Spring 2017 – Epigenetics and Systems Biology
Lecture Outline (Systems Biology)
Michael K. Skinner – Biol 476/576
Weeks 5, 6 and 7 (February 7, 14 and 21)

Epigenetics (History / Molecular Processes/ Genomics)

- Definitions and History
- Molecular Factors (DNA Methylation, Histone Modification, Chromatin Structure, ncRNA)
- Epigenetics Technology and Genomics

Required Reading

Holliday R. Epigenetics: a historical overview. Epigenetics. 2006 Apr-Jun;1(2):76-80.

Books (Reserve in Library)

Kevin V. Morris (2012) Non-coding RNAs and Epigenetic Regulation of Gene Expression: Drivers of Natural Selection. Caister Academic Press.

Russo, V.E.A., Martienssen, A.and Riggs, A.D. (eds.). 1996. Epigenetic Mechanisms of Gene Regulation. Cold Spring Harbor Press. Cold Spring Harbor.

Allis, C.D., Jenuwein, T. & Reinberg, D., Eds. (2007). Epigenetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Jeanteur, P. EPIGENETICS AND CHROMATIN. Progress in Molecular and Subcellular Biology, 2005, Volume 38, 151-167, DOI: 10.1007/3-540-27310-7_6

Literature

Haussmann IU, Bodi Z, Sanchez-Moran E, et al, (2016) m6A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. Nature. 8;540(7632):301-304.

Engreitz JM, Haines JE, Perez EM, et al (2016) Local regulation of gene expression by lncRNA promoters, transcription and splicing. Nature. 17;539(7629):452-455.

Chen CK, Blanco M, Jackson C, Aznauryan E, Ollikainen N, Surka C, Chow A, Cerase A, McDonel P, Guttman M. (2016) Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. Science. 2016 Oct 28;354(6311):468-472.

Azvolinsky A (2017) The Scientist. 01.2017: 50-52.

Leonardi M, Librado P, Der Sarkissian C, et al (2016) Evolutionary Patterns and Processes: Lessons from Ancient DNA. Syst Biol. 2016 Jul 5. pii: syw059. [Epub ahead of print]

Gao Z, Zhu X, Dou Y. (2015) The miR-302/367 cluster: a comprehensive update on its evolution and functions. Open Biol. 2015 Dec;5(12):150138. doi: 10.1098/rsob.150138.

Fagny M, Patin E, MacIsaac JL, et al. (2015) The epigenomic landscape of African rainforest hunter gatherers and farmers. Nat Commun. 2015 Nov 30;6:10047. doi: 10.1038/ncomms10047.
Li J, Li R, Wang Y, Hu X, et al (2015) Genome-wide DNA methylome variation in two genetically distinct chicken lines using MethylC-seq. BMC Genomics. 2015 Oct 23;16:851. doi: 10.1186/s12864-015-2098-8.
Vogt G. (2015) Stochastic developmental variation, an epigenetic source of phenotypic diversity with far-reaching biological consequences. J Biosci. 2015 Mar;40(1):159-204.
Aravind L, Burroughs AM, Zhang D, Iyer LM. (2014) Protein and DNA modifications: evolutionary imprints of bacterial biochemical diversification and geochemistry on the provenance of eukaryotic epigenetics. Cold Spring Harb Perspect Biol. 2014 Jul 1;6(7):a016063.
Beltrame MH, Rubel MA, Tishkoff SA. (2016) Inferences of African evolutionary history from genomic data. Curr Opin Genet Dev. 41:159-166.
Klein SL, Moody SA. (2016) When Family History Matters: The Importance of Lineage Analyses and Fate Maps for Explaining Animal Development. Curr Top Dev Biol. 2016;117:93-112.
Chen DH, Huang Y, Ruan Y, Shen WH, (2016) The evolutionary landscape of PRC1 core components in green lineage. Planta. 2016 Apr;243(4):825-46.
Labat-Robert J, Robert L. (2015) Longevity and aging. Mechanisms and perspectives. Pathol Biol (Paris). 2015 Dec;63(6):272-6.
Hepworth J, Dean C. (2015) Flowering Locus C's Lessons: Conserved Chromatin Switches Underpinning Developmental Timing and Adaptation. Plant Physiol. 2015 Aug;168(4):1237-45.
Orlando L, Gilbert MT, Willerslev E. (2015) Reconstructing ancient genomes and epigenomes. Nat Rev Genet. 2015 Jul;16(7):395-408.
Rodgers AB, Bale TL. (2015) Germ Cell Origins of Posttraumatic Stress Disorder Risk: The Transgenerational Impact of Parental Stress Experience. Biol Psychiatry. 2015 Sep 1;78(5):307-14.
Jablonka E, Lamb MJ. (2015) The inheritance of acquired epigenetic variations. Int J Epidemiol. 2015 Aug;44(4):1094-103.
Nazmul Islam M, Yadav S, Hakimul Haque M, et al (2016) Optical biosensing strategies for DNA methylation analysis. Biosens Bioelectron. 2016 Oct 19. pii: S0956-5663(16)31052-1.
Meier K, Recillas-Targa F. (2017) New insights on the role of DNA methylation from a global view. Front Biosci (Landmark Ed). 2017 Jan 1;22:644-668.
Niederhuth CE, Schmitz RJ. (2017) Putting DNA methylation in context: from genomes to gene expression in plants. Biochim Biophys Acta. 2017 Jan;1860(1):149-156.
Reis AH, Vargas FR, Lemos B. (2016) Biomarkers of genome instability and cancer epigenetics. Tumour Biol. 2016 Oct;37(10):13029-13038. Epub 2016 Jul 28.
Bunkar N, Pathak N, Lohiya NK, Mishra PK. (2016) Epigenetics: A key paradigm in reproductive health. Clin Exp Reprod Med. 2016 Jun;43(2):59-81.
Karlsson O, Baccarelli AA. (2016) Environmental Health and Long Non-coding RNAs. Curr Environ Health Rep. 2016 Sep;3(3):178-87.
Elhamamsy AR. (2016) DNA methylation dynamics in plants and mammals: overview of regulation and dysregulation. Cell Biochem Funct. 2016 Jul;34(5):289-98.
Peng J, Xia B, Yi C. (2016) Single-base resolution analysis of DNA epigenome via high-throughput sequencing. Sci China Life Sci. 2016 Mar;59(3):219-26.
Kubiak M, Lewandowska MA. (2015) Can chromatin conformation technologies bring light into human molecular pathology? Acta Biochim Pol. 2015;62(3):483-9.

O'Connell TM, Markunas CA. (2016) DNA Methylation and MicroRNA-Based Biomarkers for Risk of Type 2 Diabetes. Curr Diabetes Rev. 2016;12(1):20-9.
Ito S, Kuraoka I. (2015) Epigenetic modifications in DNA could mimic oxidative DNA damage: A double-edged sword. DNA Repair (Amst). 2015 Aug;32:52-7.
Gutierrez C, Desvoyes B, Vergara Z, Otero S, Sequeira-Mendes J. (2016) Links of genome replication, transcriptional silencing and chromatin dynamics. Curr Opin Plant Biol. 2016 Dec;34:92-99.
Li Y, Seto E. (2016) HDACs and HDAC Inhibitors in Cancer Development and Therapy. Cold Spring Harb Perspect Med. 2016 Oct 3;6(10). pii: a026831.
Chen Y, Müller F, Rieu I, Winter P. (2016) Epigenetic events in plant male germ cell heat stress responses. Plant Reprod. 2016 Jun;29(1-2):21-9.
Cao G, Li HB, Yin Z, Flavell RA. (2016) Recent advances in dynamic m6A RNA modification. Open Biol. 2016 Apr;6(4):160003.
Larriba E, del Mazo J. (2016) Role of Non-Coding RNAs in the Transgenerational Epigenetic Transmission of the Effects of Reprotoxicants. Int J Mol Sci. 2016 Mar 25;17(4):452.
Patel DJ. (2106) A Structural Perspective on Readout of Epigenetic Histone and DNA Methylation Marks. Cold Spring Harb Perspect Biol. 2016 Mar 1;8(3):a018754.
Lakhotia SC. (2015) Divergent actions of long noncoding RNAs on X-chromosome remodelling in mammals and Drosophila achieve the same end result: dosage compensation. J Genet. 2015 Dec;94(4):575-84.
Shafik A, Schumann U, Evers M, Sibbritt T, Preiss T. (2016) The emerging epitranscriptomics of long noncoding RNAs. Biochim Biophys Acta. 2016 Jan;1859(1):59-70.
Costa MC, Leitão AL, Enguita FJ. (2016) Noncoding Transcriptional Landscape in Human Aging. Curr Top Microbiol Immunol. 2016;394:177-202.
Rai G, Rai R, Saeidian AH, Rai M. (2016) Microarray to deep sequencing: transcriptome and miRNA profiling to elucidate molecular pathways in systemic lupus erythematosus. Immunol Res. 2016 Feb;64(1):14-24.
Khorkova O, Hsiao J, Wahlestedt C. (2015) Basic biology and therapeutic implications of lncRNA. Adv Drug Deliv Rev. 2015 Jun 29;87:15-24.
D'Urso A, Brickner JH. (2016) Epigenetic transcriptional memory. Curr Genet. 2016 Nov 2. [Epub ahead of print]
Perino M, Veenstra GJ. (2016) Chromatin Control of Developmental Dynamics and Plasticity. Dev Cell. 2016 Sep 26;38(6):610-20.
Perišić O, Schlick T. (2016) Computational strategies to address chromatin structure problems. Phys Biol. 2016 Jun 25;13(3):035006.
García-González E, Escamilla-Del-Arenal M, Arzate-Mejía R, Recillas-Targa F. (2016) Chromatin remodeling effects on enhancer activity. Cell Mol Life Sci. 2016 Aug;73(15):2897-910.
McFadden EJ, Hargrove AE. (2016) Biochemical Methods To Investigate lncRNA and the Influence of lncRNA:Protein Complexes on Chromatin. Biochemistry. 2016 Mar 22;55(11):1615-30.
Friedman N, Rando OJ. (2015) Epigenomics and the structure of the living genome. Genome Res. 2015 Oct;25(10):1482-90.
Pilu R. (2015) Paramutation phenomena in plants. Semin Cell Dev Biol. 2015 Aug;44:2-10.
Li G, Zhu P. (2015) Structure and organization of chromatin fiber in the nucleus. FEBS Lett. 2015 Oct 7;589(20 Pt A):2893-904.
Cuerda-Gil D, Slotkin RK. (2016) Non-canonical RNA-directed DNA methylation. Nat Plants. 2016 Nov 3;2(11):16163.

Wendte JM, Pikaard CS. (2017) The RNAs of RNA-directed DNA methylation. Biochim Biophys Acta. 2017 Jan;1860(1):140-148.
Cao G, Li HB, Yin Z, Flavell RA. (2016) Recent advances in dynamic m6A RNA modification. Open Biol. 2016 Apr;6(4):160003.
Gebert D, Rosenkranz D. (2015) RNA-based regulation of transposon expression. Wiley Interdiscip Rev RNA. 2015 Nov-Dec;6(6):687-708.
Trerotola M, Relli V, Simeone P, Alberti S. (2015) Epigenetic inheritance and the missing heritability. Hum Genomics. 2015 Jul 28;9:17. doi: 10.1186/s40246-015-0041-3.
Soubry A. (2015) Epigenetic inheritance and evolution: A paternal perspective on dietary influences. Prog Biophys Mol Biol. 2015 Jul;118(1-2):79-85.
Fagnocchi L, Mazzoleni S, Zippo A. (2015) Integration of Signaling Pathways with the Epigenetic Machinery in the Maintenance of Stem Cells. Stem Cells Int. 2016;2016:8652748.
Gligorijević V, Malod-Dognin N, Pržulj N. (2016) Integrative methods for analyzing big data in precision medicine. Proteomics. 2016 Mar;16(5):741-58.
Chen L, Ge B, Casale FP, Vasquez L, et al, (2016) Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune Cells. Cell. 2016 Nov 17;167(5):1398-1414.e24.
Durek P, Nordström K, Gasparoni G, Salhab A, et al. (2016) Epigenomic Profiling of Human CD4+ T Cells Supports a Linear Differentiation Model and Highlights Molecular Regulators of Memory Development. Immunity. 2016 Nov 15;45(5):1148-1161.
Mitra S, Samadder A, Das P, Das S, Dasgupta M, Chakrabarti J. (2016) Decrypting ENCODEd epigenetic marks of human tRN-A-RS genes in normal, stem and cancer cell lines. J Biomol Struct Dyn. 2016 Oct 6:1-13. [Epub ahead of print]
Kinkley S, Helmuth J, Polansky JK, et al. (2016) reChIP-seq reveals widespread bivalency of H3K4me3 and H3K27me3 in CD4(+) memory T cells. Nat Commun. 2016 Aug 17;7:12514.
Morozova I, Flegontov P, Mikheyev AS, et al. (2016) Toward high-resolution population genomics using archaeological samples. DNA Res. 2016 Aug;23(4):295-310.
Barsyte-Lovejoy D, Szewczyk MM, Prinos P, Lima-Fernandes E, Ackloo S, Arrowsmith CH. (2016) Chemical Biology Approaches for Characterization of Epigenetic Regulators. Methods Enzymol. 2016;574:79-103.
Nersisyan L. (2016) Integration of Telomere Length Dynamics into Systems Biology Framework: A Review. Gene Regul Syst Bio. 2016 Jun 16;10:35-42.
Chaitankar V, Karakülah G, Ratnapriya R, Giuste FO, Brooks MJ, Swaroop A. (2016) Next generation sequencing technology and genomewide data analysis: Perspectives for retinal research. Prog Retin Eye Res. 2016 Nov;55:1-31.
Suravajhala P, Kogelman LJ, Kadarmideen HN. (2016) Multi-omic data integration and analysis using systems genomics approaches: methods and applications in animal production, health and welfare. Genet Sel Evol. 2016 Apr 29;48(1):38.
Zhao H, Zhang G, Pang L, Lan Y, et al. (2016) 'Traffic light rules': Chromatin states direct miRNA-mediated network motifs running by integrating epigenome and regulatome. Biochim Biophys Acta. 2016 Jul;1860(7):1475-88.
Lowdon RF, Jang HS, Wang T. (2016) Evolution of Epigenetic Regulation in Vertebrate Genomes. Trends Genet. 2016 May;32(5):269-83.
Rajagopal N, Srinivasan S, Kooshesh K, et al. (2016) High-throughput mapping of regulatory DNA. Nat Biotechnol. 2016 Feb;34(2):167-74.
Nadel J, Athanasiadou R, Lemetre C, et al. (2015) RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. Epigenetics Chromatin. 2015 Nov 16;8:46.

Bradburne C, Graham D, Kingston HM, et al. (2015) Overview of 'Omics Technologies for Military Occupational Health Surveillance and Medicine. Mil Med. 2015 Oct;180(10 Suppl):34-48.
Laufer BI Singh SM. (2015) Strategies for precision modulation of gene expression by epigenome editing: an overview. Epigenetics Chromatin. 2015 Sep 17;8:34.
Sharma A. (2015) Systems genomics analysis centered on epigenetic inheritance supports development of a unified theory of biology. J Exp Biol. 2015 Nov;218(Pt 21):3368-73.
Zierer J, Menni C, Kastenmüller G, Spector TD. (2015) Integration of 'omics' data in aging research: from biomarkers to systems biology. Aging Cell. 2015 Dec;14(6):933-44.
Nag A, Vigneau S, Savova V, Zwemer LM, Gimelbrant AA. (2015) Chromatin Signature Identifies Monoallelic Gene Expression Across Mammalian Cell Types. G3 (Bethesda). 2015 Jun 18;5(8):1713-20.
Weber S, Hofmann A, Herms S, Hoffmann P, Doerfler W. (2015) Destabilization of the human epigenome: consequences of foreign DNA insertions. Epigenomics. 2015 Aug;7(5):745-55.
Enríquez P. (2016) CRISPR-Mediated Epigenome Editing. Yale J Biol Med. 2016 Dec 23;89(4):471-486. eCollection 2016.
Dearfield KL, Gollapudi BB, Bemis JC, et al. (2016) Next generation testing strategy for assessment of genomic damage: A conceptual framework and considerations. Environ Mol Mutagen. 2016 Sep 21. doi: 10.1002/em.22045. [Epub ahead of print]
Klosin A, Lehner B. (2016) Mechanisms, timescales and principles of trans-generational epigenetic inheritance in animals. Curr Opin Genet Dev. 2016 Feb;36:41-9.
Suravajhala P, Kogelman LJ, Kadarmideen HN. (2016) Multi-omic data integration and analysis using systems genomics approaches: methods and applications in animal production, health and welfare. Genet Sel Evol. 2016 Apr 29;48(1):38.
Talwar P, Sinha J, Grover S, et al. (2016) Dissecting Complex and Multifactorial Nature of Alzheimer's Disease Pathogenesis: a Clinical, Genomic, and Systems Biology Perspective. Mol Neurobiol. 2016 Sep;53(7):4833-64.
Hochedlinger K, Jaenisch R. (2015) Induced Pluripotency and Epigenetic Reprogramming. Cold Spring Harb Perspect Biol. 2015 Dec 1;7(12). pii: a019448.
Kim K, Lee K, Bang H, Kim JY, Choi JK. (2016) Intersection of genetics and epigenetics in monozygotic twin genomes. Methods. 2016 Jun 1;102:50-6.
Stelzer Y, Jaenisch R. (2015) Monitoring Dynamics of DNA Methylation at Single-Cell Resolution during Development and Disease. Cold Spring Harb Symp Quant Biol. 2015;80:199-206.
Zhen L, Jianhong X. (2015) The application of the high throughput sequencing technology in the transposable elements. Yi Chuan. 2015 Sep;37(9):885-98.
Kitamura A, Miyauchi N, Hamada H, et al. (2015) Epigenetic alterations in sperm associated with male infertility. Congenit Anom (Kyoto). 2015 Aug;55(3):133-44.
Li N, Shen Q, Hua J. (2016) Epigenetic Remodeling in Male Germline Development. Stem Cells Int. 2016;2016:3152173.
Gokhman D, Meshorer E, Carmel L. (2016) Epigenetics: It's Getting Old. Past Meets Future in Paleoepigenetics. Trends Ecol Evol. 2016 Apr;31(4):290-300.
O'Doherty AM, McGettigan PA. (2015) Epigenetic processes in the male germline. Reprod Fertil Dev. 2015 Jun;27(5):725-38.
Van Soom A, Peelman L, Holt WV, Fazeli A. (2014) An introduction to epigenetics as the link between genotype and environment: a personal view. Reprod Domest Anim. 49 Suppl 3:2-10.
Goriaux C, Théron E, Brasset E, Vaury C. (2014) History of the discovery of a master locus producing piRNAs: the flamenco/COM locus in Drosophila melanogaster. Front Genet. 4;5:257.

