

EDITORIAL

What can epigenomics do for you?

Alexander Meissner^{1,2}

In 1990, with the initial launch of a 15-year project to map and sequence the human genome, a new era of science began. However, even after its successful and early completion in 2001 [1], no one could have foreseen how, only a few years later, genome sequencing would explode to become a widely applied multi-purpose tool whose applications include the mapping of epigenetic modifications and the complete assessment of both coding and non-coding RNA transcripts. The game changer behind this explosion was the transition from the classic electrophoretic Sanger sequencing method, which had limited scalability, to image-based massively parallel 'sequencing-by-synthesis' platforms.

It had already become clear in the early days of the post-genome era, before these technological breakthroughs, that there were additional layers to the primary sequence waiting to be uncovered, and a small number of pilot epigenome projects, including the Human Epigenome Consortium (HEC), were launched [2, 3]. While on the right track, these early projects suffered from lacking the sequencing capacity required to tackle the multidimensional space of the epigenome. This obstacle was overcome in 2006, with the introduction of next generation sequencing platforms, and the NIH was commendably fast to capitalize on these developments by implementing both its ENCODE and 'Roadmap Epigenomics' projects. ENCODE aimed to utilize the newly generated epigenome maps to assist in discovering and assigning functional elements in the genome, while the Roadmap Epigenomics Program aimed to create reference maps for the majority of normal, primary cell types [4, 5]. The success of these projects has helped to popularize epigenomics and has proved somewhat contagious, with additional consortia, such as the recently funded BLUEPRINT (European) and DEEP (German), arriving on the scene; the International Human Epigenome Consortium (IHEC) now coordinates international efforts.

The core technologies used in these projects, and in general across the field, have stabilized over the years and standards are now largely agreed upon. Chromatin immunoprecipitation followed by sequencing (ChIP-seq) [6, 7] remains the standard assay for determining transcription factor binding, as well as for mapping the genome-wide distribution of histone modifications. Continued efforts to increase sensitivity and resolution has resulted in some recent technical improvements to the basic ChIP-seq method, in the form of nano-ChIP-seq [8] and ChIP-exo [9], respectively. By contrast, dozens of assays exist for DNA methylation [10], although most genome-wide studies are focused on just a few of these [11-13]. As costs continue to decrease, methods are converging on whole genome bisulfite sequencing [14], which had previously been prohibitively expensive. As with exome sequencing, the subject of *Genome Biology's* 2011 special issue [15], the driving force behind the ongoing explosion in epigenome studies, and data, has been an increase in sequencing capacity at reduced cost.

What can epigenomics do for you?

Epigenome data are very powerful and have multiple applications that extend beyond a simple map of a particular mark or modification in a given cell type. Below, I will highlight a few selected examples of these applications, although this is a far from exhaustive list.

Genome Annotation

Mammalian genomes are large and complex. Understanding such genomes is not trivial and comparative genomics based on the primary DNA sequence alone, while powerful [16], cannot provide all the answers. As demonstrated several years ago [7], and further highlighted by the many recent ENCODE publications [17], chromatin signatures enable efficient and precise genome annotation of regulatory elements, and can pinpoint functional or cell type-specific regions of interest.

Cell identity

It has become abundantly clear over the past years that epigenomic maps provide more information than can be gained from gene expression data alone [6, 7, 18-20].

Correspondence: alexander_meissner@harvard.edu

¹Department of Stem Cell and Regenerative Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, USA

²Broad Institute, 7 Cambridge Center, Cambridge, MA 02142, USA

While genes are either expressed or not, chromatin states can add further refinement to a gene's activity status, such as whether it is primed or poised, and can also describe varying degrees of repressed states that would all look the same by any gene expression measure. The precise chromatin state of these loci can have clear consequences for how they behave in both normal development and disease.

Disease

As highlighted by many studies of human disease, including several in this issue [21-27], epigenomic maps can be utilized to trace the origin of cells, dissect affected pathways and identify predictive biomarkers. Epigenome data have also proved to be powerful in helping to pinpoint disease-relevant regulatory elements through epigenome-wide association studies, or 'EWAS', especially when integrated with data from genome-wide association studies [17, 28].

