Journal of Theoretical and Applied Information Technology
31st January 2023. Vol.101. No 2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN XHAUSTIVE ANALYSIS OF RECENT DEVELOPMENTS IN COMPUTATIONAL EPIGENETICS

[1]SUKUMAR T, [2]BABY DEEPA V

[1]Research Scholar, [2]Research Advisor
Bharathidasan University, Tiruchirapalli, TN, India

[1]sukumarmsg@gmail.com, [2]deepamct@gmail.com

ABSTRACT

Objective: To analyze the available data models, recent advancements, and contemporary methodologies in computational epigenetics. **Methods:** The computational epigenetic modelling of dynamic epigntics was analysed and its efficacy and scope were adjudged. We consulted and compared the available primary graph data sources such as Methy-LogiX, StatEpigen, MethyCancer, PubMeth, MethDB, MethBank, etc. The schema design of Neo4j was analyzed in depth. Specific tools and instruments used in different spectrums of epigenetics were consulted and evaluated. **Findings:** It has been demonstrated through the use of combinatorial and multivariable analysis that there are design-level barriers in the database structures that prevent the transfer of transcriptional regulatory knowledge from organisms to their cell level. The authors predict that as large databases grow rapidly, current bioinformatic techniques will no longer be sufficient and that fundamental research will be required to develop a new paradigm utilizing high level prototyping and cutting edge bioinformatic techniques. A comprehensive analysis would also be required to determine the likelihood of curing diseases in novel ways using a variety of data sources. **Novelty:** There are fundamental design level barriers in database structures, and with rapidly growing databases, current bioinformatic techniques are not sufficient. Fundamental research is required to redesign the schema of epigenetic databases and to develop specialized bioinformatic techniques to make the most of the abundance of epigenetic data.

Keywords: *Computational Epigenetics, Dynamics of Epigenetics, Graph Databases, Epigenetic Databases, Epigenomic Mapping.*

1. INTRODUCTION

The desire to comprehend how chemical changes might modify gene participation and its performance has spurred the investigation of the mechanism that regulates gene flipping in the wake of the sequencing of the human genome [1]. Phenotypic plasticity is integrally connected to developmental features and differences, as well as the onset and course of illness (how much non-genotypic factors influence phenotypic form). Genetic and genomic features have been intensively investigated for cases involving humans and other organisms, and databases like Ensemble, ENCODE, GEO, GWAS, and Gene are utilized to maintain up-to-date information [2]. The compiled dataset is substantial and relies significantly on cutting-edge gene sequencing techniques and broad measurements of gene expression as well, laying the groundwork for the further study of other phenomena such as disease-specific mutations and molecular evolution. Amidst the abundance of evidence, it is now clear that gene factors can not fully account for complex structures responsible for phenotypic variety and inheritance. As a result, during the past few decades, there has been a heightened focus on how the epigenetic regulatory network might be reprogrammed [3].

Nucleosomes and the components of chromatin are made by encircling collections of core histone proteins with double-stranded DNA cells that regulate gene expression. In the vicinity of active genes, the structure of chromatin is known to be altered, notably around regions of the genome that promote and enhance transcription [5]. Consequently, variations at the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

chromatin level (DNA and histone proteins both undergo chemical modifications that cause them) affect gene expression, causing the beginning of transcription to be sped up, slowed down, or suppressed. The genetic code is unaffected by such inherited modifications in the composition of chromatin (that further control transcription either through protein excitation and expression or RNA-encoding genes). Cells control the expression of certain genes by creating nucleosomes, the chromatin building blocks, by entangling double-stranded DNA surrounded by groups of central histone proteins. It is known that the structure of chromatin is different near active genes, particularly in the genomic regions of enhancer and promoter [5]. Therefore, variations in the structure of chromatin (resulting from the modifications of chemical properties of histone proteins and DNA) control gene activity, resulting in the initiation of transcription being accelerated, slowed, or inhibited. These heritable alterations to the structure of chromatin have no effect on the genetic code (which also regulates transcription by means of RNA coding or expression of genes). Consequently, a second-order effect known as epigenesis stretches further than the genome's contents to the method by which its signal is created and put into practice during cell division, development, and proliferation.