Lane M, Robker RL, Robertson SA. (2014) Parenting from before conception. Science. 15;345(6198):756-60.
Jonsson B, Jonsson N. (2014) Early environment influences later performance in fishes. J Fish Biol. 85(2):151-88.
Sweatt JD. (2013) The emerging field of neuroepigenetics. Neuron. 2013 Oct 30;80(3):624-32.
Jodar M, Selvaraju S, Sendler E, Diamond MP, Krawetz SA; Reproductive Medicine Network. (2013) The presence, role and clinical use of spermatozoal RNAs. Hum Reprod Update. 19(6):604-24.
Jablonka E. (2013) Epigenetic inheritance and plasticity: The responsive germline. Prog Biophys Mol Biol. 111(2-3):99-107.
Santiago M, Antunes C, Guedes M, Sousa N, Marques CJ. (2014) TET enzymes and DNA hydroxymethylation in neural development and function - How critical are they? Genomics. 2014 Sep 6. pii: S0888-7543(14)00160-8.
Egea RR, Puchalt NG, Escrivá MM, Varghese AC. (2014) OMICS: Current and future perspectives in reproductive medicine and technology. J Hum Reprod Sci. 7(2):73-92
Huang B, Jiang C, Zhang R. (2014) Epigenetics: the language of the cell? Epigenomics. 6(1):73-88.
Davidsson J. (2014) The epigenetic landscape of aneuploidy: constitutional mosaicism leading the way? Epigenomics. 6(1):45-58.
Liu J, Jia G. (2014) Methylation modifications in eukaryotic messenger RNA. J Genet Genomics. 20;41(1):21-33.
Maeso I, Irimia M, Tena JJ, Casares F, Gómez-Skarmeta JL. (2013) Deep conservation of cis-regulatory elements in metazoans. Philos Trans R Soc Lond B Biol Sci. 11;368(1632) :20130020.
Amaral PP, Dinger ME, Mattick JS. (2013) Non-coding RNAs in homeostasis, disease and stress responses: an evolutionary perspective. Brief Funct Genomics.;12(3):254-78.
Cherblanc FL, Davidson RW, Di Fruscia P, Srimongkolpithak N, Fuchter MJ. (2013) Perspectives on natural product epigenetic modulators in chemical biology and medicine. Nat Prod Rep. 30(5):605-24.
Meng Q, Mäkinen VP, Luk H, Yang X. (2013) Systems Biology Approaches and Applications in Obesity, Diabetes, and Cardiovascular Diseases. Curr Cardiovasc Risk Rep. 7(1):73-83.
Ogino S1, Lochhead P, Chan AT, et al. (2013) Molecular pathological epidemiology of epigenetics: emerging integrative science to analyze environment, host, and disease. Mod Pathol. 26(4):465-84.
Xu GL, Walsh CP. (2014) Enzymatic DNA oxidation: mechanisms and biological significance. BMB Rep. 47(11):609-18.
Messerschmidt DM, Knowles BB, Solter D. (2014) DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev. 15;28(8):812-28.
Fojtová M, Fajkus J. (2014) Epigenetic regulation of telomere maintenance. Cytogenet Genome Res. 143(1-3):125-35.
Wu H, Zhang Y. (2014) Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell. 16;156(1-2):45-68.
Kumar S, Kumari R, Sharma V, Sharma V. (2013) Roles, and establishment, maintenance and erasing of the epigenetic cytosine methylation marks in plants. J Genet. 92(3):629-66.
Jurkowska RZ, Jeltsch A. (2013) Genomic imprinting--the struggle of the genders at the molecular level. Angew Chem Int Ed Engl. 16;52(51):13524-36.
Lu R, Wang GG. (2013) Tudor: a versatile family of histone methylation 'readers'. Trends Biochem Sci. 38(11):546-55.

Zhang H, Wang B, Duan CG, Zhu JK. (2013) Chemical probes in plant epigenetics studies. Plant Signal Behav. 8(9). pii: e25364.
Sandoval J, Peiró-Chova L, Pallardó FV, García-Giménez JL. (2013) Epigenetic biomarkers in laboratory diagnostics: emerging approaches and opportunities. Expert Rev Mol Diagn. 13(5):457-71.
Auclair Y, Richard S. (2013) The role of arginine methylation in the DNA damage response. DNA Repair (Amst). 12(7):459-65.
Morris KV, Mattick JS. (2014) The rise of regulatory RNA. Nat Rev Genet. 15(6):423-37.
Beckedorff FC, Amaral MS, Deocesano-Pereira C, Verjovski-Almeida S. (2013) Long non-coding RNAs and their implications in cancer epigenetics. Biosci Rep. 30;33(4). pii: e00061.
Qureshi IA, Mehler MF. (2013) Long non-coding RNAs: novel targets for nervous system disease diagnosis and therapy. Neurotherapeutics. 10(4):632-46.
Lee JT, Bartolomei MS. (2013) X-inactivation, imprinting, and long noncoding RNAs in health and disease. Cell. 14;152(6):1308-23.
Dahlin JL, Chen X, Walters MA, Zhang Z. (2014) Histone-modifying enzymes, histone modifications and histone chaperones in nucleosome assembly: Lessons learned from Rtt109 histone acetyltransferases. Crit Rev Biochem Mol Biol. 3:1-23. [Epub ahead of print]
Roidl D, Hacker C. (2014) Histone methylation during neural development. Cell Tissue Res. 356(3):539-52.
Leung KS, Cheng VW, Mok SW, Tsui SK. (2014) The involvement of DNA methylation and histone modification on the epigenetic regulation of embryonic stem cells and induced pluripotent stem cells. Curr Stem Cell Res Ther. 9(5):388-95.
Bose P, Dai Y, Grant S. (2014) Histone deacetylase inhibitor (HDACI) mechanisms of action: emerging insights. Pharmacol Ther. 143(3):323-36.
Fiorino E, Giudici M, Ferrari A, Mitro N, Caruso D, De Fabiani E, Crestani M. (2014) The sirtuin class of histone deacetylases: regulation and roles in lipid metabolism. IUBMB Life. 66(2):89-99.
Jørgensen S, Schotta G, Sørensen CS. (2013) Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. Nucleic Acids Res. 1;41(5):2797-806.
Gräff J, Tsai LH. (2013) Histone acetylation: molecular mnemonics on the chromatin. Nat Rev Neurosci. 14(2):97-111.
Wang Y, Shang Y. (2013) Epigenetic control of epithelial-to-mesenchymal transition and cancer metastasis. Exp Cell Res. 15;319(2):160-9.
Thomson JP, Moggs JG, Wolf CR, Meehan RR. (2014) Epigenetic profiles as defined signatures of xenobiotic exposure. Mutat Res Genet Toxicol Environ Mutagen. 764-765:3-9.
Ogino S, Lochhead P, Chan AT, et al. (2013) Molecular pathological epidemiology of epigenetics: emerging integrative science to analyze environment, host, and disease. Mod Pathol.;26(4):465-84.
Sekar D, Thirugnanasambantham K, Hairul Islam VI, Saravanan S. (2014) Sequencing approaches in cancer treatment. Cell Prolif. 47(5):391-5.
Ning B, Su Z, Mei N, Hong H, Deng H, Shi L, Fuscoe JC, Tolleson WH. (2014) Toxicogenomics and cancer susceptibility: advances with next-generation sequencing. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev. 32(2):121-58.
Pathak RR, Davé V. (2014) Integrating omics technologies to study pulmonary physiology and pathology at the systems level. Cell Physiol Biochem. 33(5):1239-60.
Samantarrai D, Dash S, Chhetri B, Mallick B. (2013) Genomic and epigenomic cross-talks in the regulatory landscape of miRNAs in breast cancer. Mol Cancer Res. 11(4):315-28.

Eberl HC, Mann M, Vermeulen M. (2011) Quantitative proteomics for epigenetics. Chembiochem. 24;12(2):224-34.
Tollefsbol TO. (2011) Advances in epigenetic technology. Methods Mol Biol. 791:1-10.
Xin Y, O'Donnell AH, et al. (2011) Role of CpG context and content in evolutionary signatures of brain DNA methylation. Epigenetics. 6(11):1308-18.
Lv J, Liu H, et al (2012) DiseaseMeth: a human disease methylation database. Nucleic Acids Res. 40(Database issue):D1030-5.
Miura F, Enomoto Y,et al. (2012) Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. Nucleic Acids Res. 1;40(17):e136.
Baek SJ, Yang S, et al. (2012) MENT: Methylation and expression database of normal and tumor tissues. Gene. Dec 7. pii: S0378-1119(12)01453-9. [Epub ahead of print]
Trask MC, Mager J. (2011) Complexity of polycomb group function: diverse mechanisms of target specificity. J Cell Physiol. 226(7):1719-21.
Hoki Y, Ikeda R, et al. (2011) Incomplete X-inactivation initiated by a hypomorphic Xist allele in the mouse. Development. 138(13):2649-59.
Law JA, Vashisht AA, et al. (2011) SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. PLoS Genet. 7(7):e1002195.
Baysal BE, McKay SE, (2011) Genomic imprinting at a boundary element flanking the SDHD locus. Hum Mol Genet. 15;20(22):4452-61. doi: 10.1093/hmg/ddr376.
Jung CJ, Iyengar S, et al. (2011) Epigenetic modulation of miR-122 facilitates human embryonic stem cell self-renewal and hepatocellular carcinoma proliferation. PLoS One. 2011;6(11):e27740.
Cabianca DS, Casa V, et al. (2012) A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. Cell. 11;149(4):819-31.
Rajasethupathy P, Antonov I, et al. (2012) A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. Cell. 27;149(3):693-707.
Gontan C, Achame EM, et al. (2012) RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. Nature. 29;485(7398):386-90.
Iglesias-Platas I, Martin-Trujillo A, et al. (2012) Characterization of novel paternal ncRNAs at the Plagl1 locus, including Hymai, predicted to interact with regulators of active chromatin. PLoS One. 2012;7(6):e38907.
Girardot M, Cavaillé J, Feil R. (2012) Small regulatory RNAs controlled by genomic imprinting and their contribution to human disease. Epigenetics. 15;7(12).
Veiseth SV, Rahman MA, et al. (2011) The SUVR4 histone lysine methyltransferase binds ubiquitin and converts H3K9me1 to H3K9me3 on transposon chromatin in Arabidopsis. PLoS Genet. 7(3):e1001325.
Seong KH, Li D, Shimizu H, (2011) Inheritance of stress-induced, ATF-2-dependent epigenetic change. Cell. 24;145(7):1049-61.
Grau DJ, Chapman BA, (2011) Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. Genes Dev. 15;25(20):2210-21.
Baron R, Vellore NA. (2012) LSD1/CoREST is an allosteric nanoscale clamp regulated by H3-histone-tail molecular recognition. Proc Natl Acad Sci U S A. 31;109(31):12509-14.
Eapen SA, Netherton SJ, et al. (2012) Identification of a novel function for the chromatin remodeling protein ING2 in muscle differentiation. PLoS One. 7(7):e40684.
Fernando RN, Eleuteri B, et al. (2011) Cell cycle restriction by histone H2AX limits proliferation of adult neural stem cells. Proc Natl Acad Sci U S A. 5;108(14):5837-42.

Xiong J, Wang H, et al. (2011) Male germ cell apoptosis and epigenetic histone modification induced by Tripterygium wilfordii Hook F. PLoS One. 6(6):e20751.
Krishnan V, Chow MZ, et al. (2011) Histone H4 lysine 16 hypoacetylation is associated with defective DNA repair and premature senescence in Zmpste24-deficient mice. Proc Natl Acad Sci U S A. 26;108(30):12325-30.
Hanover JA, Krause MW, Love DC. (2012) Bittersweet memories: linking metabolism to epigenetics through O-GlcNAcylation. Nat Rev Mol Cell Biol. 23;13(5):312-21.
Iglesias-Platas I, Martin-Trujillo A, et al. (2012) Characterization of novel paternal ncRNAs at the Plagl1 locus, including Hymai, predicted to interact with regulators of active chromatin. PLoS One. 2012;7(6):e38907.
Musselman CA, Avvakumov N, et al. (2012) Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. Nat Struct Mol Biol. 19(12):1266-72.
Chen Q, Chen Y, et al. (2012) TET2 promotes histone O-GlcNAcylation during gene transcription. Nature. Dec 9. doi: 10.1038/nature11742. [Epub ahead of print]
Dong F, Song Z, et al. (2013) Global transcriptional analysis of nuclear reprogramming in the transition from MEFs to iPSCs. Genes Cells. 18(1):42-55.
Sharma S, De Carvalho DD, et al. (2011) Nucleosomes containing methylated DNA stabilize DNA methyltransferases 3A/3B and ensure faithful epigenetic inheritance. PLoS Genet. 3;7(2):e1001286.
Li W, Han Y, et al. (2011) Knockdown of SAMS genes encoding S-adenosyl-l-methionine synthetases causes methylation alterations of DNAs and histones and leads to late flowering in rice. J Plant Physiol. 15;168(15):1837-43.
Guo JU, Ma DK, et al. (2011) Neuronal activity modifies the DNA methylation landscape in the adult brain. Nat Neurosci. 28;14(10):1345-51.
Knezovich JG, Ramsay M. (2012) The effect of preconception paternal alcohol exposure on epigenetic remodeling of the h19 and rasgrf1 imprinting control regions in mouse offspring. Front Genet. 2012;3:10.
Iglesias-Platas I, Martin-Trujillo A, et al. (2012) Characterization of novel paternal ncRNAs at the Plagl1 locus, including Hymai, predicted to interact with regulators of active chromatin. PLoS One. 7(6):e38907.
Lu C, Thompson CB. (2012) Metabolic regulation of epigenetics. Cell Metab. 3;16(1):9-17.
Zeng J, Konopka G, Hunt BG, et al. (2012) Divergent whole-genome methylation maps of human and chimpanzee brains reveal epigenetic basis of human regulatory evolution. Am J Hum Genet. 7;91(3):455-65.
Buck-Koehntop BA, Stanfield RL, et al. (2012) Molecular basis for recognition of methylated and specific DNA sequences by the zinc finger protein Kaiso. Proc Natl Acad Sci U S A. 18;109(38):15229-34.
Seifert M, Cortijo S, Colomé-Tatché M, et al. (2012) MeDIP-HMM: genome-wide identification of distinct DNA methylation states from high-density tiling arrays. Bioinformatics. 15;28(22):2930-9.
Du J, Zhong X, et al. (2012) Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell. 28;151(1):167-80.
Lam LL, Emberly E, et al. (2012) Factors underlying variable DNA methylation in a human community cohort. Proc Natl Acad Sci U S A. 16;109 Suppl 2:17253-60.
Yu CC, Furukawa M, et al. (2012) Genome-wide DNA methylation and gene expression analyses of monozygotic twins discordant for intelligence levels. PLoS One. 7(10):e47081.

Zhang D, Li S, Tan Q, Pang Z. (2012) Twin-based DNA methylation analysis takes the center stage of studies of human complex diseases. J Genet Genomics. 20;39(11):581-6.
Arabsolghar R, Azimi T, Rasti M. (2012) Mutant p53 binds to estrogen receptor negative promoter via DNMT1 and HDAC1 in MDA-MB-468 breast cancer cells. Mol Biol Rep. Dec 15. [Epub ahead of print]
Mohamed Ariff I, Mitra A, Basu A. (2012) Epigenetic regulation of self-renewal and fate determination in neural stem cells. J Neurosci Res. 90(3):529-39.
Nicol-Benoît F, Le-Goff P, et al. (2012) Epigenetic memories: structural marks or active circuits?
Cell Mol Life Sci. 69(13):2189-203.
Hochberg Z, Feil R, et al (2011) Child health, developmental plasticity, and epigenetic programming. Endocr Rev. 32(2):159-224.
Choudhuri S. (2011) From Waddington's epigenetic landscape to small noncoding RNA: some important milestones in the history of epigenetics research. Toxicol Mech Methods. 21(4):252-74.
Faulk C, Dolinoy DC. (2011) Timing is everything: the when and how of environmentally induced changes in the epigenome of animals. Epigenetics. 6(7):791-7.
Baylin SB, Jones PA. (2011) A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer. 23;11(10):726-34.
Murren CJ. (2012) The integrated phenotype. Integr Comp Biol. 52(1):64-76.
Karmaus W, Ziyab AH, Everson T, Holloway JW. (2013) Epigenetic mechanisms and models in the origins of asthma. Curr Opin Allergy Clin Immunol. 13(1):63-9.
Van Speybroeck L. From epigenesis to epigenetics: the case of C. H. Waddington. Ann N Y Acad Sci. 2002 Dec;981:61-81. Review.
Haig D. The (dual) origin of epigenetics. Cold Spring Harb Symp Quant Biol. 2004;69:67-70.
Morange M. The relations between genetics and epigenetics: a historical point of view. Ann N Y Acad Sci. 2002 Dec;981:50-60. Review.
Goldberg AD, Allis CD, Bernstein E. Epigenetics: a landscape takes shape.
Cell. 2007 Feb 23;128(4):635-8. Review.
Holliday R. DNA methylation and epigenotypes. Biochemistry (Mosc). 2005 May;70(5):500-4. Review.
Rivera RM, Bennett LB. Epigenetics in humans: an overview. Curr Opin Endocrinol Diabetes Obes. 2010 Dec;17(6):493-9.
Vaissière T, Sawan C, Herceg Z. Epigenetic interplay between histone modifications and DNA methylation in gene silencing. Mutat Res. 2008 Jul-Aug;659(1-2):40-8
Mitsuyoshi Nakao Epigenetics: interaction of DNA methylation and chromatin.
Gene, Volume 278, Issues 1-2, 31 October 2001, Pages 25-31
Nagase H, Ghosh S. Epigenetics: differential DNA methylation in mammalian somatic tissues. FEBS J. 2008 Apr;275(8):1617-23.
Zee BM, Levin RS, Dimaggio PA, Garcia BA. Global turnover of histone post-translational modifications and variants in human cells. Epigenetics Chromatin. 2010 Dec 6;3(1):22.
Pinskaya M, Morillon A. Histone H3 lysine 4 di-methylation: a novel mark for transcriptional fidelity? Epigenetics. 2009 Jul 1;4(5):302-6.
Cazzonelli CI, Millar T, Finnegan EJ, Pogson BJ. Promoting gene expression in plants by permissive histone lysine methylation. Plant Signal Behav. 2009 Jun;4(6):484-8.
Miller SA, Weinmann AS. An essential interaction between T-box proteins and histone-modifying enzymes. Epigenetics. 2009 Feb 16;4(2):85-8.

