



# Editorial: Finding New Epigenomics and Epigenetics Biomarkers for Complex Diseases and Significant Developmental Events With Machine Learning Methods

Lu Xie<sup>1</sup>, Tao Huang<sup>2\*</sup>, Michael Liebman<sup>3</sup> and Yu-Hang Zhang<sup>4</sup>

<sup>1</sup>Shanghai Center for Bioinformation Technology, Shanghai, China, <sup>2</sup>Bio-Med Big Data Center, CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences, Shanghai, China, <sup>3</sup>IPQ Analytics, Philadelphia, PA, United States, <sup>4</sup>Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, United States

**Keywords:** machine learning, epigenomics and epigenetics, biomarkers, complex diseases, developmental events

## Editorial on the Research Topic

### Finding New Epigenomics and Epigenetics Biomarkers for Complex Diseases and Significant Developmental Events with Machine Learning Methods

The epigenome has been regarded as one of the most important regulators for genome on its functioning and downstream regulation, therefore, the maintenance of stable and normal functioning epigenome is quite crucial for living organisms. However, the epigenome changes over time, either naturally or triggered by various endogenous and exogenous factors like environment, disease state and infections. The alteration of the epigenome may be crucial intermediate between environments and biological phenotypes during development and pathogenesis. Therefore, developmental or pathological alterations can also be monitored using epigenome characteristics, which is exactly what epigenetics and epigenomics studies focus on.

Currently, studies on epigenetics and epigenomics mainly focused on the plasticity of epigenome during two major biological processes: development and pathogenesis, both of which are inextricably linked to the environment through the epigenetic modifications. From the research layers, current epigenomic and epigenetic studies can be further divided into multiple layers: 1) direct methylation on DNA molecules; 2) histone protein modification; 3) chromatin structure and 4) related noncoding RNAs. Integration all layers of epigenomics and epigenetics studies, the ultimate research goal in this field is to reveal the specific role of epigenome during the development and pathogenesis of human beings and explain the related biological mechanisms using typical epigenetics/epigenomics biomarkers.

Biologically, epigenetics and epigenomics describe complex interactions between environment and genomics, resulting in diverse modifications on histone and DNA molecules. With the development of detecting techniques (like microarray and Methyl-Seq), an explosive increase occurs in epigenetics and epigenomics data. To handle such massive complex data, machine learning models have been introduced in the analyses on data at this omics-level and contribute to the identification of potential disease/developmental events associated epigenetic biomarkers. However, several restrictions and challenges still remain in current epigenetics and epigenomics studies:

- (1) For most epigenomics studies, patients are hard to recruit (comparing to normal controls), lacking samples with diseases characteristics;

## OPEN ACCESS

### Edited and reviewed by:

Michael E. Symonds,  
University of Nottingham,  
United Kingdom

### \*Correspondence:

Tao Huang  
tohuangtao@126.com

### Specialty section:

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

**Received:** 07 January 2022

**Accepted:** 12 January 2022

**Published:** 27 January 2022

### Citation:

Xie L, Huang T, Liebman M and  
Zhang Y-H (2022) Editorial: Finding  
New Epigenomics and Epigenetics  
Biomarkers for Complex Diseases and  
Significant Developmental Events With  
Machine Learning Methods.  
Front. Genet. 13:850367.  
doi: 10.3389/fgene.2022.850367

- (2) For each epigenome, epigenomic alterations with biological significance is highly imbalanced distributed across the genome, making it hard for us to detect;
- (3) Comparing to the sample number, methylation sites targeted by current probes are too many, forming a matrix with much more variables than samples. Larger datasets and pre-modeling features selection may be potential solutions for current restrictions on epigenetics and epigenomics studies.

In this research topic, we focused on the application of machine learning models on data at epigenomics and epigenetic levels to identify potential biomarkers for complex diseases/developmental events, including cancers. There are 33 studies in this special issue.

We hope this collection will inspire the epigenomics and epigenetics researchers to use machine learning methods for biomarker identification and regulatory mechanism investigation.

## AUTHOR CONTRIBUTIONS

TH wrote the editorial and all authors have approved the editorial.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Xie, Huang, Liebman and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*