Numerous enzymes play a role in the synthetic annealing process and are connected with epigenetic changes and alteration biomarkers that "encapsulate" the genetic sequence. DNA methylation is an example of a chemical alteration, while protein modifications include bonded post-transcriptional aims to change histone, with methylation and acetylation being the most thoroughly studied [36]. The process of incorporation of a DNA strand with a methyl group is called DNA methylation. Typically, histone modifications serve as a permanent "off" switch for the gene, whereas the structure of the chromotin and level of gene expression are greatly influenced by histone alterations [6]. However, H Different variations in histone (encoded by distinct genes) are defined differently in chromatin areas that are "open" as opposed to "closed" or "compact", adding a further layer of complexity. This knowledge is also passed on to future descendants. Recent research has linked diverse micromolecular defects to the risk and progression of numerous ailments and diseases, including overeating, psychiatric disorders and other psychotic disorders, and ageing [7]. In contrast, the changeable nature and rapid dynamism of epigenetic modifications are of significant interest in the precision of interference.

As a result of the need to comprehend epigenetic alterations and their consequences on disorder, many computational techniques and technologies for use in data production, modeling, administration, analysis, and treatments have been developed. In contrast to the Human Genome Project (1990–2003), which sequenced all 20,000 human genes (consisting of approximately three billion base pairs), the Epigenomics Road Map (2008–present) investigates unique patterns of epigenetic changes with the purpose of developing an epigenome map for multiple types of tissue and malignancies [8].

The wealth of data in multiple medical and biological domains, fueled by technological and processing capacity advancements, has made it possible to investigate the patterns of mutations that disrupt the normal genetic transcription factors. The difficulties provided by the "additional layer of defense" have also helped to highlight the significance of multidisciplinary methods combining classical components of computing science, math, and the physical sciences with adjacent subjects of genomics [9]. A crucial characteristic of the new paradigm is the building of models to characterize biochemical mechanisms in parallel to the examination of many data sources and comprehensive integrated analysis.

This article conducts an exhaustive study on computational epigenetics, analyzing recent advancements in computational models, database models, and the available methodologies and techniques for generating quantified epigenetic data.

2. COMPUTERIZED SIMULATION OF DYNAMIC EPIGENETICS

Time scales distinguish epigenetic changes (DNA methylation and histone alterations). Given the inherent variety of combinations of epigenetic and epigenomic modifications and their effects, regulation is based on several distinctive methods and the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

complex relationships between examples of these, which influence the functionality and expression of genes [10]. Consequently, the control of epigenetics relates to developing system behavior with complicated global dynamics. The computational modeling community has taken a keener interest in epigenetic dynamics, frequently in the context of illness onset and development. Recent initiatives attempt to investigate interdependencies and the process through which systems evolve, as opposed to the earlier concentration on single changes.

In an early application for stomach tumors, computer modeling served as a supplement to in vivo and in vitro research in validating hypotheses by shedding light on global methylation dynamics. In addition, the following investigation of the function of abnormal promoter methylation in the modification of transcription patterns indicated the enduring objective for a model of a system, namely the several scaling of impacts from abnormal cell modifications to the onset and development of illness. For epigenetic processes, model phenomenology (which the physical and complexity sciences make extensive use of) was constructed to facilitate the construction of theories supported by scant evidence that could be revised eventually. Such micromodels used in computers employ the family of techniques known as Markov Chain Monte Carlo to simulate epigenetic events' interdependence by arbitrary state sampling [11]. Based on empirical evidence, information about transcription facilitated the transition to new states of the histone on the basis of relevant transition probabilities.

The significant role of data on DNA methylation in genomic constancy and the plasticity of cells, in genomic imprinting, as well as other typical cell processes, has inspired significant modeling, prediction, and analytic efforts [12]. Consequently, research on epigenetic leukaemia treatment describes a dynamic multicompartmental model of the levels of DNA methylation depending on the function of proteins and methyltransferase. The partial differential equation model's solution of the first order revealed the process for local regulation of hypo-methylation of such proteins.