Thorvaldsen JL, Verona RI, Bartolomei MS. X-tra! X-tra! News from the mouse X chromosome. Dev Biol. 2006 Oct 15;298(2):344-53.
McEwen KR, Ferguson-Smith AC. Distinguishing epigenetic marks of developmental and imprinting regulation. Epigenetics Chromatin. 2010 Jan 15;3(1):2.
Pradhan S, Chin HG, Estève PO, Jacobsen SE. SET7/9 mediated methylation of non-histone proteins in mammalian cells. Epigenetics. 2009 Aug 16;4(6):383-7.
Hiragami-Hamada K, Xie SQ, Saveliev A, Uribe-Lewis S, Pombo A, Festenstein R. The molecular basis for stability of heterochromatin-mediated silencing in mammals. Epigenetics Chromatin. 2009 Nov 4;2(1):14.
Grimaud C, Nègre N, Cavalli G. From genetics to epigenetics: the tale of Polycomb group and trithorax group genes. Chromosome Res. 2006;14(4):363-75.
Nikolaou C, Althammer S, Beato M, Guigó R. Structural constraints revealed in consistent nucleosome positions in the genome of S. cerevisiae. Epigenetics Chromatin. 2010 Nov 12;3(1):20.
Roloff TC, Nuber UA. Chromatin, epigenetics and stem cells. Eur J Cell Biol. 2005 Mar;84(2-3):123-35.
Heard E, Chaumeil J, Masui O, Okamoto I. Mammalian X-chromosome inactivation: an epigenetics paradigm. Cold Spring Harb Symp Quant Biol. 2004;69:89-102
Costa FF. Non-coding RNAs: Meet thy masters. Bioessays. 2010 Jul;32(7):599-608.
Mosher RA, Melnyk CW. siRNAs and DNA methylation: seedy epigenetics. Trends Plant Sci. 2010 Apr;15(4):204-10.
Djupedal I, Ekwall K. Epigenetics: heterochromatin meets RNAi. Cell Res. 2009 Mar;19(3):282-95.
Costa FF. Non-coding RNAs, epigenetics and complexity. Gene. 2008 Feb 29;410(1):9-17.
Chuang JC, Jones PA. Epigenetics and microRNAs. Pediatr Res. 2007 May;61(5 Pt 2):24R-29R.
Morris KV. siRNA-mediated transcriptional gene silencing: the potential mechanism and a possible role in the histone code. Cell Mol Life Sci. 2005 Dec;62(24):3057-66.
Kawasaki H, Taira K, Morris KV. siRNA induced transcriptional gene silencing in mammalian cells. Cell Cycle. 2005 Mar;4(3):442-8
Hattori N, Shiota K. Epigenetics: the study of embryonic stem cells by restriction landmark genomic scanning. FEBS J. 2008 Apr;275(8):1624-30.
van de Nobelen S, Rosa-Garrido M, Leers J, Heath H, Soochit W, Joosen L, Jonkers I, Demmers J, van der Reijden M, Torrano V, Grosveld F, Delgado MD, Renkawitz R, Galjart N, Sleutels F. CTCF regulates the local epigenetic state of ribosomal DNA repeats. Epigenetics Chromatin. 2010 Nov 8;3(1):19.
Docherty SJ, Davis OS, Haworth CM, Plomin R, Mill J. Bisulfite-based epityping on pooled genomic DNA provides an accurate estimate of average group DNA methylation. Epigenetics Chromatin. 2009 Mar 10;2(1):3.
Tierling S, Schuster M, Tetzner R, Walter J. A combined HM-PCR/SNuPE method for high sensitive detection of rare DNA methylation. Epigenetics Chromatin. 2010 Jun 2;3(1):12.

[Epigenetics 1:2, 76-80; April/May/June 2006]; ©2006 Landes Bioscience

Review

Epigenetics

A Historical Overview

Robin Holliday

Correspondence to: Robin Holliday; 12 Roma Court; West Pennant Hills; N.S.W. 2125, Australia; Tel.: +61.2.9873.3476; Fax: +61.2.9871.2159; Email: RandL.Holliday@bigpond.com

Received 01/12/06; Accepted 03/15/06

Previously published online as an *Epigenetics* E-publication: http://www.landesbioscience.com/journals/epigenetics/abstract.php?id=2762

KEY WORDS

epigenetics, development, inheritance, DNA methylation, epimutation, epigenotype, epigenome, RNA, chromatin

ACKNOWLEDGEMENTS

I thank Julian Sale for providing some up-to-date references and a reviewer for several helpful suggestions.

ABSTRACT

In the first half of the twentieth century, developmental biology and genetics were separate disciplines. The word epigenetics was coined by Waddington to link the two fields. Epigenetics could be broadly defined as the sum of all those mechanisms necessary for the unfolding of the genetic programme for development. Several decades later specific mechanisms were proposed in which information was superimposed on DNA sequences. In particular, it was suggested that 5-methyl cytosine had a role in controlling gene expression, and also that the pattern of methylation was heritable. These predictions are now supported by a large body of evidence which shows that methylation is strongly associated with gene silencing in a variety of biological contexts. There are now also many examples of epigenetic inheritance through the germ line There are several other important epigenetic mechanisms involving chromatin and histone modifications, and also the expanding field of regulatory RNAs. The human epigenome project will unravel the pattern of DNA methylation in different tissues, and will this determine whether the regulation of gene expression is at the level of DNA or chromatin, or both.

INHERITANCE AND DEVELOPMENT

In the nineteenth century the leading biologists considered inheritance and development to be one and the same problem. The genius of Gregor Mendel was to realize, and then to demonstrate, that inheritance could be studied on its own, without including development. In a scholarly review and discussion of nearly 70 pages, Sandler and Sandler[1] explain that this was the major reason why Mendel's work was ignored by the leading biologists of his day. When it was finally re-discovered thirty five years later, the science of genetics subsequently flourished. Again, the problem of development was sidelined, and it is remarkable that one of the pioneers of the new genetics, Thomas Hunt Morgan, was by background an embryologist, but his laboratory did not study *Drosophila* development. It was only in his books that he re-visited embryology.

Whilst the science of genetics was making rapid progress, embryologists and developmental biologists were using methods and procedures that took little account of genes and gene action. Towards the middle of the twentieth century, there were a few leading biologists who realized that genetics and developmental biology were indeed related and should eventually come together in a common discipline. One was Conrad Waddington, who was knowledgeable in both fields of research. He took the Greek word epigenesis, a theory of development which proposed that the early embryo was undifferentiated, and changed it to epigenetics.[2] He was the Buchanan Professor of Genetics at Edinburgh University, and he also set up an Epigenetics Research Unit supported by the Medical Research Council for some years. Epigenetics could be broadly defined as the unfolding of the genetic program for development, but to Waddington, epigenetics was not very different from embryology. For example, his book *The Epigenetics of Birds* is largely an account of the development of the chick.[3] He also coined the term epigenotype, which was defined as "The total developmental system consisting of interrelated developmental pathways through which the adult form of the an organism is realized."[2] This is so broad that it is not very useful, and I will return to a more specific definition of the epigenotype later on.

The another leading biologist interested in both genetics and development was Ernst Hadorn in Zurich. Many of his studies were on mutations that affect *Drosophila* development, and he also wrote a book *Developmental Genetics and Lethal Factors*.[4] He also worked for many years on the remarkable properties of the imaginal discs of *Drosophila*. These are regions of embryonic tissue that are present in fly larvae. Each disc will later

develop into a specific adult structure: two for each wing, two for antennae, and so on. The disc cells are completely undifferentiated, but it can be said that they are determined to differentiate later on. Hadorn and his colleagues grew disc tissue in the abdomen of adult flies, and passaged it from fly to fly. When the disc tissue was treated with the hormone ecdysone, it differentiated into the appropriate adult structure. In other words, the determined state was heritable, sometimes for hundreds of cell divisions. However, from time to time the disc changed from one determined state to another, for example from a leg to a wing. This event was called transdetermination, and in innumerable studies it was shown that transdetermination followed certain pathways. For example, disc A could change into disc B, and B to C, but A never changed directly into C. This remarkable experimental system (reviewed in ref. 5) has not been exploited in modern experimental studies. Everything that is known about it comes from Hadorn's laboratory years ago.

Waddington and Hadorn were not the only important biologists who wanted to make connections between genetics and development. Another was Richard Goldschmidt, but his views were quite controversial (see ref. 6). Others, such as Julian Huxley[7] and J.B.S. Haldane, certainly understood the importance of the relationship, but the latter was particularly interested in the the biochemistry of gene activity. In this area there had been the early insights of Garrod, who realized that some inherited defects in man blocked specific steps in metabolic pathways.[8] This interpretation was ignored for many years, until Haldane became involved in the genetics of pigment formation in plants, and Ephrussi and Beadle attempted similar studies in *Drosophila*. Finally, Beadle and Tatum started to isolate biochemical mutants and their effects on metabolic pathways in Neurospora. Their work was very successful and culminated in the concept of one gene-one enzyme,[9] which was eventually verified in the 1950s. However, it was independent of studies of development.

After Waddington, there was spasmodic discussion of epigenetics by several scientists; much of this was reviewed by Nanney[10] and much more recently by Haig.[11] Some of the examples related to cytoplasmic inheritance, the phenotypes of cultured mammalian cells, or cancer cells. In general, observations that were not easily interpreted in genetic terms but had a heritable component, were liable to be labeled epigenetic. However, each author had his own idea of the meaning or definition of epigenetics. and no specific mechanisms were proposed. This was also true of the earlier work of Waddington, although he did introduce important new concepts such as canalization.[12]

THE NEED FOR EPIGENETIC MECHANISMS

The importance of the work of Waddington and Hadorn was to relate genes and gene action to development, in an environment in which most geneticists and most developmental biologists were not communicating with each other. As time went on, it became apparent that there were certain fundamental features of development that demanded explanation. One was the fact that differentiated cells, such as fibroblasts or lymphocytes, stably maintain their phenotypes through cell division. This means that some specialized genes which determine the phenotype of differentiated cells are permanently turned on, and other genes—active in some other cell type—are permanently turned off. These controls are heritable, just as the determined state of *Drososphila* disc cells are heritable. Traditionally, inheritance refers to the transmission of genes from generation to generation, but it was now realized that there is also mitotic inheritance in somatic cells of higher organisms. Of course, such inheritance had long been studied in yeasts and fungi, and then in cultured mammalian cells, but it had rarely been spelled out that it also regularly occurred in vivo, that is, in the normal somatic cells of higher organisms with specialized phenotypes. Another feature of higher organisms is the stem cell. Here an undifferentiated cell divides to produce a differentiated cell, and another undifferentiated stem cell. In the case of bone marrow stem cells, a variety of blood cell types are produced. In this situation there are clearly switches in gene activity associated with cell division. A third example is the X chromosome of female eutherian mammals. Early in development one X chromosome is randomly inactivated in every cell, whilst the other remains active. These two chromosome have almost identical DNA sequences, and they reside in a common nucleoplasm and cytoplasm, so the differences in gene activity are intrinsic to the chromosomes themselves. It is evident that there is a switch mechanism early in development, the result of which is the inactivity of one chromosome and the activity of the other. The switch is random and once made it is permanent. This example therefore embodies both a switch in gene activities and also its subsequent heritability.

The first suggestion that DNA methylation (or demethylation) might have an important biological role was made by Griffith and Mahler, who proposed in 1969 that it could provide a basis for long term memory in the brain.[13] In 1975 two papers were published which outlined a molecular model for the switching of gene activities, and also the heritability of gene activity or inactivity. It was based on the enzymic methylation of cytosine in DNA, which can also be referred to as DNA modification. The proposals by Riggs[14] and Holliday and Pugh[15] were very similar, but were made completely independently of each other. The suggestion was that DNA methylation could have strong effects on gene expression, and that changes in DNA methylation might therefore explain the switching on and off of genes during development. The enzyme(s) methylating a particular region of DNA would be sequence specific, or interact with another protein that was sequence specific. It was also proposed that the pattern of methylation could be heritable, if there was an enzyme called a maintenance methylase that recognized hemimethylated DNA soon after replication, but did not act on unmethylated DNA. This provides a mechanism for the heritability of the methylated and non-methylated state of DNA, and therefore for the heritability of a given pattern of gene activities. The issue of X chromosome inactivation was addressed particularly by Riggs. There might be an initial methylation that was immediately shut off, so that only one chromosome is marked. There would also have to be a spreading mechanism which inactivated the whole chromosome. Since it was much easier to envisage a processive methylating enzyme than the reverse, this implies that methylation of DNA is associated with gene inactivity. This can also explain the inactivation or silencing of autosomal DNA in several cases of X-autosome translocations.

There was also the possibility that developmental clocks might be important in unfolding the genetic program for development. This would be a mechanism that counts a specific number of cell divisions before a given gene or genes is activated or inactivated, and several molecular models were discussed.[15] Although there is scattered evidence for developmental clocks, it is not a commonly discussed topic, and only time will tell whether they are a significant component of development. As well as DNA methylation, there was also the possibility that specific base changes might occur, for example, the enzymic deamination of 5-methyl cytidine to form thymidine, and thus the substitution of an G-C base pair by a A-T base pair, a mechanism

that had previously been proposed by Scarano.[16] The existence of the enzyme cytidine deaminase which converts cytosine to uracil in DNA is now very well documented in the immune system and also in pluripotent cells.[17,18]

A third paper on DNA methylation by Sager and Kitchin also appeared in 1975, which proposed that there are enzymes in eukaryotic organisms that restrict unmodified DNA.[19] They explored the possibility that the many known examples of chromosome elimination or silencing might involve such a mechanism. It also became apparent that changes in DNA methylation might be important in tumor progression.[20,21] There was much accumulating evidence that changes in gene expression in cancer cells was due to mutation, but if the methylation model was correct, then aberrant changes in the distribution of 5-methyl cytosine in cancer cells could also result in changes in gene expression. The word epigenetics was not used in any of the 1975 papers on DNA methylation and gene expression, possibly because it had previously been used in several quite different contexts and remained undefined.[10,11]

EVIDENCE RELATING DNA METHYLATION TO GENE EXPRESSION

In 1975 when the DNA methylation models were proposed, there was no experimental evidence to support them. Nor did the models predict that specific DNA methylation would be associated with the activity or inactivity of genes. However, the spreading model for X chromosome inactivation did propose that methylation was the basis for such inactivation. With the cloning and sequencing of DNA, means were discovered for screening DNA methylation in specific DNA sequences. There were restriction enzymes which recognize and cut unmethylated sequences of DNA (usually four or six bases). In some cases there were two restriction enzymes which recognized the same base sequences, but only one of them would cut this sequence when it was methylated. This pair of enzymes were called isoschizomers, and examples were Hpa II and Msp I. Both cut DNA at GCGC sites, but only Msp cuts this sequence if the internal C is methylated. Using Southern blots it became possible to determine whether a given sequence containing a GCGC site was methylated or not. It was soon discovered that many genes with methylated promoter regions were inactive, and also that the corresponding active gene was unmethylated. This early work was reviewed by Doerfler.[22,23] A limitation of the method is that it detects only a subset of possible methylation sites, usually about 10%. Later on a more powerful method was introduced which can detect all methylated and non-methylated cytosine sites in a given stretch of DNA (see below).

Other evidence for the significance of DNA methylation came from the use of the nucleoside analogue 5-azacytidine. This is incorporated into DNA, inactivates DNA methyl transferase and thereby demethylates DNA. It was shown in many contexts that azacytidine reactivates silent genes, often at very high frequency (reviewed in ref. 24). This included the reactivation of genes on the inactive X chromosome. It had been shown that strains could be isolated in cultured mammalian cells which had biochemical deficiencies. Originally it was thought that these were mutations, but it now became apparent that they were often genes silenced by methylation, reactivable by 5-azacytidine.[25]

DEFINITIONS OF EPIGENETICS

Waddington did not use a specific definition for epigenetics. What he had in mind was: "All those events which lead to the unfolding of the genetic program for development." There is nothing wrong with that, except that it is not very specific. By the mid-1980s it was clear that there was a new type of inheritance, not based on changes in DNA sequence. In 1987 I wrote a paper "The inheritance of epigenetic defects."[26] In this I re-visited Waddington's use of the term, and I applied it to situations where changes in DNA methylation also changed gene activity. Possible epigenetic changes in cancer and also in ageing were discussed, and it was also suggested that some transgenerational effects that could not easily be explained by Mendelian genetics, might sometimes be due to the transmission of DNA methylation, or lack of it, through the germ line. It was also possible that some epigenetic defects might be recognized and repaired by genetic recombination at meiosis. At this time the word epimutation was introduced to describe heritable changes in genes which were not due to changes in DNA sequence. It has been suggested that this 1987 publication "was the critical paper that lit the fuse for the explosion in use of 'epigenetic' in the 1990s".[11]

Genomic imprinting in mammals had by now been discovered, and it was apparent that this was due to information superimposed on DNA that could be reversed at meiosis or during gametogenesis. New definitions of epigenetics were needed, and two were suggested in 1994: 1) The study of the changes in gene expression which occur in organisms with differentiated cells, and the mitotic inheritance of given patterns of gene expression, and 2) Nuclear inheritance which is not based on changes in DNA sequence.[27] The first definition is quite broad, which can include DNA methylation, but also a number of other mechanisms. The second definition includes imprinting and many other documented cases of epigenetic inheritance. It excludes cytoplasmic events, but they can be included in the first definition. Both definitions are in fact incomplete, but they seem to cover most known epigenetic processes. They do not include development itself, and for that we can use Waddington's general definition, which will become more specific as new information accumulates in the future.

DIFFERENCES BETWEEN GENETIC AND EPIGENETIC SYSTEMS

Much less is known about the epigenetic inheritance system than traditional genetics. Genetics is based on cell lineages and clonal inheritance. Gametogenesis produces single haploid cells that fuse to form a diploid zygote. The organism thus starts as a single cell, and ends up as a clone of cells. If a mutation or chromosomal change occurs in a somatic cell, then all its descendents would be expected to have the same genotype. In contrast, epigenetic changes often occur in groups of cells, for example, the induction of muscle tissue in mesoderm cells. This is due to a specific signal which impinges on a group of cells with the same receptor. Some epigenetic events are clonal, and X chromosome inactivation is an excellent example. Genetic changes are stable and rarely reversed, whereas epigenetic changes are often reversed. A good example of that is genomic imprinting, where the changes imposed on DNA sequences may be lost during development, or if they persist, are erased and re-set during gametogenesis. Environmental influences do not change the genotype (leaving aside mutagens), and there is no inheritance of acquired characteristics. Epigenetics is quite different, because normal development depends on communication between cells. Thus, a hormone, morphogen or growth factor may induce an epigenetic change that

may be heritable. This means that the environment of a cell may be all important in determining its properties or its fate in the developing organism. In this sense, epigenetics encompasses Lamarckian inheritance.

Maynard Smith[28] introduced the term dual inheritance, by which he meant that there is classical inheritance based on changes in DNA sequence, and also epigenetic inheritance which is not based on changes in DNA sequence. He was responding to the proposals by Jablonka and Lamb[29,30] that epigenetic inheritance in the germ line might introduce the possibility that environmental influences which induce phenotypic changes could become heritable. There are now many well documented examples of transgenerational effects, presumed to have an epigenetic basis.[31-38] Dual inheritance has also been demonstrated in experiments with cultured mammalian cells.[39] In some cases, what had long been thought to be a classical mutation has been shown to be due to a heritable change in DNA methylation, and a good example of that is a well known change in floral symmetry,[40] which can now be labeled as an epimutation. Transgenerational epigenetic inheritance and related topics have been recently reviewed by Jablonka and Lamb.[41]

It is well established that DNA methylation is involved in genomic imprinting, but the biological reasons for the existence of imprinting remain a matter for debate. (reviewed in ref. 42). One interesting possibility arises from the fact that imprinting results in haploid gene activity, because one of the gametes has an inactive gene. It may be important in early development to have single copies of single genes, particularly if a switch in gene activity takes place prior to or during division. Switching two copies has more than one consequence, but switching one simply leads to a plus and minus situation.[43] A challenge for the future is the unravelling the specificity of genomic re-programming when the germ cells and fertilized egg are formed. Little is yet known about this, although it is established that there are massive changes in DNA methylation at this time and also in early development. These are global changes, whereas information is needed about specific changes, as have been established in the case of imprinting.

OTHER EPIGENETIC MECHANISMS

Chromatin structure and gene expression has become an intensively active field of research. Chromatin can be in the open form that allows access of the machinery for transcription, and a closed form which does not allow transcription. The modification of histones, particularly acetylation and methylation, play a crucial role in this change, and many believe that it is this switch, rather than DNA methylation, which is the more important (reviewed in ref. 44). However, it is not at all obvious how chromatin configurations can be stably inherited. The evidence that DNA methylation can provide a primary switch is very strong,[45] and one likely possibility is that the presence of such methylation triggers the changes that lead to the closed chromatin configuration.