Challenges

Several of the reports in this special issue expand the catalog of user-friendly tools for the visualization of epigenomic datasets [27, 29-31], and much work has previously been done elsewhere in this area, including the development of advanced epigenome browsers [32, 33]. Making data even more accessible will be critical if the field is to continue its rapid growth and strengthen its impact. As the number of epigenomic datasets grows into the thousands and tens of thousands, of key importance will be plans made to ensure that sufficient standards are met and that data can be navigated in well curated, high quality databases. An additional challenge in data integration is that much of the full complexity of the epigenome lies in uncharted waters, with many known modifications remaining unmapped and other modifications, such as hydroxymethylation [34], moving to the center stage. This dynamism in the types of data being produced will be sure to generate increasing demands for new, refined, bioinformatic tools.

Conclusions

The overall impact of the growing number of epigenomes, including the NIH Roadmap Epigenomics Project reference maps, will likely be underestimated. For example, almost every study uses the reference genome sequence, whether it is to design primers, target constructs or align sequencing reads, yet those studies rarely acknowledge the reference genome, because it is simply there and so you can just use it.

I predict that, as investigators become more accustomed to epigenome browsers and to utilizing the existing data for various purposes, the reference set of epigenomics maps will also become a routine resource

used in many studies. Applications would include providing a quick overview of a gene/locus of interest; helping to refine a hypothesis; assisting primer design by narrowing down the exact region of dynamic regulation; forming the bases of reporter assays by selecting with precision the functional elements of an upstream regions; and so forth.

This special issue covers epigenomics over a wide range of organisms, systems and methods, all of which provide an informative sampling to illuminate the possibilities for future studies in this expanding and exciting field.

Competing interests

The author declares that he has no competing interests.

Published: 23 October 2012

References

1. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P: **Initial sequencing and comparative analysis of the mouse genome.** *Nature* 2002, **420**:520-562.
2. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, Haefliger C, Horton R, Howe K, Jackson DK, Kunde J, Koenig C, Liddle J, Niblett D, Otto T, Pettett R, Seemann S, Thompson C, West T, Rogers J, Olek A, Berlin K, Beck S: **DNA methylation profiling of human chromosomes 6, 20 and 22.** *Nat Genet* 2006, **38**:1378-1385.
3. Rakyan VK, Hildmann T, Novik KL, Lewin J, Tost J, Cox AV, Andrews TD, Howe KL, Otto T, Olek A, Fischer J, Gut IG, Berlin K, Beck S: **DNA Methylation Profiling of the Human Major Histocompatibility Complex: A Pilot Study for the Human Epigenome Project.** *PLoS Biol* 2004, **2**:e405.
4. ENCODE Project Consortium. **The ENCODE (ENCyclopedia Of DNA Elements) Project.** *Science* 2004, **306**:636-640.
5. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet AL, Ecker JR, Farnham PJ, Hirst M, Lander ES, Mikkelsen TS, Thomson JA: **The NIH Roadmap Epigenomics Mapping Consortium.** *Nat Biotechnol* 2010, **28**:1045-1048.
6. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K: **High-resolution profiling of histone methylations in the human genome.** *Cell* 2007, **129**:823-837.
7. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE: **Genome-wide maps of chromatin state in pluripotent and lineage-committed cells.** *Nature* 2007, **448**:553-560.
8. Adli M, Bernstein BE: **Whole-genome chromatin profiling from limited numbers of cells using nano-ChIP-seq.** *Nat Protoc* 2011, **6**:1656-1668.
9. Rhee HS, Pugh BF: **Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution.** *Cell* 2011, **147**:1408-1419.
10. Beck S, Rakyan VK: **The methylome: approaches for global DNA methylation profiling.** *Trends Genet* 2008.
11. Beck S: **Taking the measure of the methylome.** *Nat Biotechnol* 2010, **28**:1026-1028.
12. Bock C, Tomazou EM, Brinkman AB, Muller F, Simmer F, Gu H, Jager N, Gnirke A, Stunnenberg HG, Meissner A: **Quantitative comparison of genome-wide DNA methylation mapping technologies.** *Nat Biotechnol* 2010.
13. Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, Johnson BE, Fouse SD, Delaney A, Zhao Y, et al: **Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications.** *Nat Biotechnol* 2010.
14. Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, Johnson BE, Fouse SD, Delaney A, Zhao Y, Olshen A, Ballinger T, Zhou X, Forsberg KJ, Gu J, Echipare L, O'Geen H, Lister R, Pelizzola M, Xi Y, Epstein CB, Bernstein BE, Hawkins RD, Ren B, Chung WY, Gu H, Bock C, Gnirke A, Zhang MQ, Haussler D, et al: **Human DNA methylomes at base resolution show widespread epigenomic differences.** *Nature* 2009.