Recent research on the interaction between DNA methylation and histone modifications has revealed that particular combinations have epigenetic effects irrespective of whether the chromatin is compact or open. Unique theories for the dynamics of the epigenome, the histone code, and patterns of histone modification have gained more interest in recent years. Moreover, a suggested model of stochastic mathematics specifies, using nonphenomenological physical parameters, the biological pathways engaged in producing patterns of modifications of histone in a single cell.

The development of a design of a transitional control process of epigenetic attempted to reconcile existing theories with experimental observations with an effort to establish a relationship between the structure of chromotin, DNA methylation and histone modifications. Performance was evaluated in terms of quality of memory and stability.

Epigenetic processes appear to tightly regulate genomic control and gene expression throughout several lifetimes, with dysregulation affecting phenotypic diversity and increasing illness risk. A full, ideal model that adheres to ECREM principles would contain details on all associated systems and continuously track cumulative changes [11].

3. DATA SOURCES AND VARIETIES

While there are more than 50,000 publications on epigenetics in PubMed, a majority of them have debuted since 2013. Instead of intergenerational processes, discussions on data processing have emphasized intra-generational processes. (i.e., modification of phenotype inheritance). Literature and experimental research provide elevated explanations of biochemical procedures and their accompanying entities, but quantitative data, mostly gathered by databases for epigenetics and molecules, is becoming more and more prevalent. These tools include a variety of tiny and specialized expansions of large and long-standing databases containing transcription factors, websites that regulate transcription, and nucleotide sequences for the genomes of humans and, in addition, microarray and gene expression information. The following sources of them:

EMBL-EBI:
https://www.ebi.ac.uk/training/online/course/bioi

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

nformatics-terrified/what-database/relational-databases/primary-and-secondary-databases

NAR:
http://www.oxfordjournals.org/nar/database/c/

NGS:
https://www.nextgenerationsequencing.info/bioinformatics/genetic-databases/general-genetic-databases

The most common nucleotide data sets are obtained by new-generation sequencing techniques, such as RNA, exome, and the entire genome, as well as targeted technologies. There is also substantial microarray and RNA sequencing (including nanopore variations) data on gene expression. Selective studies for epigenetic measurements have centered on patterns and content of DNA methylation, methylated genes, chromatin-associated proteins, and cancer and other disorders as well.

4. GRAPH DATABASES: A NEW STRATEGY

Over the past two decades, epigenetic and epigenomic databases have grown significantly. The Krembil Family Epigenetics Laboratory keeps track of the amount of methylation on human chromosomes 21 and 22, and its focus is on the DNA methylation in twins and the male germ cells, whereas the DNA methylation database Methy-LogiX focuses on Alzheimer's disease with late-onset. Also, many epigenetic databases have been created to help researchers learn more about different kinds of diseases.

Small-scale examples are StatEpigen [14], which was created for examining biological variables and correlations in statistics in the cancer type of colon. MethyCancer [15] and the larger Catalog of Somatic Mutations in Cancer are two larger-scale relational databases (COSMIC) [16]. Other examples include PubMeth [17] and MethInfoText [18], both of which are well-known. MethDB contains information regarding the patterns and content of DNA methylation of several phenotypes, tissues, and species. MethBank [19] concentrates on integrating methylation programming data of the next generation, while MethPrimerDB [20] gathers human DNA methylation investigations while For hundreds of DNA methyltransferase genes, REBASE [21] has GenBank [22] data. Histone [23], which has information about human histone proteins and enzymes that change them, has sequences for histones. ChromDB [24] and CREMO-FAC [25] provide information on chromatin-associated proteins and chromatin-remodeling factors in eukaryotes, whereas CR Cistrome [26] offers CHiP-seq information on the chromatin regulators and histone modification linkages in humans and mice.

