (2015).

26. B. I. Hutchins, X. Yuan, J. M. Anderson, G. M. Santangelo, Relative Citation Ratio (RCR): A new metric that uses citation rates to measure influence at the article level. *PLOS Biol.* **14**, e1002541 (2016).

27. E. Pennisi, Research funding. Are epigeneticists ready for big science? *Science* **319**, 1177 (2008).

# ScienceAdvances

## The NIH Common Fund/Roadmap Epigenomics Program: Successes of a comprehensive consortium

John S. Satterlee, Lisa H. Chadwick, Frederick L. Tyson, Kim McAllister, Jill Beaver, Linda Birnbaum, Nora D. Volkow, Elizabeth L. Wilder, James M. Anderson and Ananda L. Roy

| | |
|---|---|
| **ARTICLE TOOLS** | http://advances.sciencemag.org/content/5/7/eaaw6507 |
| **REFERENCES** | This article cites 27 articles, 3 of which you can access for free http://advances.sciencemag.org/content/5/7/eaaw6507#BIBL |
| **PERMISSIONS** | http://www.sciencemag.org/help/reprints-and-permissions |