15. *Genome Biology*: special issue on exome sequencing. [<http://genomebiology.com/content/12/9>]
16. Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, Washietl S, Kheradpour P, Ernst J, Jordan G, Mauceli E, Ward LD, Lowe CB, Holloway AK, Clamp M, Gnerre S, Alfoldi J, Beal K, Chang J, Clawson H, Cuff J, Di Palma F, Fitzgerald S, Flicek P, Guttman M, Hubisz MJ, Jaffe DB, Jungreis I, Kent WJ, Kostka D, Lara M, et al.: **A high-resolution map of human evolutionary constraint using 29 mammals.** *Nature* 2011, **478**:476-482.
17. Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M: **An integrated encyclopedia of DNA elements in the human genome.** *Nature* 2012, **489**:57-74.
18. Koche RP, Smith ZD, Adli M, Gu H, Ku M, Gnirke A, Bernstein BE, Meissner A: **Reprogramming factor expression initiates widespread targeted chromatin remodeling.** *Cell Stem Cell* 2011, **8**:96-105.
19. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J: **A unique chromatin signature uncovers early developmental enhancers in humans.** *Nature* 2011, **470**:279-283.
20. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA: **A chromatin landmark and transcription initiation at most promoters in human cells.** *Cell* 2007, **130**:77-88.
21. Sproul D, Kitchen RR, Nestor CE, Dixon JM, Sims AH, Harrison DJ, Ramsahoye BH, Meehan RR: **Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns.** *Genome Biol* 2012, **13**:R84.
22. Thomson JP, Lempiäinen H, Hackett JA, Nestor CE, Müller A, Bolognani F, Oakeley EJ, Schübeler D, Terranova R, Reinhardt D, Moggs JG, Meehan RR: **Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome.** *Genome Biol* 2012, **13**:R93.
23. Song CX, He C: **Balance of DNA methylation and demethylation in cancer development.** *Genome Biol* 2012, **13**:173.
24. Decock A, Ongenaert M, Hoebeeck J, De Preter K, Van Peer G, Crieckinge WV, Ladenstein R, Schulte JH, Noguera R, Stallings RL, Van Damme A, Laureys G, Vermeulen J, Van Maerken T, Speleman F, Vandesompele J: **Genome-wide promoter methylation analysis in neuroblastoma identifies prognostic methylation biomarkers.** *Genome Biol* 2012, **13**:R95.
25. Horvath S, Zhang Y, Langfelder P, Kahn RS, Boks MPM, van Eijk K, van den Berg LH, Ophoff RA: **Aging effects on DNA methylation modules in human brain and blood tissue.** *Genome Biol* 2012, **13**:R97.
26. Haider S, Cordeddu L, Robinson E, Movassagh M, Siggens L, Vujic A, Choy MK, Goddard M, Lio P, Foo R: **The landscape of DNA repeat elements in human heart failure.** *Genome Biol* 2012, **13**:R90.
27. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE: **methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles.** *Genome Biol* 2012, **13**:R87.
28. Ernst J, Kheradpour P, Mikkelsen TS, Shores N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, Ku M, Durham T, Kellis M, Bernstein BE: **Mapping and analysis of chromatin state dynamics in nine human cell types.** *Nature* 2011, **473**:43-49.
29. Hansen KD, Langmead B, Irizarry RA: **BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions.** *Genome Biol* 2012, **13**:R83.
30. Halachev K, Bast H, Albrecht F, Lengauer T, Bock C: **EpiExplorer: live exploration and global analysis of large epigenomic datasets.** *Genome Biol* 2012, **13**:R96.
31. Diaz A, Nellore A, Song JS: **CHANCE: comprehensive software for quality control and validation of ChIP-seq data.** *Genome Biol* 2012, **13**:R98.
32. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP: **Integrative genomics viewer.** *Nat Biotechnol* 2011, **29**:24-26.
33. Zhou X, Maricque B, Xie M, Li D, Sundaram V, Martin EA, Koebe BC, Nielsen C, Hirst M, Farnham P, Kuhn RM, Zhu J, Smirnov I, Kent WJ, Haussler D, Madden PA, Costello JF, Wang T: **The Human Epigenome Browser at Washington University.** *Nat Methods* 2011, **8**:989-990.
34. Penn NW, Suwalski R, O'Riley C, Bojanowski K, Yura R: **The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid.** *Biochem J* 1972, **126**:781-790.

doi:10.1186/gb-2012-13-10-420

Cite this article as: Meissner A: What can epigenomics do for you? *Genome Biology* 2012, **13**:420.