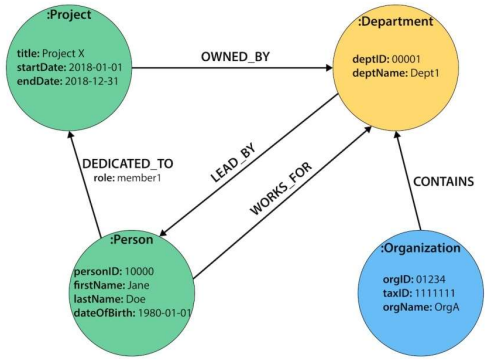

Figure 1: Schema Design of Neo4j

Relational statements relating to the many interdependencies between genetic and epigenetic alterations are required for this complicated system's representation and querying. It has recently been recognized that it is not a primary necessity to maintain organized data; the actual requirements are 1) linking and integrating several data types conforming to distinct data schemas. and 2) analysis of complex hypotheses requires time-consuming and complex query forms. As a result, a unique graph database strategy has been suggested, which allows both integration and query acceleration and has a broad context. Nodes and edges represent concepts and affiliations, respectively, in the graph database, with the structure being adaptable to densely linked data. In comparison to standard relational databases, the use of next-neighbor node-linked traversal searches and graphical search algorithms provides greater suppleness. The Neo4j framework is particularly useful in the biomedical sciences, and it can be utilized for data exploration, data integration, and visualization. There are instances of responsive querying instruments for the assimilation and management of many forms of medical and biological data, as

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

well as Neo4j connects to manage and analyze data as shown in Figure - 1 [27].

Similarly, Neo4j-based frameworks are also used to evaluate the performance of computational and mathematical archetypes of cancer (and the BioModels database).Additionally, FlockDB enables applications rather than traverse searches to reset, offering a foundation for scaled execution comparable to that of Scope, MapReduce and GraphLab. In their paper, Low et al. present a proportional analysis of different deep learning algorithms. In their study, Corbellini et al. provide a summary of graph analyzing platforms for massive networks using graph databases [29].

5. SPECIFIC ANALYSES: PROCEDURES, AND INSTRUMENTS

The rising wealth of resources for various data formats has necessitated the development of related algorithms and tools. Birney et al. and Michels et al. have provided suggestions for the development and assessment of EWAS and also for the analysis of the complex data generated [30]. In epigenome mapping activities, computational modeling has prompted the application of a variety of innovative methodologies in an effort to connect inherent behaviors and to investigate patterns of joint methylation [28]. A comparable regression framework is utilized in modeling genome-wide bisulfite data in RADMeth's [31] discovery of differentially methylated sites. The variance-covariance structure of bisulfite sequence-mapped count data is challenging. However, MACAU [32], an instrument to discover differential DNA methylation that has been known recently, both genetic relatedness and over-dispersion have been introduced. Tools based on Bayesian models, like Bis-SNP [33], find epigenetic events of allele-types, and BS-SNPer [34], a quicker variant. In addition to the multiplicity of technical platforms and necessary sequencing levels, they have inspired contributions to contemporary literature.

Scientists trying to decipher myriad methylation data from numerous specimens require increasingly advanced bio-informatic techniques such as model-based grouping for subsequent analyses, exemplified by Houseman et al. [35], who customized the data generated from microarrays that detect methylation. Multiple variable statistical techniques, especially for unsupervised and supervised grouping, visualization tools such as heat-maps, regression, and principal component analysis, have proven helpful for the understanding of these intricate data sets combining epigenetic changes and molecular processes.

To combat intra-sample cell-type heterogeneity, several statistical techniques have been developed (variations in the sample's component cell type proportions), which is a key hurdle for EWAS. These approaches can be divided into reference-based and reference-free categories.

Examples of text-and data-mining techniques for locating information on epigenetics in scientific publications, in addition to appropriate computational, mathematical, and statistical techniques, are frequently referred to as combined search criteria that show different levels of control. Even for a single change, combined epigenome-wide and summary-level genome analysis can be outlined. Additionally, the epi-informatic approach has changed its focus from Additionally, the epi-informatic approach has changed its focus. Additionally, the epi-informatic approach has changed its focus, as in the case of protein interaction, which determines how methylation and gene expression are related. Various cell types could have erroneous epigenetic markers and enhanced phenotypic plasticity connected to network features and connected to these anomalies, which can offer understanding for diagnoses and treatments. More epi-informatic efforts are being drawn to the use of genetic resources, such as genome-scale libraries, and protein stability regulators are examined more often using posttranscriptional modifications, which facilitates loss-of-function screening, for instance in cancer.