The role of RNA in epigenetic events has become increasingly important. The alternative splicing of gene transcripts can be regarded as an epigenetic mechanism This can produce many isoforms of a given protein that have subtly different properties, and distinct cell types are likely to have specific isoforms. The specificity of splicing events remains a problem, which might be solved if there were small RNA molecules that hybridized across splice junctions.[46] Another prediction is that there are large RNA molecules in the egg or early embryo that have an essential spatial, positional or structural role.[47] This could be essential for the correct 3-dimensional distribution of proteins. If substantiated, this can also be regarded as an epigenetic mechanism. It is now evident that there are a huge number of small regulatory RNA molecules in cells (reviewed in refs. 48–50), and their activities comprise new epigenetic controls. An exciting possibility is that some of these molecules can transmit signals by moving from one cell to another.

The DNA sequence remains constant in most somatic cells, but there is a special epigenetic mechanism in cells of the immune system that can join one constant and one variable sequence, from a pool of such sequences in the whole region, to form a particular genotype that is clonally inherited. Another mechanism to generate antibody variability depends on enzymes that can deaminate cytosine to uracil, or 5methyl cytosine to thymine.[17] This is in effect a mutation, but induced by an enzyme. It could be argued that such a mutation is not an epigenetic change, but it is certainly the result of a protein-DNA interaction and in this respect is epigenetic.[18]

THE GENOME, THE EPIGENOME AND EPIGENOTYPES

In the sequence of the human genome there are just four bases, yet with cytosine in methylated or non-methylated form, there are five, and there is the possibility of six.[51] The epigenome project sets out to determine the pattern of cytosine methylation in a variety of cell types.[52] This depends on the bisulphite sequencing technique introduced in 1992.[53] Since then the technique has been greatly improved, but the underlying chemistry remains the same. It relies on the fact that bisulphite can deaminate cytosine to uracil under conditions in which 5-methyl cytosine is not deaminated. Thus when bisulphite-treated DNA is amplified and sequenced, all the 5-methyl cytosine residues remain as cytosine, but the non-methylated cytosines have become thymines. This technique has been applied in a large number of contexts, and particularly to demonstrate the methylation of many inactive tumor suppressor genes in cancer cells.[54]

The epigenome project will take a long time to complete; nevertheless along the way, we can expect that interesting information will be continually uncovered. We might expect that some regions of the DNA will have the same, or a very similar pattern of methylation in all cell types. These sequences will include many repetitive or transposable elements which have entered the genome at some time and have been silenced by DNA methylation. Much more interesting information will come from specialized genes that are active in one cell type and inactive in another. The importance of DNA methylation in determining the cell phenotype will then be revealed. In the epigenome project, a new terminology will be necessary to classify differences in DNA methylation between cell types.

This introduces the concept of the epigenotype. It has been suggested that the epigenotype is the actual pattern of gene activity in a specialized cell type.[55] These cells are said to have household enzymes and proteins, necessary for normal metabolism in all cell types, and also luxury proteins which have specialized functions. The epigenotype includes all those genes necessary for both household and luxury functions, and also those that are silent or repressed in a given cell type. Thus, fibroblasts and lymphocytes have the same genotype, inherited from the fertilized egg, but they have very different epigenotypes. Of course, as in the case of genotypes, any terminology may apply just to one gene or a subset of genes.

CONCLUSIONS

This overview began with a brief historical account of genetics and developmental biology, and how they diverged for a major part of the twentieth century. Epigenetics is the field that attempted to unite them, and provide new insights into the mechanisms for unfolding the genetic program for development. In the last two decades of the twentieth century much progress has been made on the relationship between DNA methylation and gene expression in a variety of biological contexts, and the experimental study of epigenetics was established. The field has now widened to include another of other mechanisms, especially those involving RNA. Many new insights into the mechanisms for development will be gained in this century.

The sequencing of the human genome is being followed by the epigenome project, which will eventually unravel the significance of DNA methylation in the control of specialized gene functions. It will become apparent whether the primary controls are at the DNA or at the chromatin level. In either case, the nature of the continual interactions between proteins and DNA will further advance the field of epigenetics, and illuminate current problems, such as the re-programming of the genome which initiates the normal processes of development.

References

1. Sandler I, Sandler L. A conceptual ambiguity that contributed to the neglect of Mendel's paper. History Phil Life Sciences 1985; 7:3-70.
2. Waddington CH. Introduction to Modern Genetics. London: Allen and Unwin 1939.
3. Waddington CH. Epigenetics of Birds. Cambridge: Cambridge University Press:1952.
4. Hadorn E. Developmental Genetics and Lethal Factors. 1960:Methuen, London. First published in German, 1955:G.T. Verlag, Stuttgart.
5. Ursprung H, Nothiger R. Biology of Imaginal Discs. Berlin: Springer Verlag 1972.
6. Dietrich MR. Richard Goldschmidt: hopeful monsters and other heresies. Nature Rev. Genet. 2003; 4:68-74.
7. Huxley J. Epigenetics. Nature 1956; 177:806-8.
8. Garrod AE. Inborn Errors of Metabolism. London: Froude, Hodder and Sloughton 1909. 2nd Edition, Oxford: Oxford University Press 1923.
9. Beadle GW, Tatum EL. Genetic control of biochemical reactions in Neurospora. Proc Nat Acad Sci USA 1941; 27:499-506.
10. Nanney DL. Epigenetic control systems. Proc Nat Acad Sci USA; 1958; 44:712-7.
11. Haig D. The (dual) origin of epigenetics. Cold Spring Harbor Symp Quant Biol 2004; LXIX :1-4.
12. Slack JMW. Conrad Hal Waddington: was he the last Rennaissance biologist? Nature Rev Genet 2002; 3:889-95.
13. Griffith JS, Mahler HR. DNA ticketing theory of memory. Nature 1969; 223:580 -2.
14. Riggs AD. X inactivation, differentiation and DNA methylation. Cytogenet.Cell Genet. 1975; 14:9-25.
15. Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. Science. 1975; 187:226-32.
16. Scarano E. The control of gene function in cell differentiation and in embryogenesis. Adv Cytopharmacol. 1971; 1:13-24.
17. Petersen-Mahrt S. DNA deamination in immunity. Immunol Rev 2005; 203:80-97.
18. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-induced cytidine deaminase deaminates 5-methyl cytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. J Biol Chem 2004; 279:52353-60.
19. Sager R, Kitchin R. Selective silencing of eukaryotic DNA. Science. 1975; 189:426-33.
20. Pugh JE. Holliday R. Do chemical carcinogens act by altering epigenetic controls through DNA repair rather than by mutations? Heredity 1978; 40:329.
21. Holliday R. A new theory of carcinogenesis. Brit J Cancer 1979; 40:512-3.
22. Doerfler W. DNA Methylation- A regulatory signal in eukaryotic gene expression. J Gen Virol 1981; 57:1-20.
23. Doerfler W. DNA methylation and gene activity. Ann Rev Biochem 1983; 52:93-124.
24. Jones PA. Altering DNA methylation with 5-azacytidine. Cell 1985; 40:485-6.
25. Holliday R. Mutations and epimutations in mammalian cells. Mutat. Res 1991; 250:345-63.
26. Holliday R. The inheritance of epigenetic defects. Science 1987; 238:163-70.
27. Holliday R. Epigenetics: an overview. Dev Genet 1994; 15:453-7.
28. Maynard Smith J. Models of a dual inheritance system. J Theoret Biol 1990; 143:41-53.
29. Jablonka E, Lamb M. The inheritance of acquired epigenetic variations. J Theoret Biol 1989; 139:59-83.
30. Jablonka E, Lamb M. Epigenetic inheritance and Evolution: the Lamarckian dimension. 1995: Oxford University Press, Oxford.
31. Morgan HD, Sutherland HGE, Martin DIK and Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. Nature Genetics 1999; 23:314-8.
32. Dubrova YE, Plumb MA, Guttierez B, Boulton E, Jeffreys A. Transgerational mutation by irradiation. Nature 2000; 405:37.
33. Barber R, Plumb MA, Boulton E, Roux I, Dubrova YE. Elevated mutation rates in the germ lineof first- and second-generation offspring of irradiated male mice. Proc Nat Acad Sci USA 2002; 99:6877-81.
34. Dubrova YE. Radiation-induced transgenerational instability. Oncogene 2003; 22:7087-93.
35. Rakyan VK, Chong S, Champ ME, Cuthbert PC, MorganHD, Luu KVK and Whitelaw. E 2003. Transgenerational inheritance of epigenetic states at the murine AxinFu allele occurs after maternal and paternal transmission. Proc Nat Acad Sci USA 2003; 100:2538-43.
36 Morgan WE. Non-targeted and delayed effects of exposure to ionising radiation: II radiation-induced genomic instability and bystander effects in vivo, clastogenic factors and transgenerational effects. Radiation Res 2003; 159:381-92.
37. Pogribny I, Raiche J, Slovack M, Kovalchuck O. Dose-dependence, sex and tissue specificity, and persistence of radiation-induced genomic DNA methylation changes. Biochem Biophys Res Comm 2004; 320:1253-61.
38. Anway MD, Cupp AS, Uzumcu M , and Skinner MK. Epigenetic transgenerational actions of endocrine disruptors and male fertility. Science 2005; 308:1466-9.
39. Paulin RP, Ho T, Balzer HJ, Holliday R. Gene silencing by DNA methylation and dual inheritance in Chinese hamster ovary cells. Genetics 1998; 149:1081-8.
40. Cubas P, Vincent C, and E. Coen E. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. Nature 1999; 401:157-61.
41 Jablonka E. and Lamb M. The changing concept of epigenetics. Ann NY Acad Sci 2002; 981:82-96.
42. Constancia M, Kelsey G, Reik W. Resourceful imprinting, Nature 2004; 432:53-7.
43. Holliday R. Genomic imprinting and allelic exclusion.. In: Monk M, Surani A eds Genomic Imprinting. Development Suppl. Company of Biologists, Cambridge 1990:125-9.
44. Cosgrove MS, Wolberger C. How does the histone code work? Biochem. Cell Biol. 2005; 83:468-76.
45. Holliday R. DNA methylation in eukaryotes: 20 years on. In Russo VEA, Riggs AD, Martienssen R, eds. Epigenetic mechanisms of gene regulation. New York. Cold Spring Harbor Laboratory Press ,1996:5-27.
46. Holliday R, Murray V. Specificity in splicing. BioEssays 1994; 16:771-4.
47. Holliday R. A molecular approach to the problem of positional information in eggs and early embryos. New Biologist 1989; 1:336-43.
48. Baulcome D. RNA silencing. Trends in Biochem Sci 2005; 30:290-3.
49. Sontheimer EJ, Carthew RW. Silence from within: endogenous siRNAs and miRNAs. Cell 2005; 122:9-12.
50. Filipowitz W. RNAi: the nuts and bolts of the RISC machine. Cell 2005; 122:17-20.
51. Kay PH, Pereira E, Marlow SA, Turbett G, Mitchell CA, Jacobsen PF, Holliday R, Papadimitriou. Evidence for adenine methylation within the mouse myogenic gene Myo-D1, Gene 1994; 151:89-95.
52. Beck S, Olek, A eds. The Epigenome: Molecular Hide and Seek. Weinheim: Wiley-VCH, 2003.
53. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CI. A genomic sequencing protocol which yield a positive display of 5-methyl cytosine residues in individual strands. Proc Nat Acad Sci USA 1992; 89:1827-31.
54. Millar DS, Holliday R. Grigg, GW. Five not four: history and significance of the fifth base. In: Beck S, Olek A, eds The Epigenome: Molecular Hide and Seek. Weinheim: Wiley-VCH, 2003:3-38.
55. Holliday R. DNA methylation and epigenotypes. Biochemistry 2005; 70:612-7.

Spring 2017 – Epigenetics and Systems Biology
Lecture Outline (Systems Biology)
Michael K. Skinner – Biol 476/576
Weeks 5, 6 and 7 (February 7, 14 and 21)

Epigenetics (History / Molecular Processes/ Genomics)

- Definitions and History
- Molecular Factors (DNA Methylation, Histone Modification, Chromatin Structure, ncRNA)
- Epigenetics Technology and Genomics

Required Reading

Holliday R. Epigenetics: a historical overview. Epigenetics. 2006 Apr-Jun;1(2):76-80.

Books (Reserve in Library)

Kevin V. Morris (2012) Non-coding RNAs and Epigenetic Regulation of Gene Expression: Drivers of Natural Selection. Caister Academic Press.

Russo, V.E.A., Martienssen, A.and Riggs, A.D. (eds.). 1996. Epigenetic Mechanisms of Gene Regulation. Cold Spring Harbor Press. Cold Spring Harbor.

Allis, C.D., Jenuwein, T. & Reinberg, D., Eds. (2007). Epigenetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Jeanteur, P. EPIGENETICS AND CHROMATIN. Progress in Molecular and Subcellular Biology, 2005, Volume 38, 151-167, DOI: 10.1007/3-540-27310-7_6

Epigenetic Definition

Conrad Waddington in the 1940s introduced the term epigenetic, but he had a very broad definition, namely:

"The total developmental system consisting of interrelated developmental pathways through which the adult form of an organism is realized".

This is too general and all-embracing to be useful, hence the more specific definition introduced here.

Huxley (1957) used epigenetics "to denote the analytic study of individual development (ontogeny) with its central problem of differentiation."

For Huxley, "The method by which tissues and organs differentiate in the course of normal development is at the moment the main blank space in biology's map. . .

Nanney (1958) "template replicating mechanism" that determined the "library of specificities." However, he believed that "auxiliary mechanisms with different principles of operation are involved in determining which specificities are to be expressed in any particular cell."

These auxiliary mechanisms he called epigenetic control systems.

In 1990, Holliday defined epigenetics as 'the study of the mechanisms of temporal and spatial control of gene activity during the development of complex organisms'.

His definition rescues Waddington's original meaning of developmental biology, although it does not differentiate between the action of what we currently know as epigenetic mechanisms and the action of genetic regulators of gene expression such as transcription factors.

Herring (1993), for whom epigenetics refers to "*the entire series of interactions among cells and cell products which leads to morphogenesis and differentiation.*"

She continues that "among the numerous epigenetic factors influencing the vertebrate face is mechanical loading" and that "epigenetic influences range from hormones and growth factors to ambient temperature and orientation in a gravitational field."

Riggs and colleagues in 1996 states that epigenetics is

'the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence'.

However, the term heritable is generally used in reference to generational inheritance and is not associated with growth of cells or tissues.

Bird (1990s) defines epigenetics as the 'structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states'.

Because there are several epigenetic elements that do not fit into this definition such as non-coding RNA and minor modifications of histones and DNA methylation of promoters, this definition appears insufficiently global to encompass all of epigenetics.

Epigenetics

Molecular factors/processes around the DNA that regulate genome activity, independent of DNA sequence, and are mitotically stable

Opposition

Epigenetics is an attempt to oppose the reductionist, piecemeal approach of genetics: epigenetics—and what belongs to epigenetics—can be defined only as a reaction against the current, dominant, reductionist approach of genetics. (Michel Morange 2002)

When it becomes possible to integrate data from different domains, I strongly suspect that epigenetics as such will disappear, because the epigenetic ideas will be omnipresent in the practice and ideas of geneticists. Its disappearance will be the sign of its victory. (Michel Morange 2002)

Epigenetic History

History of Epigenetics

1940s Conrad Waddington defined epigenetics as environment-gene interactions induce phenotype.

1975 Holliday and Pugh, and Riggs identify DNA methylation

1988 X- chromosome inactivation and DNA methylation

1990s Imprinted genes, allelic expression and DNA methylation

1995s Histone modifications and chromatin structure

2000s Small micro RNAs

2005s Epigenome mapping

Waddington argued the development of differences within a single organism, for example:

the difference between an eye and a nose...is clearly *neither* genotypic nor phenotypic. It is due...to the different sets of developmental processes which have occurred in the two masses of tissue; and these again can be traced back to local interactions between the various genes of the genotype and the already differentiated regions of the cytoplasm in the egg. One might say that the set of organizers and organizing relations to which a certain piece of tissue will be subject during development make up its *epigenetic constitution or epigenotype*; then the appearance of a particular organ is the product of the genotype and the epigenotype, reacting with the external environment." [italics added]

development as an *epigenetic process*

FROM

genotype + environment = phenotype

TO

genotype + *epigenotype* + environment = particular phenotype

⇓

= epigenetic constitution of tissue/cell
= set of organizers and organizing relations to which tissue is subject during development

Figure 2. Scheme of Waddington's expansion of the classical model on the phenotype-genotype distinction.

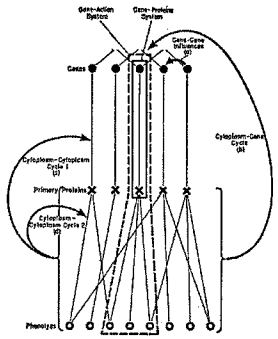
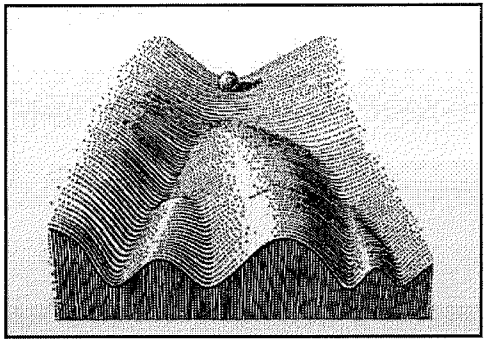


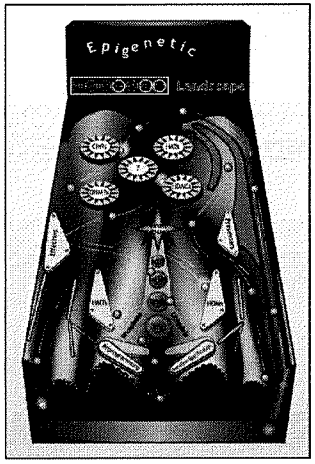
Figure 3. Waddington's original scheme of an epigenetic action system of a cell: the feedback reactions from (a) gene to gene, (b) cytoplasm to gene, (c) cytoplasm to gene-protein system, and (d) cytoplasm to the primary-protein-to-phenotype processes illustrate the gene contexts that Waddington focused on. (Reprinted from New Patterns in Genetics and Development by Conrad Hal Waddington © 1962, Columbia University Press; used with permission.)

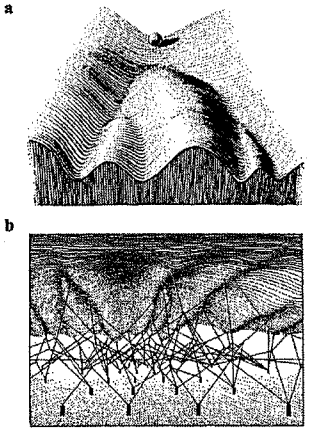
Waddington's Evo-Devo Program

Stressing the developing phenotype in evolutionary theory expands the classical focus on the transmission of genetic information with a second focus on gene regulation or instructions on how to use the genetic information. To bring evolution and development to full synthesis, however, a developmental theory is needed. Therefore, Waddington's epigenetics mainly situates itself on the developmental plane, as a model to link the genotype and the phenotype during development in a specific environmental context.



Waddington's Classical Epigenetic Landscape
In 1957, Conrad Waddington proposed the concept of an epigenetic landscape to represent the process of cellular decision-making during development. At various points in this dynamic visual metaphor, the cell (represented by a ball) can take specific permitted trajectories, leading to different outcomes or cell fates. Figure reprinted from Waddington, 1957.





History of Epigenetics

1940s Conrad Waddington defined epigenetics as environment-gene interactions induce phenotype.