6. CANCER AND DATA INTEGRATION

Cancer has been the subject of numerous studies on epigenetics and epigenomic. However, it is difficult to distinguish between healthy and cancerous states because cancer is not a single disease and does not have a consistent course or set of markers. Given that epigenetic variation underlies normal tissue, establishing consistent objectives in order to detect and treat malignancies depends significantly on the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

biochemical attributes that distinguish various types of cells [37]. The abnormal tumor-suppressor gene activity and mutation cause a significant breakdown in the mechanisms of molecular adhesion and control during the cell cycle, resulting in tumors and nearby tissue [13]. Gene mutations and transcriptional states are just two of the many characteristics that describe cancer phenotypes that operate at the genome level. More epigenetic data is necessary to make these classifications better. The dysregulation of nuclear activities in cells is caused by fundamental interactions between DNA and histone proteins, such as replication, transcription, and DNA damage repair. A more recent assessment of an older model reveals that, according to its creators, some genes experience epigenetically disturbed conditions before developing mutations that result in cancer, causing aberrant differentiation throughout the evolution of the tumor [38].

Histone posttranslational modifications also play an important role in the interactions between DNA-dependent processes and histones. Instead of causing changes in nucleosome structure, this results in the use of specialised binding domains to attract non-histone proteins (Proposition of the Histone Code). These histone modifications result in modifications to chromatin conformation and that transcription is inhibited by the merging of DNA cytosines to methyl residues derived from CpG dinucleotides. Transcription is inhibited by the interaction of DNA cytosines with methyl residues from CpG dinucleotides. Significant focus is also placed on aberrant island locations proximal to promoter regions of hypermethylation of CpG and DNA methylation, with clear evidence that poor regulation of essential gene expression and unique symptoms follow [39].

When it comes to cancer, biological sub-types represent the disease aetiology, the disease as well as the composition of cells. Because different sub-types have varying responses to various therapies, molecular sub-typing is essential. But only recently, the importance of information methylation in the molecular sub-typing of cancer has been recognized. Focused computational methods are currently used to identify the subtypes of neoplastic diseases by identifying groups of genes with differential expression (i.e., biomarkers) that can most effectively distinguish between them.

These strategies, when used on data from other studies, result in distinct sets of biomarkers, which can be problematic [40]. In addition to enhancing and better describing the current subtype signatures using network techniques, combining various -omics data sources could lead to an improvement in the molecular sub-typing of malignant neoplastic diseases. Genome-wide methylation is taken into account in relation to information on subtypes of expression that has resulted from various datasets. The two molecular sub-types were identified with minimal expression differences that could be distinguished. This was confirmed for larger samples with regard to methylation specific to a locus.

7. RESULTS AND DISCUSSION

Ageing has sparked extensive epigenetic research in recent years because it is a significant risk factor for many diseases. The significance of variations in epigenetic stability as a driving force in the development of both health and sickness has been investigated, whereas genome-wide methylation profiles have been used to investigate quantifiable elements of ageing rates in humans [41].

For metabolic illnesses including obesity, diabetes, cardiovascular disease, and others, various investigations have been conducted on the modulation, regulation, and imprinting of epigenetics and concerns of inherence have been raised. In recent years, significant evidence has connected neuropsychiatric illnesses with epigenetic markers as indicators of disease causes and development as well as lifestyle awareness. The authors recently explored the unification of disparate data sets and also evaluated the utility of trans-tissue evaluation, especially when paired with investigation into blood, for assessing the success of long-term treatment programs. The examination of epigenetics in behavioral research and the function of the brain is in its infancy, but interest is expanding quickly, particularly in pediatrics psychology and in general. The influence of childhood experiences on the development of neuronal epigenetics can have impacts that persist into adolescence, and in the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

context of family structure, child care, and parenting techniques, these effects are being researched.

To elucidate the essential epigenetic models and methods that underlie phenotypic diversity, since insects are capable of producing different phenotypic variations through transcriptional reprogramming in cells derived from the same genotype, one proposal proposes using insect models to reflect environmental or lifestyle stressors that alter epigenetic control. The authors contend that this not only implies cost benefit in realizing investigation outcomes, but it also permits examination of the epigenetic trans-generational impacts of environmental variables regarding neuro-degeneration, cancer, infectious illnesses, and ageing.

There are still many unanswered questions regarding inheritance via epigenetics and its place in evolutionary biology. For instance, it has been proposed that epigenetic drifting offers clear evolutionary advantages, and while researching epigenetic regulators and their effects, future research will focus primarily on gene expression and therapeutics.

8. CONCLUSION

The generation of data has dominated numerous discussions of computational epigenetics, and the methods for mining large-scale databases are expanding in number and variety. System complexity suggests that the problems being addressed already require fundamental research; also, it's obvious that higher-level prototype paradigms and cutting-edge bioinformatics techniques will be necessary. Recent research indicates that this major issue is now being resolved. Efforts to conduct integrative analyses with a diversity of data formats and the discovery of epigenetic molecular markers present the chance to cure disease in new directions.

REFERENCES

[1]. Bentley DR. Decoding the human genome sequence. Human Molecular Genetics. 2000 Oct 1;9(16):2353–8. Available from: http://dx.doi.org/10.1093/hmg/9.16.2353

[2]. Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: archive for functional genomics data sets--10 years on. Nucleic Acids Research. 2010 Nov 21;39(Database):D1005–10. Available from: http://dx.doi.org/10.1093/nar/gkq1184

[3]. Lee T, Young R. Transcriptional Regulation and Its Misregulation in Disease. Cell. 2013 Mar;152(6):1237–51. Available from: http://dx.doi.org/10.1016/j.cell.2013.02.014

[4]. Hu Z, Tee WW. Enhancers and chromatin structures: regulatory hubs in gene expression and diseases. Bioscience Reports. 2017 Apr 28;37(2). Available from: http://dx.doi.org/10.1042/bsr20160183

[5]. Basso M, Sleiman S, Ratan RR. Looking Above but Not Beyond the Genome for Therapeutics in Neurology and Psychiatry: Epigenetic Proteins and RNAs Find a New Focus. Neurotherapeutics. 2013 Oct;10(4):551–5. Available from: http://dx.doi.org/10.1007/s13311-013-0225-2

[6]. Ioudinkova ES, Barat A, Pichugin A, Markova E, Sklyar I, Pirozhkova I, et al. Distinct Distribution of Ectopically Expressed Histone Variants H2A.Bbd and MacroH2A in Open and Closed Chromatin Domains. Defossez PA, editor. PLoS ONE. 2012 Oct 30;7(10):e47157. Available from: http://dx.doi.org/10.1371/journal.pone.0047157

[7]. T A. Epigenetic Modulation of Mood Disorders. Journal of Genetic Syndromes & Gene Therapy. 2013;04(01). Available from: http://dx.doi.org/10.4172/2157-7412.1000120

[8]. Romanoski CE, Glass CK, Stunnenberg HG, Wilson L, Almouzni G. Roadmap for regulation. Nature. 2015 Feb 18;518(7539):314–6. Available from: http://dx.doi.org/10.1038/518314a

[9]. Erdel F, Greene EC. Generalized nucleation and looping model for epigenetic memory of histone modifications. Proceedings of the National Academy of Sciences. 2016 Jul 5;113(29). Available from: http://dx.doi.org/10.1073/pnas.1605862113

[10]. Raghavan K, Ruskin HJ, Perrin D, Goasmat F, Burns J. Computational Micromodel for Epigenetic Mechanisms. Navarro A, editor. PLoS ONE. 2010 Nov 30;5(11):e14031. Available from:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

http://dx.doi.org/10.1371/journal.pone.0014031

[11]. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Research. 2011 Feb 15;21(3):381–95. Available from: http://dx.doi.org/10.1038/cr.2011.22

[12]. Ushijima T, Sasako M. Focus on gastric cancer. Cancer Cell. 2004 Feb;5(2):121–5. Available from: http://dx.doi.org/10.1016/s1535-6108(04)00033-9

[13]. Nepali K, Liou JP. Recent developments in epigenetic cancer therapeutics: Clinical advancement and emerging trends. Journal of Biomedical Science. 2021 Dec;28(1):1-58. Available from: https://doi.org/10.1186/s12929-021-00721-x