1975 Holliday and Pugh identify DNA methylation

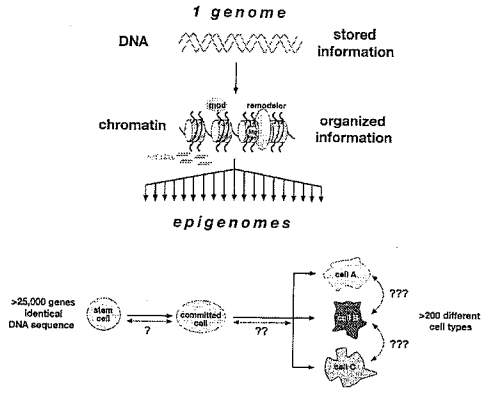
1988 X- chromosome inactivation and DNA methylation

1990s Imprinted genes, allelic expression and DNA methylation

1995s Histone modifications and chromatin structure

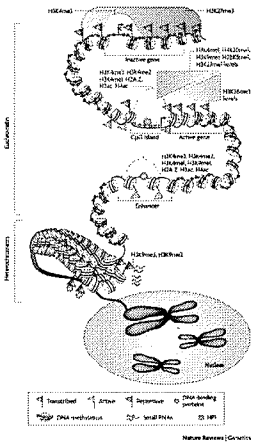
2000s Small micro RNAs

2005s Epigenome mapping



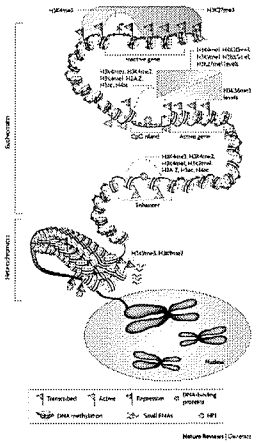
Epigenetic Mechanisms of Gene Regulation

-DNA Methylation
-Histone Modification
-Chromatin Structure
-DNA Organization into Domains (eg Loops)
-Nuclear Compartmentalization (eg nuclear matrix)
-Noncoding functional RNAs

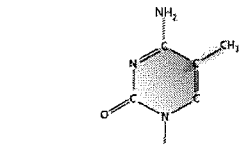


Epigenetic Mechanisms of Gene Regulation

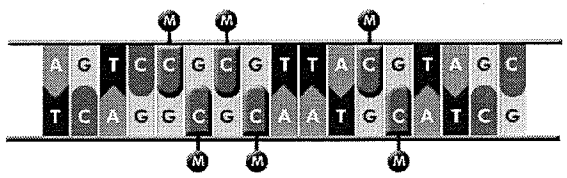
- -DNA Methylation
- -Histone Modification
- -Chromatin Structure
- -DNA Organization into Domains (eg Loops)
- -Nuclear Compartmentalization (eg nuclear matrix)
- -Noncoding functional RNAs
- -RNA Methylation



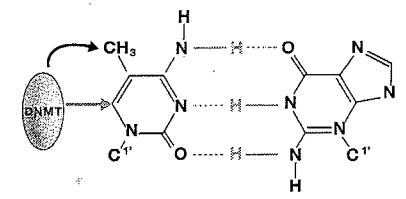
Epigenetics
DNA Methylation



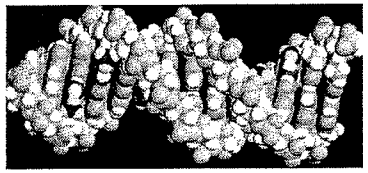
DNA Methylation is the addition of a methyl group (M) to the DNA base cytosine (C) in a CpG sequence

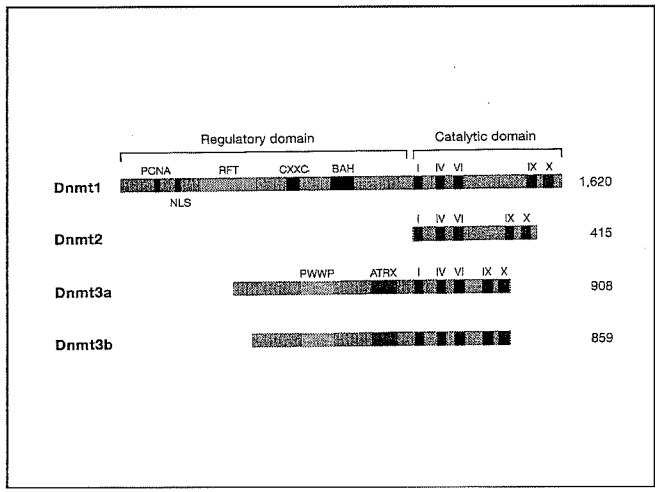


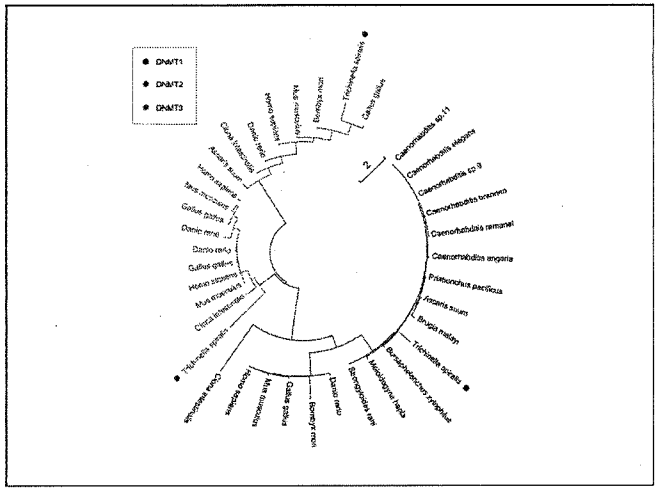
a



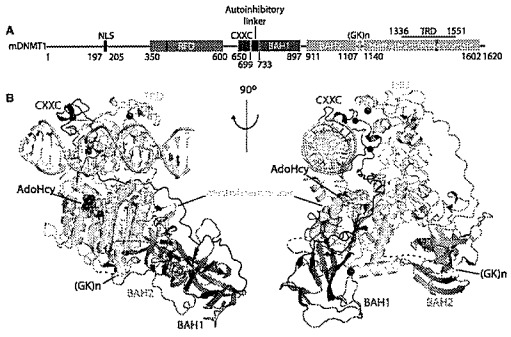
b







Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation.
Song J, et al. Science. 2011 Feb 25;331(6020):1036-40



Structural overview of mDNMT1(650–1602)–DNA 19-nucleotide oligomer complex with bound AdoHcy. (A) Color-coded domain architecture and numbering of mDNMT1 sequence. The thin vertical light blue bars indicate binding positions of zinc ions. (B) Ribbon representation of the complex in two orthogonal views. The CXXC, BAH1, BAH2, and methyltransferase domain are colored in red, light purple, orange and light blue, and DNA and zinc ions are colored in light brown and dark purple, respectively; CXXC-BAH1 linker in dark blue, BAH1-BAH2 linker in silver, (GK)n-containing BAH2-methyltransferase linker in black, and bound AdoHcy as in space-filling representation.

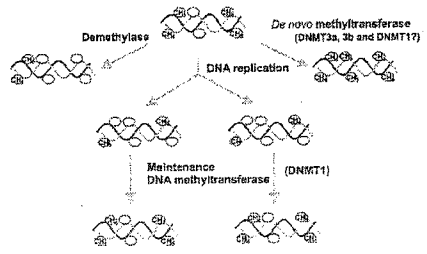
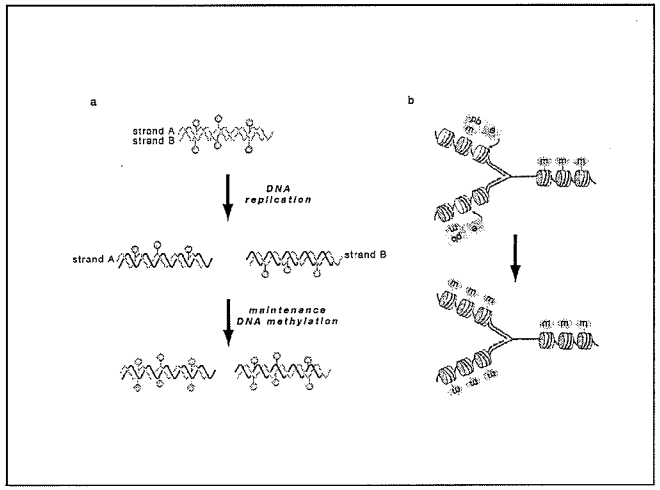
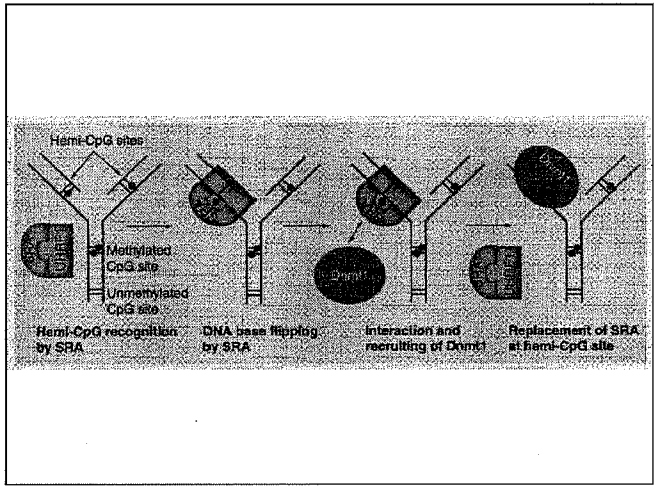
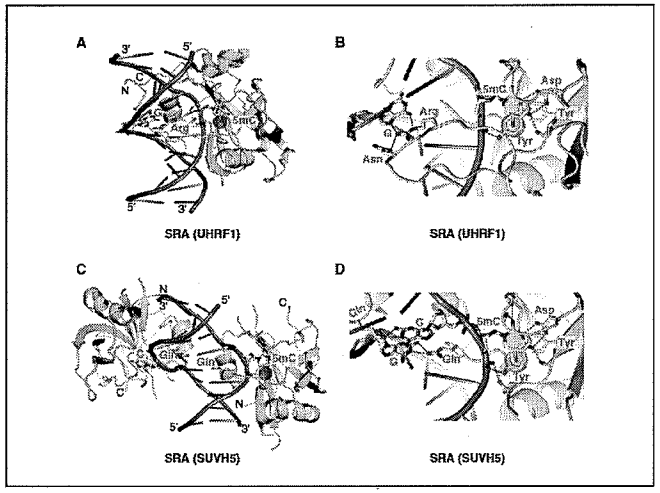
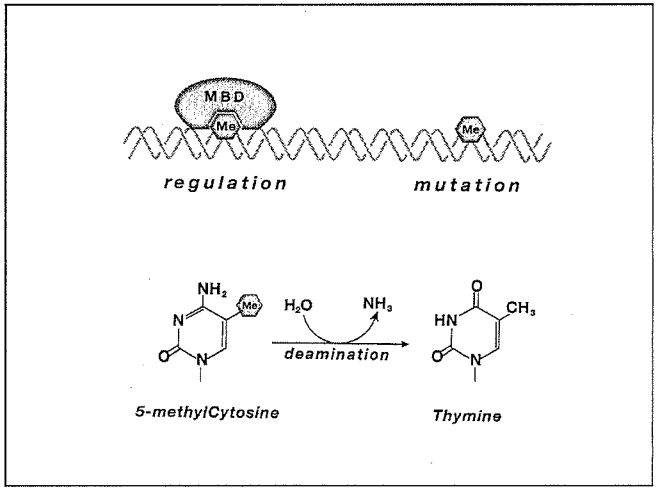


Figure 1.2 DNA methylation reactions. During embryogenesis the pattern of methylation is sculpted by a combination of DNA demethylation by demethylases and *de novo* methylation by DNA methyltransferases (DNMT). The pattern is then maintained in a tissue specific manner by maintenance methyltransferases (DNMT1) that copy the methylation pattern from parental to daughter strands.


a
strand A
strand B
DNA replication
strand A
strand B
maintenance DNA methylation
b


Hemi-CpG sites
Methylated CpG site
Unmethylated CpG site
Hemi-CpG recognition by SRA
DNA base flipping by SRA
Interaction and recruiting of Dnmt1
Replacement of SRA at hemi-CpG site


A
B
SRA (UHRF1)
SRA (UHRF1)
C
D
SRA (SUVH5)
SRA (SUVH5)


MBD
Me
Me
regulation
mutation
deamination
5-methylCytosine
Thymine

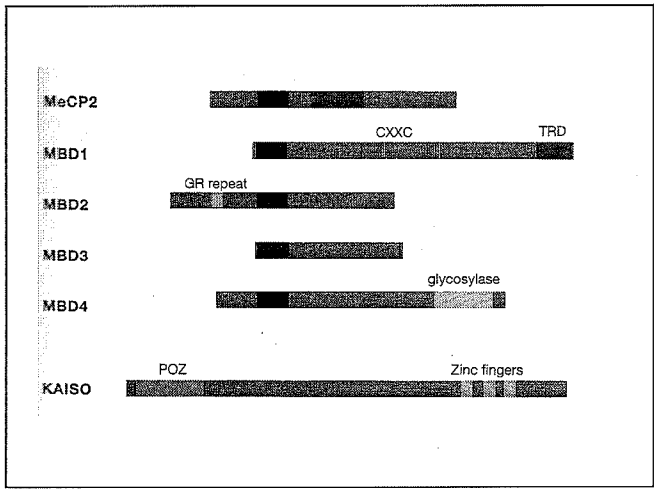
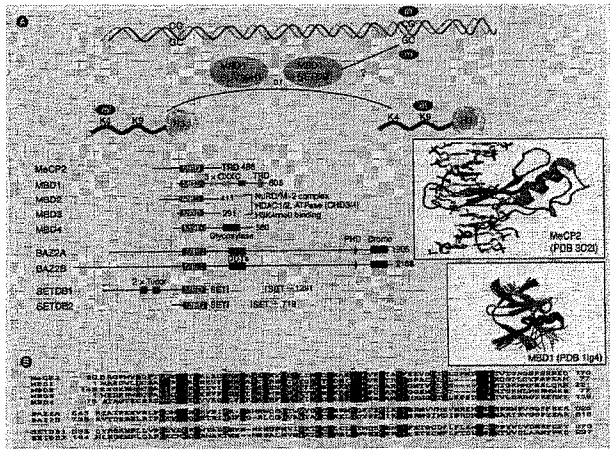
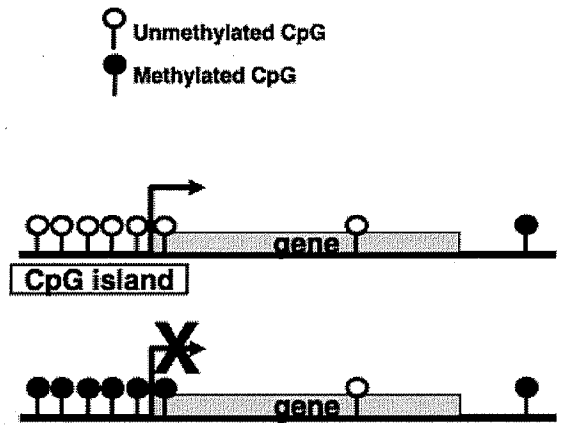
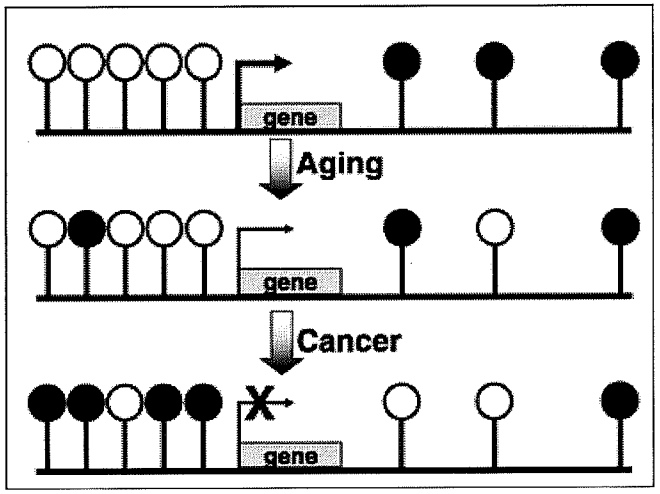


Table 2. Functions of methyl-CpG-binding proteins

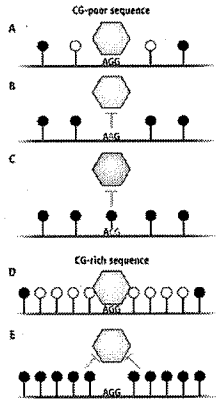
MBP	Major activity	Species	Major phenotypes of loss-of-function mutations
MeCP2	binds mCpG with adjacent A/T run; transcriptional repressor	mouse	delayed onset neurological defects including inertia, hindlimb clasping, nonrhythmic breathing, and abnormal gait; postnatal survival ~10 weeks
MECP2	binds mCpG with adjacent AT run; transcriptional repressor	human	heterozygotes suffer from Rett syndrome, a profound neurological disorder characterized by apraxia, loss of purposeful hand use, breathing irregularities, and microcephaly
Mbd1	binds mCpG via MBD; a major splice form is also able to bind CpG via a CxxC domain	mouse	no overt phenotype, but subtle defects in neurogenesis detected
Mbd2	binds mCpG; transcriptional repressor	mouse	viable and fertile, but show reduced maternal nurturing behavior; defects in gene regulation in T-helper-cell differentiation leading to altered response to infection; highly resistant to intestinal tumorigenesis
Mbd3	core component of NuRD co-repressor complex; does not show strong binding to mCpG	mouse	early embryonic lethal
Mbd4	DNA repair protein that binds mCpG and T:G mismatches at mCpG sites; thymine DNA glycosylase that excises T from T:G mismatches	mouse	viable and fertile; three- to fourfold increase in mutations at CpG sites; increased susceptibility to intestinal cancer correlates with C-to-T transitions within the Apc gene; Mbd4 functions to minimize the mutability of 5-methylcytosine
Kaiso	binds mCGmCG and CTGCNA; transcriptional repressor	mouse	no overt phenotype; small but significant delay in tumorigenesis on Min background







Molecular biology. Epigenetic islands in a genetic ocean.
Schübeler D.
Science. 2012 Nov 9;338(6108):756-7. doi: 10.1126/science.1227243. No abstract available.



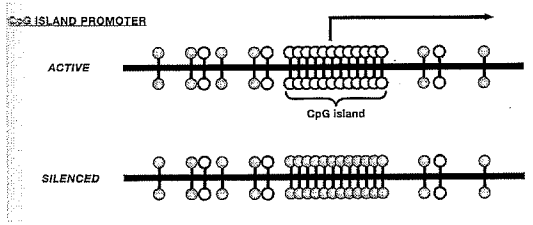
Genetics and epigenetics of transcription factor binding.

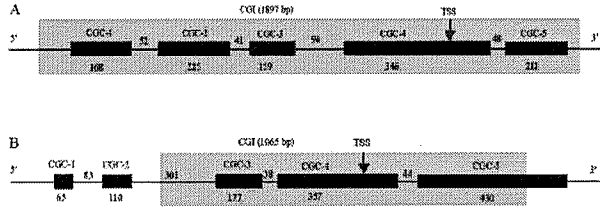
(A) Binding of a transcription factor in CG-poor regions leads to a local unmethylated state. (B) Mutations in the binding site prevent binding and result in increased methylation. (C) Some transcription factors could be sensitive to methylation even in CG-poor regions. (D) Transcription factor binding in a CG-rich area (CG island) requires the region to be unmethylated, and (E) can be blocked if methylated. (Hexagon) Transcription factors, (black circles) methylated CG, (white circles) unmethylated CG.