[14]. Chen J, Liu X, Shen L, Lin Y, Shen B. CMBD: a manually curated cancer metabolic biomarker knowledge database. Database. 2021 Mar 1;2021. Available from: http://dx.doi.org/10.1093/database/baaa094

[15]. He X, Chang S, Zhang J, Zhao Q, Xiang H, Kusonmano K, et al. MethyCancer: the database of human DNA methylation and cancer. Nucleic Acids Research. 2007 Dec 23;36(Database):D836–41. Available from: http://dx.doi.org/10.1093/nar/gkm730

[16]. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Research. 2014 Oct 29;43(D1):D805–11. Available from: http://dx.doi.org/10.1093/nar/gku1075

[17]. Ongenaert M, Van Neste L, De Meyer T, Menschaert G, Bekaert S, Van Criekinge W. PubMeth: a cancer methylation database combining text-mining and expert annotation. Nucleic Acids Research. 2007 Dec 23;36(Database):D842–6. Available from: http://dx.doi.org/10.1093/nar/gkm788

[18]. Fang YC, Lai PT, Dai HJ, Hsu WL. MeInfoText 2.0: gene methylation and cancer relation extraction from biomedical literature. BMC Bioinformatics. 2011 Dec;12(1). Available from: http://dx.doi.org/10.1186/1471-2105-12-471

[19]. Zou D, Sun S, Li R, Liu J, Zhang J, Zhang Z. MethBank: a database integrating next-generation sequencing single-base-resolution DNA methylation programming data. Nucleic Acids Research. 2014 Oct 7;43(D1):D54–8. Available from: http://dx.doi.org/10.1093/nar/gku920

[20]. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002 Nov 1;18(11):1427–31. Available from: http://dx.doi.org/10.1093/bioinformatics/18.11.1427

[21]. Roberts RJ, Vincze T, Posfai J, Macelis D. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Research. 2014 Nov 5;43(D1):D298–9. Available from: http://dx.doi.org/10.1093/nar/gku1046

[22]. Marino-Ramirez L, Levine KM, Morales M, Zhang S, Moreland RT, Baxevanis AD, et al. The Histone Database: an integrated resource for histones and histone fold-containing proteins. Database. 2011 Oct 23;2011(0):bar048–bar048. Available from: http://dx.doi.org/10.1093/database/bar048

[23]. Khare SP, Habib F, Sharma R, Gadewal N, Gupta S, Galande S. HIstome--a relational knowledgebase of human histone proteins and histone modifying enzymes. Nucleic Acids Research. 2011 Dec 2;40(D1):D337–42. Available from: http://dx.doi.org/10.1093/nar/gkr1125

[24]. Gendler K, Paulsen T, Napoli C. ChromDB: The Chromatin Database. Nucleic Acids Research. 2007 Oct 16;36(suppl_1):D298–302. Available from: http://dx.doi.org/10.1093/nar/gkm768

[25]. Shipra A, Chetan K, Rao MRS. CREMOFAC--a database of chromatin remodeling factors. Bioinformatics. 2006 Oct 4;22(23):2940–4. Available from: http://dx.doi.org/10.1093/bioinformatics/btl509

[26]. Wang Q, Huang J, Sun H, Liu J, Wang J, Wang Q, et al. CR Cistrome: a ChIP-Seq database for chromatin regulators and histone modification linkages in human and mouse. Nucleic Acids Research. 2013 Nov 18;42(D1):D450–8. Available from: http://dx.doi.org/10.1093/nar/gkt1151

[27]. Balaur I, Saqi M, Barat A, Lysenko A, Mazein A, Rawlings CJ, et al. EpiGeNet: A Graph Database of Interdependencies Between Genetic and Epigenetic Events in Colorectal Cancer. Journal of Computational Biology. 2017 Oct;24(10):969–80. Available from: http://dx.doi.org/10.1089/cmb.2016.0095

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[28]. Guevara EE, Hopkins WD, Hof PR, Ely JJ, Bradley BJ, Sherwood CC. Comparative analysis reveals distinctive epigenetic features of the human cerebellum. PLoS genetics. 2021 May 6;17(5):e1009506. Available from: https://doi.org/10.1371/journal.pgen.1009506