CpG Island

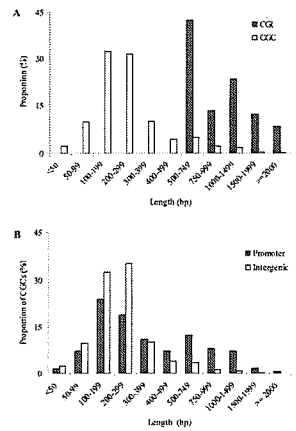
- 70- 80% of all CpG dinucleotides are heavily methylated in human cells.
- CpG islands (CGI) are 0.2 to 1-kb long DNA sequence stretches of GC-rich (G+C content: >50-60%) DNA that appear to be protected from the modification in somatic cells.
- CpG islands are frequently located in the promoters and first exon regions of 40 to 50% of all genes, but can be in introns and exons or between genes.
- Methylation typically results in loss of expression of adjacent genes.

WO 2007050706 20070503: DNA METHYLATION BIOMARKERS IN LYMPHOID AND HEMATOPOIETIC

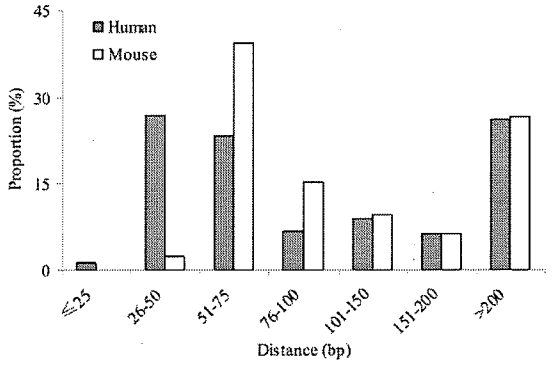




Multiple short CGCs embedded in one CGI in the promoter region. Dark box: CGCs identified by CpGcluster. Grey box: CGI identified by Takai and Jones' algorithm. The length of each CGC is labeled below the dark box and the distance between two neighboring CGCs is above the line. The transcription start site (TSS) is marked by an arrow. **(A)** CAP1. **(B)** ADAM33.



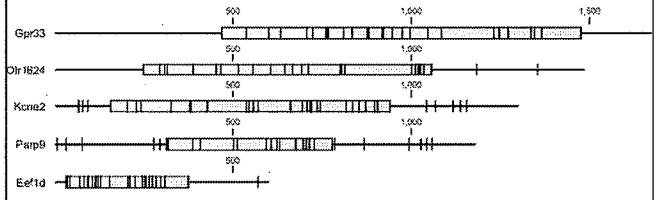
Length distribution of CGIs or CGCs in the human genome. (A) CGIs versus CGCs. **(B)** For CGCs, promoter regions versus intergenic regions.



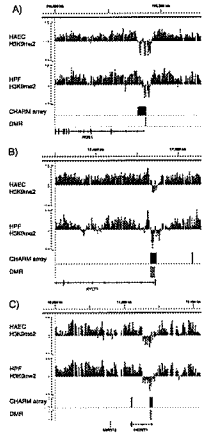
Distribution of distance between two neighbouring CGCs in the promoter region of a gene.

BMC Genomics. 2014 Aug 20;15:692. doi: 10.1186/1471-2164-15-692.
Role of CpG deserts in the epigenetic transgenerational inheritance of differential DNA methylation regions.
Skinner MK[1] and Guerrero-Bosagna C
[1]Center for Reproductive Biology, School of Biological Sciences, Washington State University, Pullman, WA 99164-4236, USA. skinner@wsu.edu.

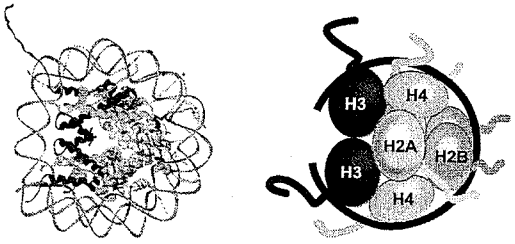
DMR and CpG Desert

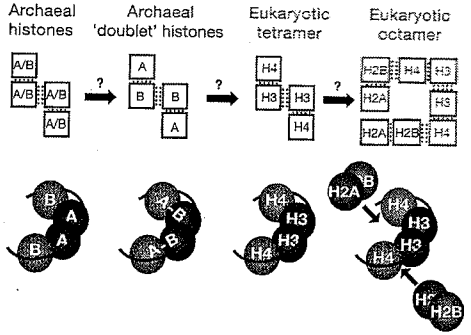


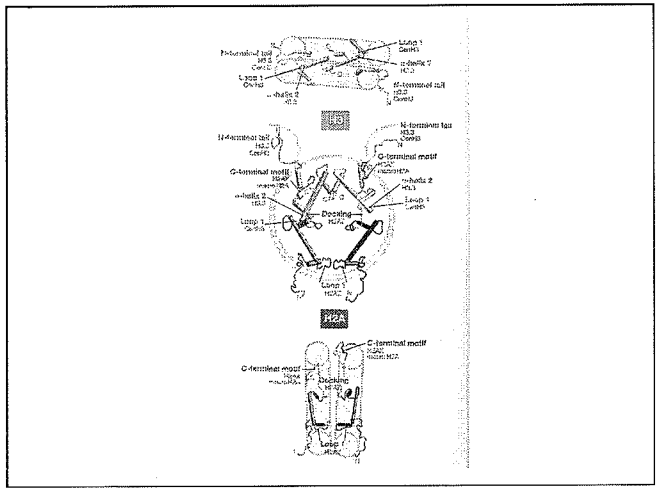
Euchromatin islands in large heterochromatin domains are enriched for CTCF binding and differentially DNA-methylated regions.
Wen B, et al. BMC Genomics. 2012 Oct 26;13:566.

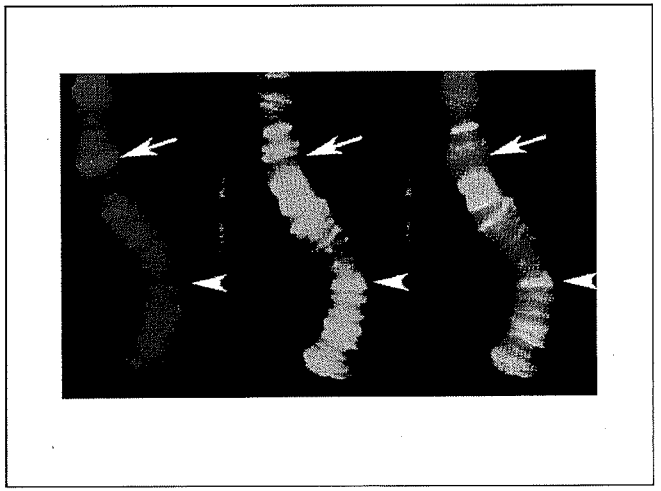


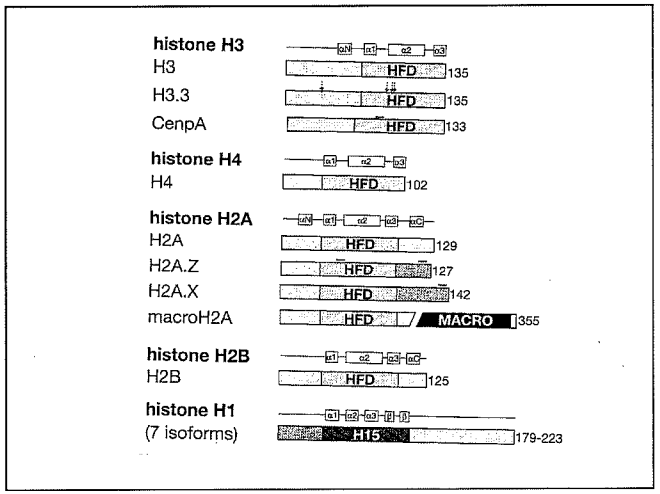
Epigenetics
Histone Modification

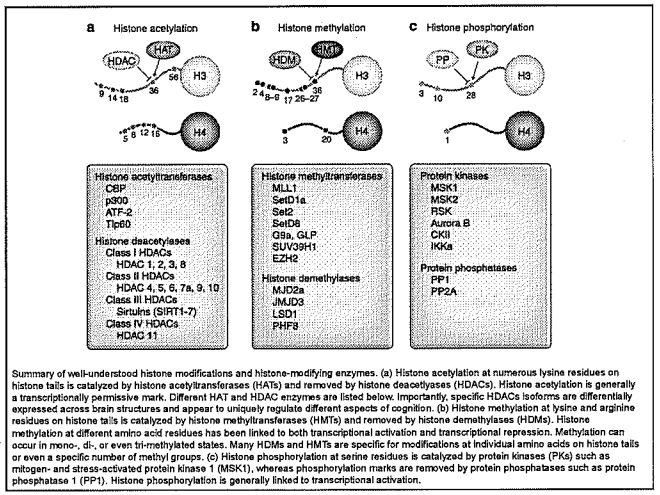




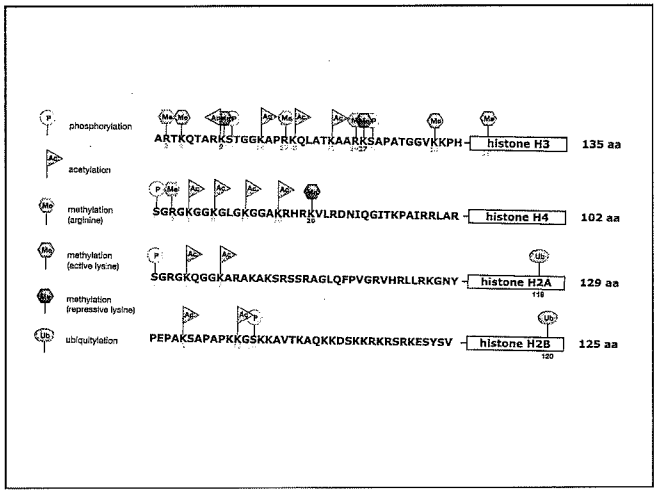


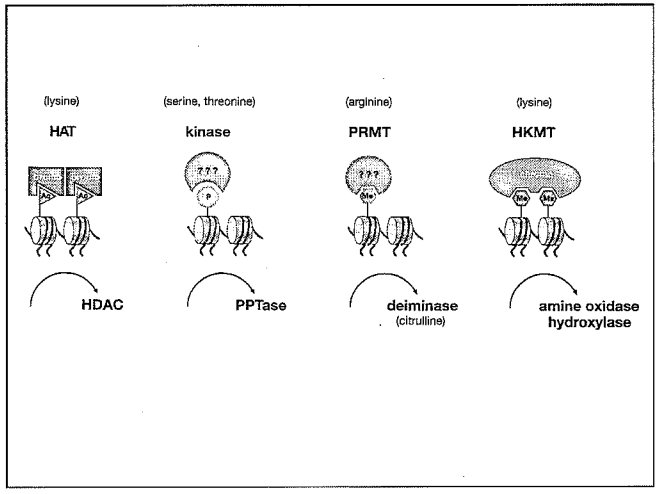


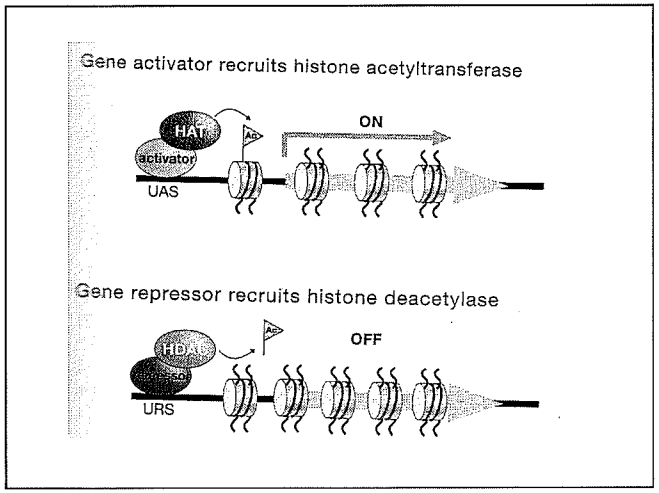


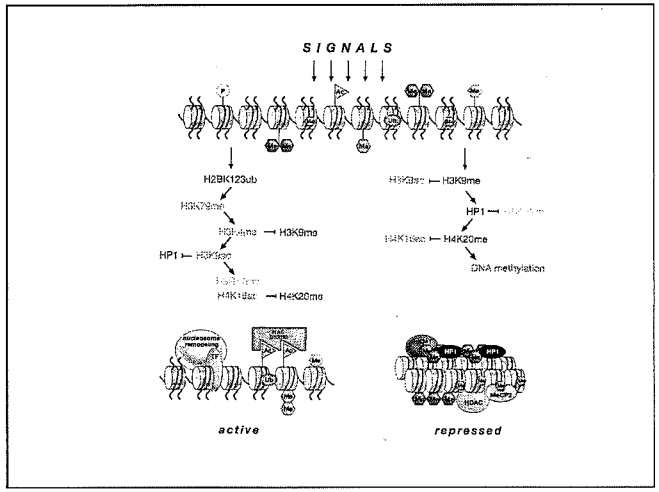


Summary of well-understood histone modifications and histone-modifying enzymes. (a) Histone acetylation at numerous lysine residues on histone tails is catalyzed by histone acetyltransferases (HATs) and removed by histone deacetylases (HDACs). Histone acetylation is generally a transcriptionally permissive mark. Different HAT and HDAC enzymes are listed below. Importantly, specific HDACs isoforms are differentially expressed across brain structures and appear to uniquely regulate different aspects of cognition. (b) Histone methylation at lysine and arginine residues on histone tails is catalyzed by histone methyltransferases (HMTs) and removed by histone demethylases (HDMs). Histone methylation at different amino acid residues has been linked to both transcriptional activation and transcriptional repression. Methylation can occur in mono-, di-, or even tri-methylated states. Many HDMs and HMTs are specific for modifications at individual amino acids on histone tails or even a specific number of methyl groups. (c) Histone phosphorylation at serine residues is catalyzed by protein kinases (PKs) such as mitogen- and stress-activated protein kinase 1 (MSK1), whereas phosphorylation marks are removed by protein phosphatases such as protein phosphatase 1 (PP1). Histone phosphorylation is generally linked to transcriptional activation.

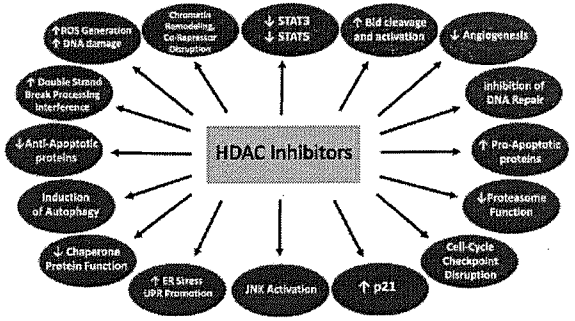

phosphorylation
acetylation
methylation (arginine)
methylation (active lysine)
methylation (repressive lysine)
ubiquitylation
ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPH
histone H3
135 aa
SGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLAR
histone H4
102 aa
SGRGKQGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGNY
histone H2A
129 aa
PEPAKSAPAPKKGSKKAVTKAQKKDSKKRKRSRKESYSV
histone H2B
125 aa


(lysine)
HAT
(serine, threonine)
kinase
(arginine)
PRMT
(lysine)
HKMT
HDAC
PPTase
deiminase
(citrulline)
amine oxidase
hydroxylase


Gene activator recruits histone acetyltransferase
HAT
activator
UAS
ON
Gene repressor recruits histone deacetylase
HDAC
URS
OFF


SIGNALS
H2BK123ub
H3K9me
HP1
H4K20me
DNA methylation
active
repressed

Histone deacetylase inhibitor (HDACI) mechanisms of action: emerging insights.
Bose P, Dai Y, Grant S.
Pharmacol Ther. 2014 Sep;143(3):323-36.



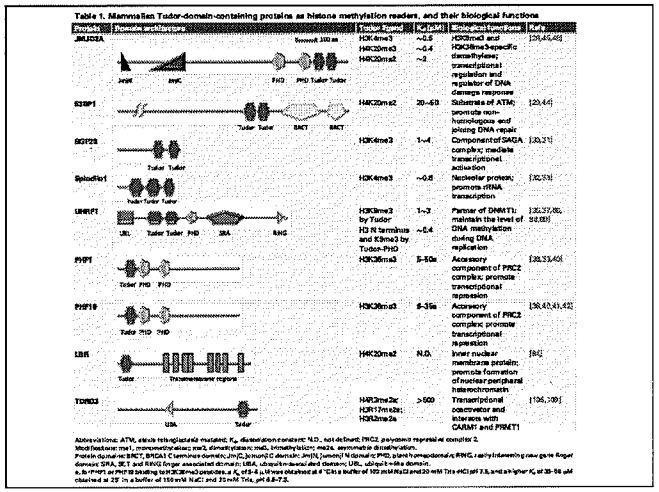
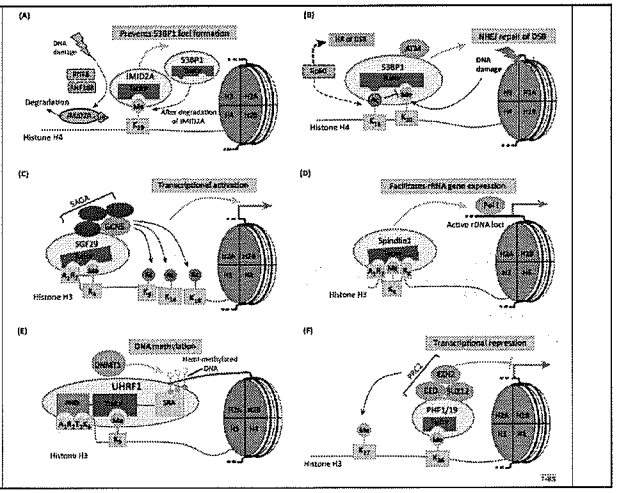
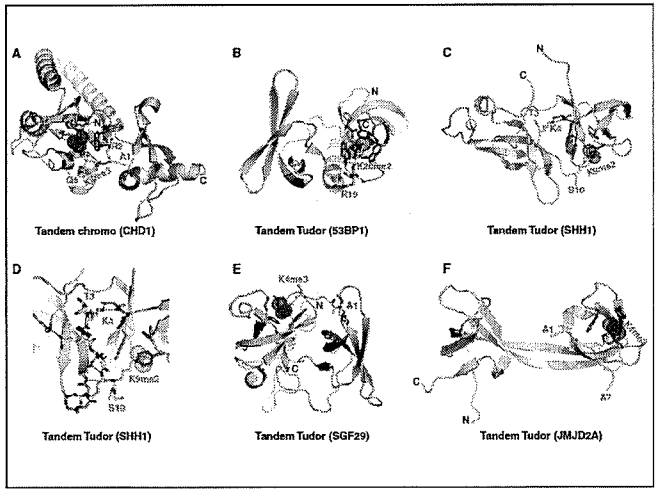


Table 1. Mammalian Tudor-domain-containing proteins as histone methylation readers, and their biological functions

Protein	Preferred methylation	Tudor count	K_d (μM)	Biological function	Ref.
JMJD2A					
	H3K4me3		~0.8	H3K9me3 and H3K20me3-specific demethylase; transcriptional regulation and regulator of DNA damage response	[39,45,46]
	H4K20me3		~0.4		
	H4K20me2		~2		
53BP1	H4K20me2	20–60		Substrate of ATM; promote non-homologous end joining DNA repair	[29,44]
SGF29	H3K4me3	1–4		Component of SAGA complex; mediate transcriptional activation	[30,31]
Spindlin1	H3K4me3	~0.8		Nucleolar protein; promote rRNA transcription	[32,33]
UHRF1	H3K9me3 by Tudor; H3 N terminus and K9me3 by Tudor-PHD	1–2	~0.4	Partner of DNMT1; maintain the level of DNA methylation during DNA replication	[35,37,38, 39,40]
PHF1	H3K36me3	5–50c		Accessory component of PRC2 complex; promote transcriptional repression	[39,53,405]
PHF19	H3K36me3	5–35c		Accessory component of PRC2 complex; promote transcriptional repression	[39,40,41,402]
LBR	H4K20me2	N.D.		Inner nuclear membrane protein; promote formation of nuclear peripheral heterochromatin	[84]
TDRD3	H3R2me2a; H3R17me2a; H3R2me2a	>500		Transcriptional coactivator and interacts with CARM1 and PRMT1	[106,109]





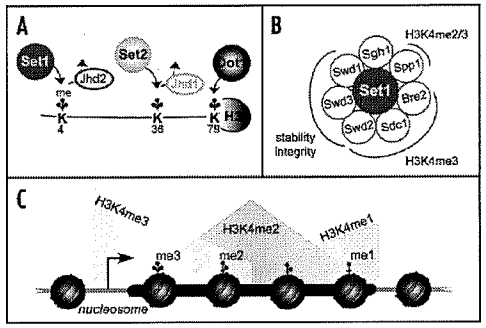


Figure 1. Histone H3 methylation in *S. cerevisiae*. (A) Histone methyl transferases Set1, Set2 and Dot1 mono-, di- or tri-methylate K4, K36 and K79 of histone H3, respectively in *S. cerevisiae*. Jhd2 and Jhd1 histone demethylases remove the H3K4me3 and H3K36me3, respectively. (B) The SET1C consists of eight subunits having distinct regulatory roles on complex activity and integrity. (C) Distribution of H3K4me1, me2 and me3 along a transcribed gene.

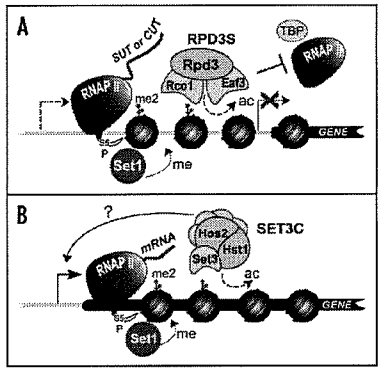


Figure 2. H3K4me2 recruits histone deacetylase complexes in *S. cerevisiae*. (A) H3K4me2 is deposited by upstream cryptic transcription targeting the RPD3S HDAC to repress promoter activity. (B) H3K4me2 recruits the SET3C HDAC to actively transcribed gene contributing to efficient transcription.

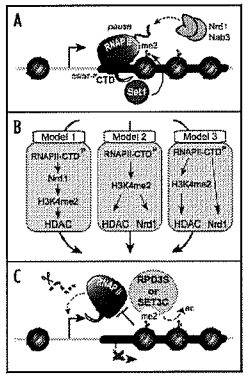
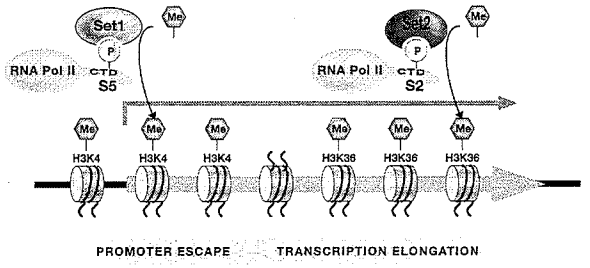
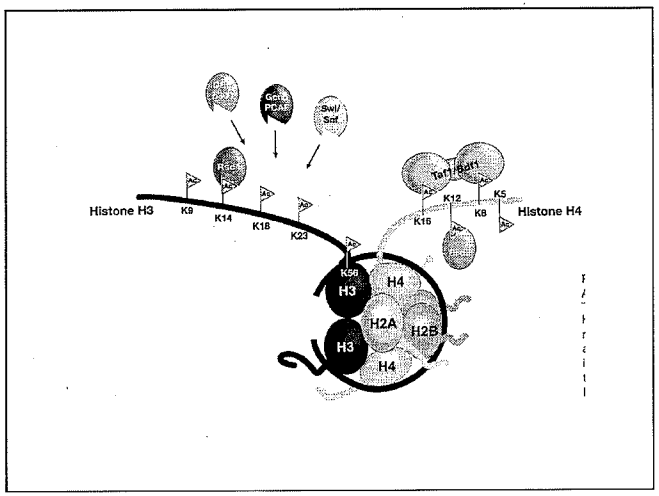


Figure 3. H3K4me2 and Nrd1/Nab3 might be interrelated to control promoter fidelity and RNA quality control. (A) After promoter release the RNAPII-machinery is poised on the transcribed chromatin. (B) Three models (model 1, 2 and 3) represent the possible cascade of events combining H3K4me2, Nrd1/Nab3 complex and RNAPII-CTD phosphorylation and leading to (C) the abortive termination, cleavage of the nascent RNA and recycling of the released RNAPII.





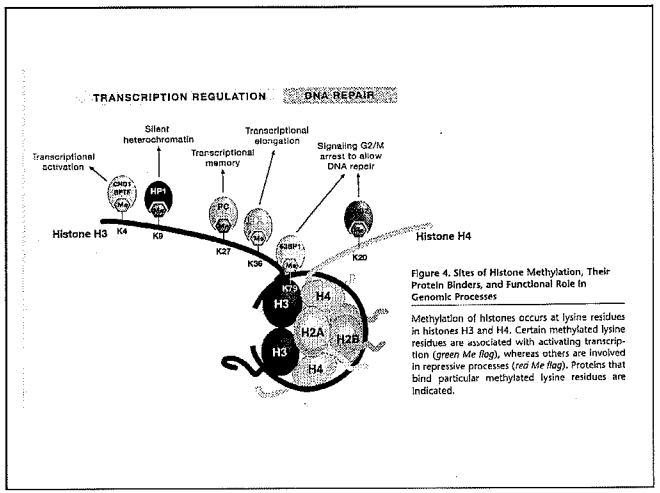


Figure 4. Sites of Histone Methylation, Their Protein Binders, and Functional Role in Genomic Processes

Methylation of histones occurs at lysine residues in histones H3 and H4. Certain methylated lysine residues are associated with activating transcription (*green Me flag*), whereas others are involved in repressive processes (*red Me flag*). Proteins that bind particular methylated lysine residues are indicated.

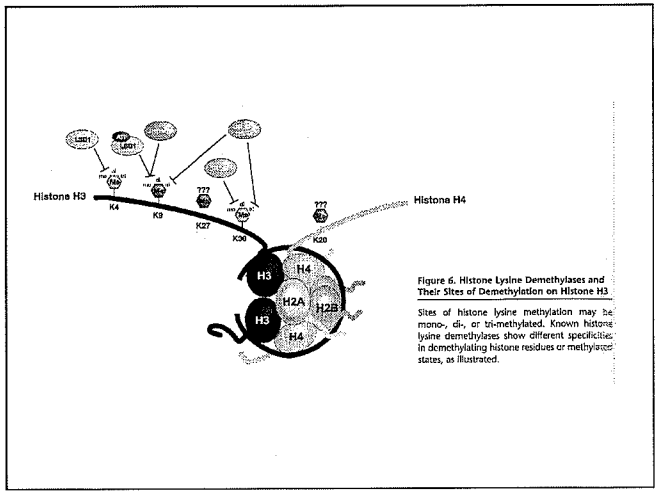


Figure 6. Histone Lysine Demethylases and Their Sites of Demethylation on Histone H3

Sites of histone lysine methylation may be mono-, di-, or tri-methylated. Known histone lysine demethylases show different specificities in demethylating histone residues or methylated states, as illustrated.

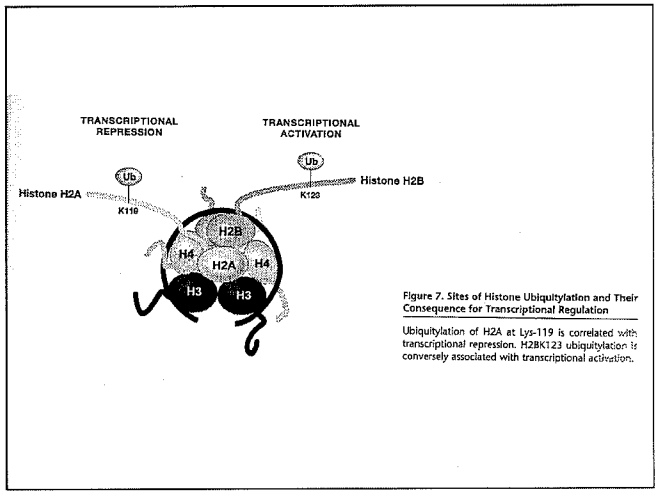
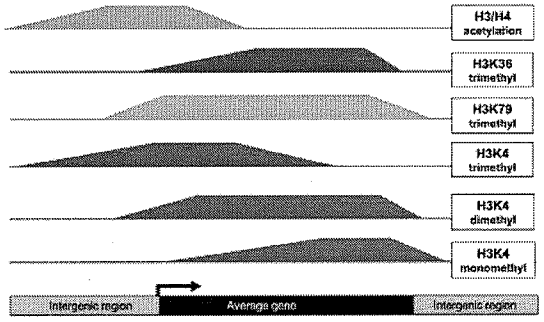
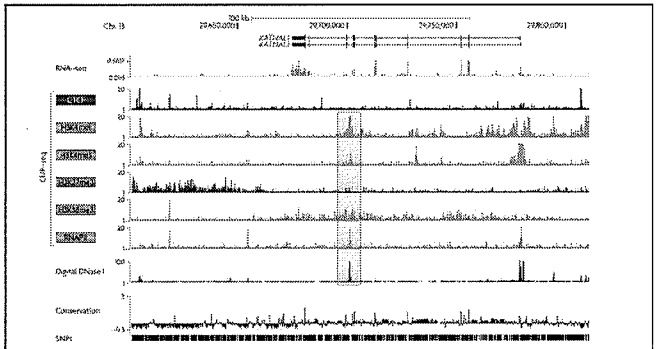


Figure 7. Sites of Histone Ubiquitylation and Their Consequence for Transcriptional Regulation

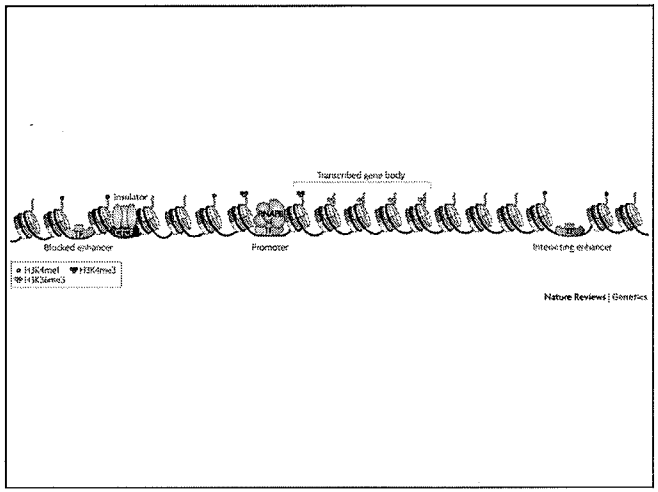
Ubiquitylation of H2A at Lys-119 is correlated with transcriptional repression. H2BK123 ubiquitylation is conversely associated with transcriptional activation.

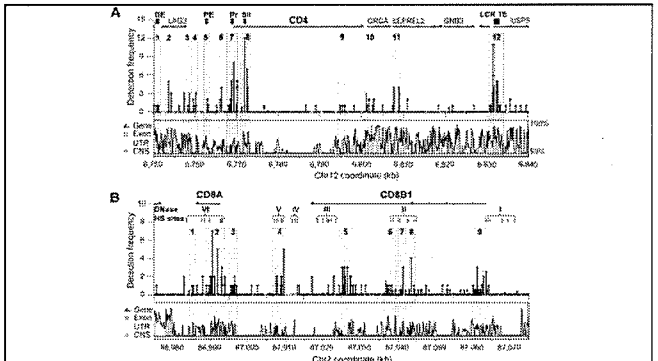


Distribution of Histone Modifications across a Typical Yeast GeneIn their ChIP-chip assay, Pokholok et al. (2005) immunoprecipitated formaldehyde-crosslinked yeast chromatin with antibodies to acetylated histones H3 and H4, and to H3 histone mono-, di-, or trimethylated at lysine 36 (H3K36), lysine 79 (H3K79), and lysine 4 (H3K4). The authors amplified immunoprecipitated and input DNA by ligation-mediated PCR, labeled them with different fluorescent dyes, and mixed and hybridized them to a single DNA microarray. The ratio of both fluorescent dyes on each spot provides a measure of the relative enrichment of the histone modification under investigation. Collectively, this provides information with which to construct a genome-wide map of the consensus distribution of histone modifications in yeast. Each of the distributions shown represents a composite profile derived from several hundred genes.

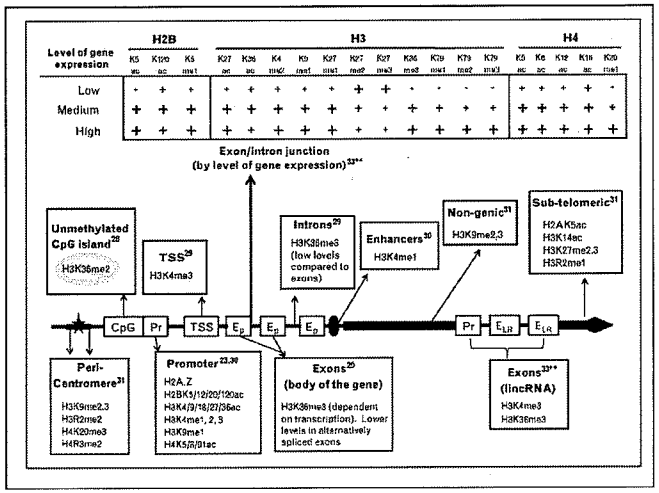


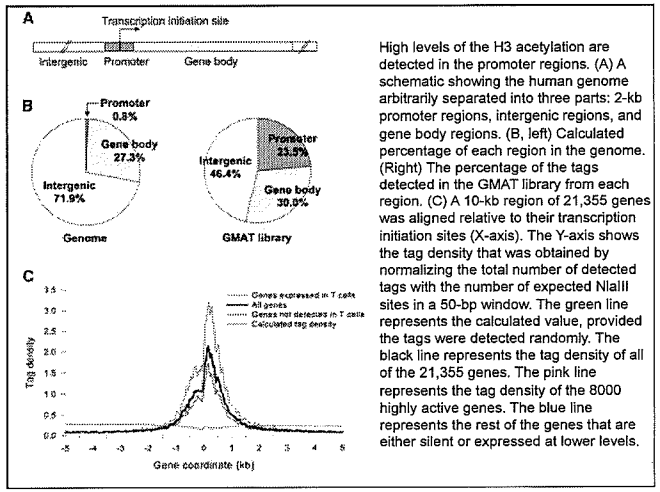
The University of California-Santa Cruz (UCSC) Genome Browser is a tool for viewing genomic data sets. A vast amount of data is available for viewing through this browser. This example from the browser shows numerous data types in K562 cells from the ENCODE Consortium. A random gene was selected — katanin p60 subunit A-like 1(KATNAL1) — that shows several points that can be identified by using this tool. The promoter has a typical chromatin structure (a peak of histone 3 lysine 4 trimethylation (H3K4me3) between the bimodal peaks of H3K4me1), is bound by RNA polymerase II (RNAPII) and is DNase hypersensitive. The gene is transcribed, as indicated by RNA sequencing (RNA-seq) data, as well as H3K36me3 localization. The gene lies between two CCCTC-binding factor (CTCF)-bound sites that could be tested for insulator activity. An intronic H3K4me1 peak (highlighted) predicts an enhancer element, corroborated by the DNase I hypersensitivity site peak. There is a broad repressive domain of H3K27me3 downstream, which could have an open chromatin structure in another cell type.





Colocalization of acetylation islands with known regulatory elements. (A) CD4 locus. The upper panel shows the acetylation data, above which gene positions and known functional regulatory elements are indicated. The lower panel shows the VISTA human and mouse sequence comparison. (DE) Distal enhancer; (PE) proximal enhancer; (Pr) promoter; (Sil) silencer; (LCR) locus control region; (TE) thymocyte enhancer. The acetylation islands colocalized with known regulatory elements are highlighted in pink. The acetylation islands with no known functions are highlighted in green. (B) CD8 locus. The acetylation data and VISTA sequence analysis are shown as in A. The positions of the six clusters of DNase hypersensitive sites (HS) are indicated below the genes.



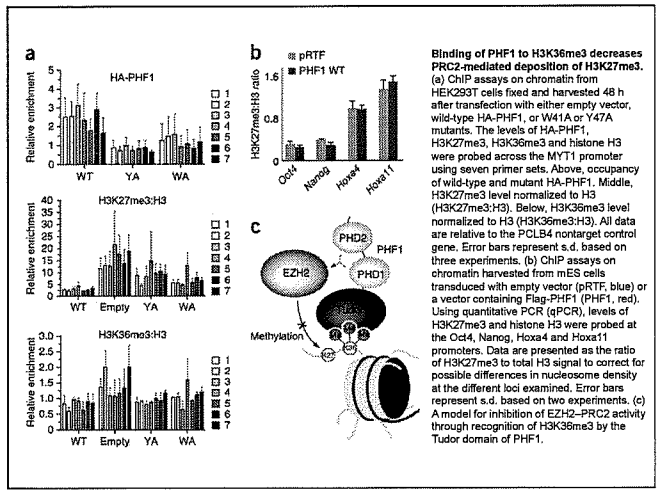


High levels of the H3 acetylation are detected in the promoter regions. (A) A schematic showing the human genome arbitrarily separated into three parts: 2-kb promoter regions, intergenic regions, and gene body regions. (B, left) Calculated percentage of each region in the genome. (Right) The percentage of the tags detected in the GMAT library from each region. (C) A 10-kb region of 21,355 genes was aligned relative to their transcription initiation sites (X-axis). The Y-axis shows the tag density that was obtained by normalizing the total number of detected tags with the number of expected NlaIII sites in a 50-bp window. The green line represents the calculated value, provided the tags were detected randomly. The black line represents the tag density of all of the 21,355 genes. The pink line represents the tag density of the 8000 highly active genes. The blue line represents the rest of the genes that are either silent or expressed at lower levels.

Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1.
Musselman CA, et al. Nat Struct Mol Biol. 2012; 19(12):1266-72.

Abstract

The PHD finger protein 1 (PHF1) is essential in epigenetic regulation and genome maintenance. Here we show that the Tudor domain of human PHF1 binds to histone H3 trimethylated at Lys36 (H3K36me3). We report a 1.9-Å resolution crystal structure of the Tudor domain in complex with H3K36me3 and describe the molecular mechanism of H3K36me3 recognition using NMR. Binding of PHF1 to H3K36me3 inhibits the ability of the Polycomb PRC2 complex to methylate Lys27 of histone H3 in vitro and in vivo. Laser microirradiation data show that PHF1 is transiently recruited to DNA double-strand breaks, and PHF1 mutants impaired in the H3K36me3 interaction exhibit reduced retention at double-strand break sites. Together, our findings suggest that PHF1 can mediate deposition of the repressive H3K27me3 mark and acts as a cofactor in early DNA-damage response.