[29]. Corbellini A, Mateos C, Godoy D, Zunino A, Schiaffino S. An architecture and platform for developing distributed recommendation algorithms on large-scale social networks. Journal of Information Science. 2015 Oct;41(5):686-704. Available from: https://doi.org/10.1177%2F0165551515588669

[30]. Michels KB, Binder AM, Dedeurwaerder S, Epstein CB, Greally JM, Gut I, Houseman EA, Izzi B, Kelsey KT, Meissner A, Milosavljevic A. Recommendations for the design and analysis of epigenome-wide association studies. Nature methods. 2013 Oct;10(10):949-55. Available from: https://doi.org/10.1038/nmeth.2632

[31]. Dolzhenko E, Smith AD. Using beta-binomial regression for high-precision differential methylation analysis in multifactor whole-genome bisulfite sequencing experiments. BMC bioinformatics. 2014 Dec;15(1):1-8. Available from: https://doi.org/10.1186/1471-2105-15-215

[32]. Dolzhenko E, Smith AD. Using beta-binomial regression for high-precision differential methylation analysis in multifactor whole-genome bisulfite sequencing experiments. BMC Bioinformatics. 2014 Jun 24;15(1). Available from: http://dx.doi.org/10.1186/1471-2105-15-215

[33]. Liu Y, Siegmund KD, Laird PW, Berman BP. Bis-SNP: Combined DNA methylation and SNP calling for Bisulfite-seq data. Genome Biology. 2012;13(7):R61. Available from: http://dx.doi.org/10.1186/gb-2012-13-7-r61

[34]. Gao S, Zou D, Mao L, Liu H, Song P, Chen Y, et al. BS-SNPer: SNP calling in bisulfite-seq data. Bioinformatics. 2015 Aug 28;btv507. Available from: http://dx.doi.org/10.1093/bioinformatics/btv507

[35]. Houseman EA, Christensen BC, Yeh RF, Marsit CJ, Karagas MR, Wrensch M, et al. Model-based clustering of DNA methylation array data: a recursive-partitioning algorithm for high-dimensional data arising as a mixture of beta distributions. BMC Bioinformatics. 2008 Sep 9;9(1). Available from: http://dx.doi.org/10.1186/1471-2105-9-365

[36]. Khogeer AA, AboMansour IS, Mohammed DA. The Role of Genetics, Epigenetics, and the Environment in ASD: A Mini Review. Epigenomes. 2022 Jun;6(2):15. Available at: https://doi.org/10.3390/epigenomes6020015

[37]. Teschendorff AE, Jones A, Fiegl H, Sargent A, Zhuang JJ, Kitchener HC, Widschwendter M. Epigenetic variability in cells of normal cytology is associated with the risk of future morphological transformation. Genome medicine. 2012 Mar;4(3):1-4. Available from: https://doi.org/10.1186/gm323

[38]. Dinalankara W, Bravo HC. Gene Expression Signatures Based on Variability can Robustly Predict Tumor Progression and Prognosis. Cancer Informatics. 2015 Jan;14:CIN.S23862. Available from: http://dx.doi.org/10.4137/cin.s23862

[39]. Barat A, Ruskin HJ. Comparative Correlation Structure of Colon Cancer Locus Specific Methylation: Characterisation of Patient Profiles and Potential Markers across 3 Array-Based Datasets. Journal of Cancer. 2015;6(8):795–811. Available from: http://dx.doi.org/10.7150/jca.9883

[40]. Banerji CRS, Miranda-Saavedra D, Severini S, Widschwendter M, Enver T, Zhou JX, et al. Cellular network entropy as the energy potential in Waddington's differentiation landscape. Scientific Reports. 2013 Oct 24;3(1). Available from: http://dx.doi.org/10.1038/srep03039

[41]. Benayoun BA, Pollina EA, Brunet A. Epigenetic regulation of ageing: linking environmental inputs to genomic stability. Nature reviews Molecular cell biology. 2015 Oct;16(10):593-610.Available from: https://doi.org/10.1038/nrm4048