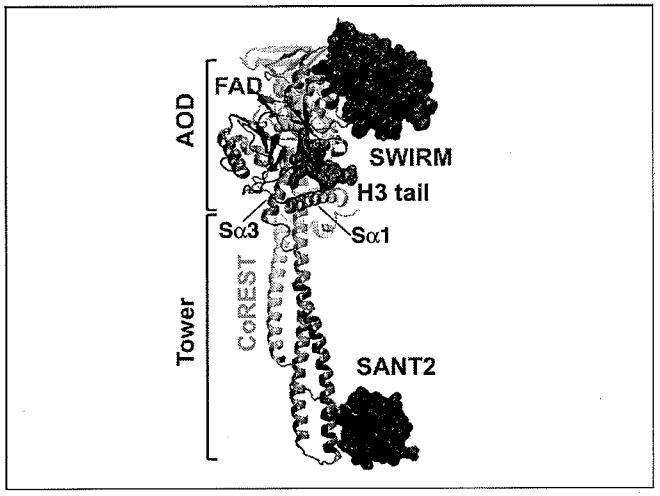


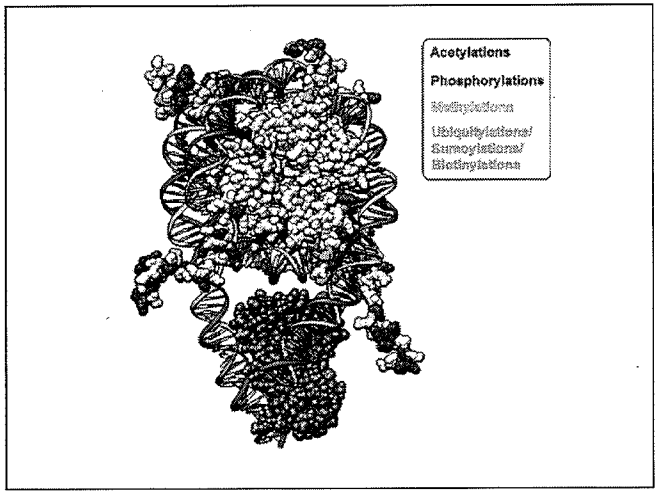
Binding of PHF1 to H3K36me3 decreases PRC2-mediated deposition of H3K27me3. (a) ChIP assays on chromatin from HEK293T cells fixed and harvested 48 h after transfection with either empty vector, wild-type HA-PHF1, or W41A or Y47A mutants. The levels of HA-PHF1, H3K27me3, H3K36me3 and histone H3 were probed across the MYT1 promoter using seven primer sets. Above, occupancy of wild-type and mutant HA-PHF1. Middle, H3K27me3 level normalized to H3 (H3K27me3:H3). Below, H3K36me3 level normalized to H3 (H3K36me3:H3). All data are relative to the PCLB4 nontarget control gene. Error bars represent s.d. based on three experiments. (b) ChIP assays on chromatin harvested from mES cells transduced with empty vector (pRTF, blue) or a vector containing Flag-PHF1 (PHF1, red). Using quantitative PCR (qPCR), levels of H3K27me3 and histone H3 were probed at the Oct4, Nanog, Hoxa4 and Hoxa11 promoters. Data are presented as the ratio of H3K27me3 to total H3 signal to correct for possible differences in nucleosome density at the different loci examined. Error bars represent s.d. based on two experiments. (c) A model for inhibition of EZH2–PRC2 activity through recognition of H3K36me3 by the Tudor domain of PHF1.

LSD1/CoREST is an allosteric nanoscale clamp regulated by H3-histone-tail molecular recognition.
Baron R, Vellore NA. Proc Natl Acad Sci U S A. 2012 Jul 31;109(31):12509-14.

Abstract

The complex of lysine-specific demethylase-1 (LSD1/KDM1A) with its corepressor protein CoREST is an exceptionally relevant target for epigenetic drugs. Here, we provide insight into the local and global changes of LSD1/CoREST conformational dynamics that occur upon H3 binding on the basis of a total cumulative time of one microsecond molecular dynamics simulation. The LSD1/CoREST complex functions as an allosteric nanoscale-binding clamp, which is regulated by substrate binding. In the unbound state, LSD1/CoREST reversibly visits clamp states that are more open or significantly more closed compared with the available X-ray crystal structures. The Lys triad of residues Lys355, Lys357, and Lys359 gates the entrance of the H3 pocket. H3 binding shifts the pocket breathing dynamics toward open, higher-volume states while reducing the overall flexibility of the LSD1/CoREST nanoscale clamp. We show that the H3 pocket is an allosteric site for the regulation of the rotation of the amino oxidase domain with respect to the Tower domain. The allosteric mechanism relies on the specific reduction of nanoscale domain rotation upon local H3-tail binding. Instead, clamp opening/closing motions that do not involve domain rotation only reduce in amplitude yet are dominant in the bound state. Overall, our data suggest that the H3 binding pocket is a central target site to (i) switch off LSD1 amino oxidase activity, thus H3-tail demethylation; (ii) block the competitive binding of transcription factors; and (iii) prevent chromatin anchoring to LSD1/CoREST. This study underscores the importance of receptor flexibility for future epigenetic drug discovery.

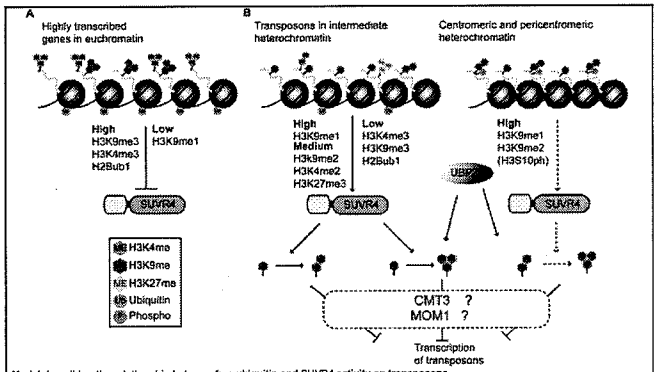




The SUVR4 histone lysine methyltransferase binds ubiquitin and converts H3K9me1 to H3K9me3 on transposon chromatin in Arabidopsis.
Veiseth SV, et al. PLoS Genet. 2011 Mar;7(3):e1001325.

Abstract

Chromatin structure and gene expression are regulated by posttranslational modifications (PTMs) on the N-terminal tails of histones. Mono-, di-, or trimethylation of lysine residues by histone lysine methyltransferases (HKMTases) can have activating or repressive functions depending on the position and context of the modified lysine. In Arabidopsis, trimethylation of lysine 9 on histone H3 (H3K9me3) is mainly associated with euchromatin and transcribed genes, although low levels of this mark are also detected at transposons and repeat sequences. Besides the evolutionarily conserved SET domain which is responsible for enzyme activity, most HKMTases also contain additional domains which enable them to respond to other PTMs or cellular signals. Here we show that the N-terminal WIYLD domain of the Arabidopsis SUVR4 HKMTase binds ubiquitin and that the SUVR4 product specificity shifts from di- to trimethylation in the presence of free ubiquitin, enabling conversion of H3K9me1 to H3K9me3 in vitro. Chromatin immunoprecipitation and immunocytological analysis showed that SUVR4 in vivo specifically converts H3K9me1 to H3K9me3 at transposons and pseudogenes and has a locus-specific repressive effect on the expression of such elements. Bisulfite sequencing indicates that this repression involves both DNA methylation-dependent and -independent mechanisms. Transcribed genes with high endogenous levels of H3K4me3, H3K9me3, and H2Bub1, but low H3K9me1, are generally unaffected by SUVR4 activity. Our results imply that SUVR4 is involved in the epigenetic defense mechanism by trimethylating H3K9 to suppress potentially harmful transposon activity.



Model describing the relationship between free ubiquitin and SUVR4 activity on transposons.
(A) SUVR4 is repressed by H3K4me3 in vitro, and has no activity on genes with high H3K4me3, H3K9me3, H2Bub1 and a low level of H3K9me1. (B) SUVR4's preference for heterochromatic transposons intercalated within euchromatin is maintained by its specificity for H3K9me1 which is highly enriched at transposons, and its repression by activating marks like H3K4me3. The deubiquitinase UBP26 regulates H3K9me2/me3 at the same targets as SUVR4, and might produce free ubiquitin that stimulates the H3K9me2/me3 activity of SUVR4 at target transposons. Although SUVR4 normally is repressed by H3K9me2 and H3S10ph which is high in pericentric heterochromatin, these regions may be targets for SUVR4 activity when ubiquitin levels are high. Since the transposons also contain a medium level of H3K27me3 in addition to H3K9me3, this could possibly create a binding site for CMT3 in order to repress transcription in a DNA methylation-dependent manner at some transposons. At other transposons, transcription may be repressed in a DNA methylation- independent manner by the MOM transcriptional repressor (See text for details).

Spring 2017 – Epigenetics and Systems Biology
Lecture Outline (Systems Biology)
Michael K. Skinner – Biol 476/576
Weeks 5, 6 and 7 (February 7, 14 and 21)

Epigenetics (History / Molecular Processes/ Genomics)

- Definitions and History
- Molecular Factors (DNA Methylation, Histone Modification, Chromatin Structure, ncRNA)
- Epigenetics Technology and Genomics

Required Reading

Holliday R. Epigenetics: a historical overview. Epigenetics. 2006 Apr-Jun;1(2):76-80.

Books (Reserve in Library)

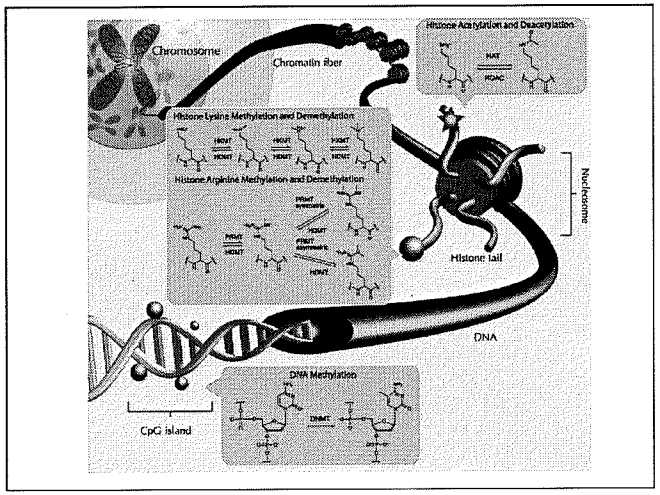
Kevin V. Morris (2012) Non-coding RNAs and Epigenetic Regulation of Gene Expression: Drivers of Natural Selection. Caister Academic Press.

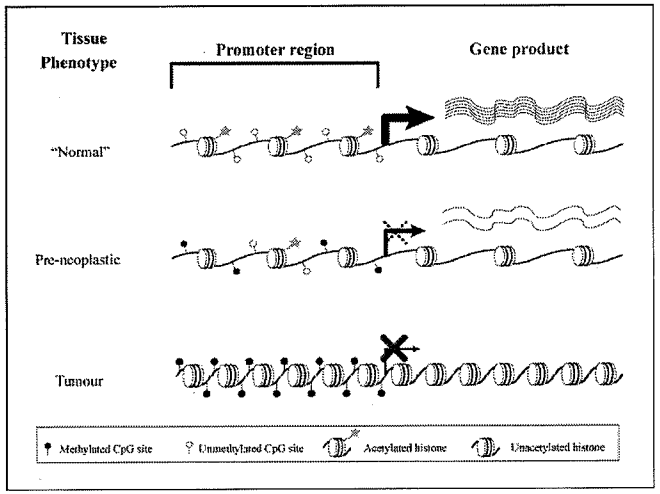
Russo, V.E.A., Martienssen, A.and Riggs, A.D. (eds.). 1996. Epigenetic Mechanisms of Gene Regulation. Cold Spring Harbor Press. Cold Spring Harbor.

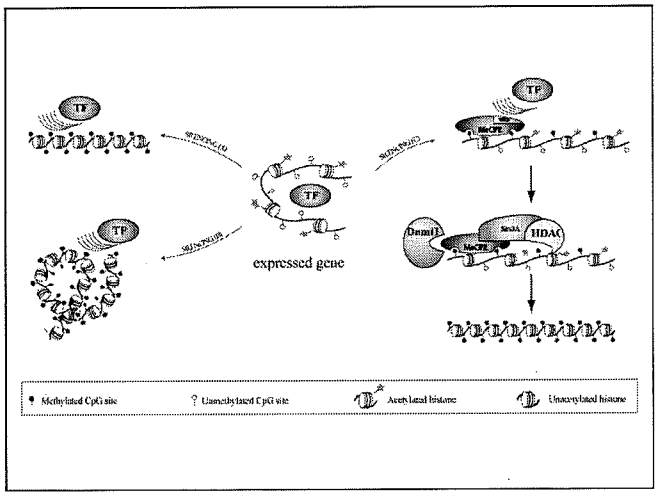
Allis, C.D., Jenuwein, T. & Reinberg, D., Eds. (2007). Epigenetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

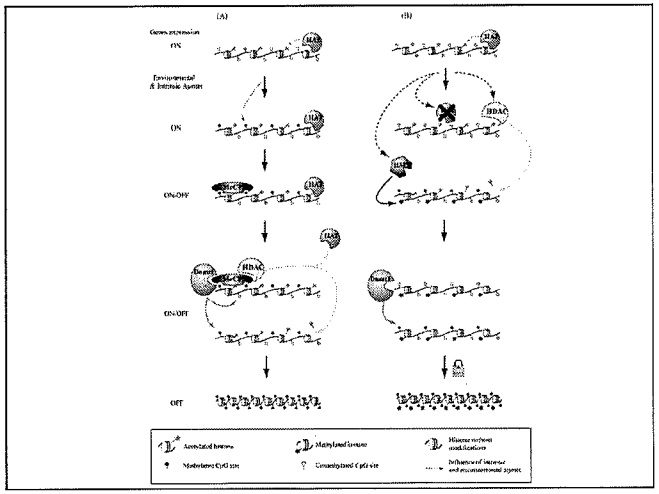
Jeanteur, P. EPIGENETICS AND CHROMATIN. Progress in Molecular and Subcellular Biology, 2005, Volume 38, 151-167, DOI: 10.1007/3-540-27310-7_6

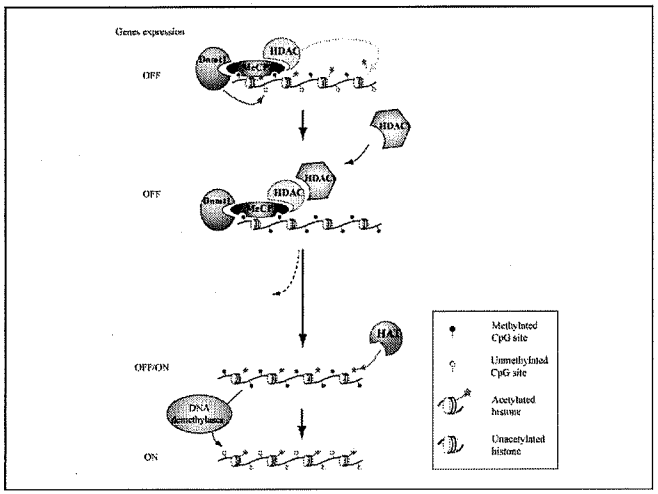
Epigenetics
DNA Methylation & Histone Modification Integration

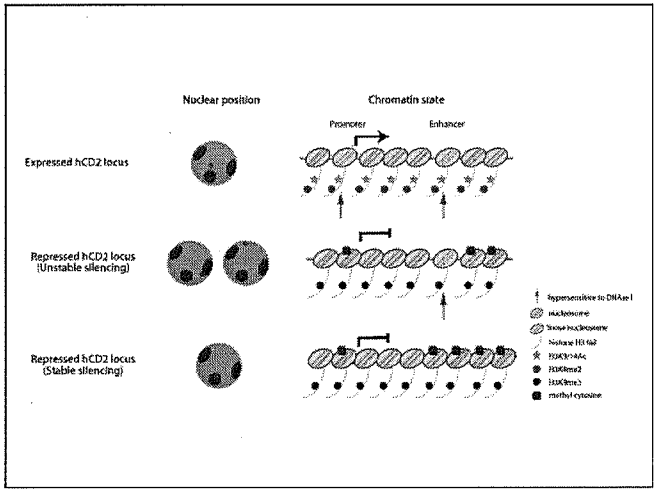

Chromosome
Chromatin fiber
Histone Acetylation and Deacetylation
HAT
HDAC
Histone Lysine Methylation and Demethylation
HKMT
Histone Arginine Methylation and Demethylation
PRMT
Nucleosome
Histone tail
DNA
DNA Methylation
DNMT
CpG island

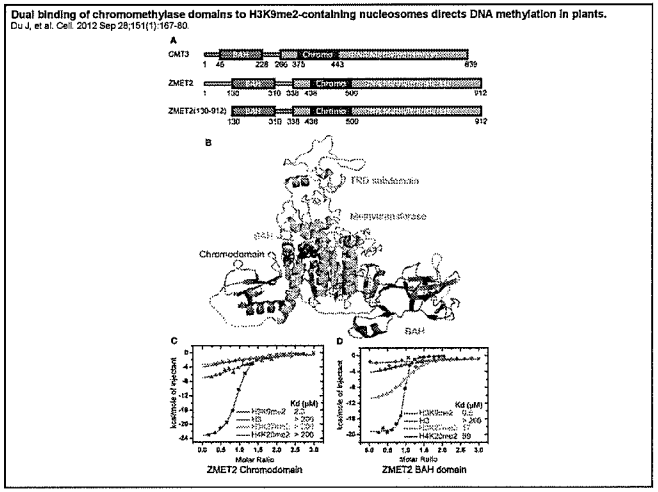

Tissue Phenotype
Promoter region
Gene product
"Normal"
Pre-neoplastic
Tumour
Methylated CpG site
Unmethylated CpG site
Acetylated histone
Unacetylated histone

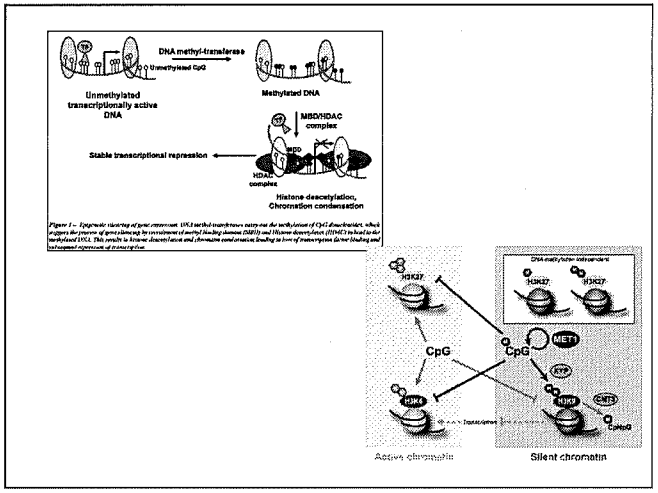

TF
expressed gene
DnmT
MeCP2
SinA
HDAC
Methylated CpG site
Unmethylated CpG site
Acetylated histone
Unacetylated histone

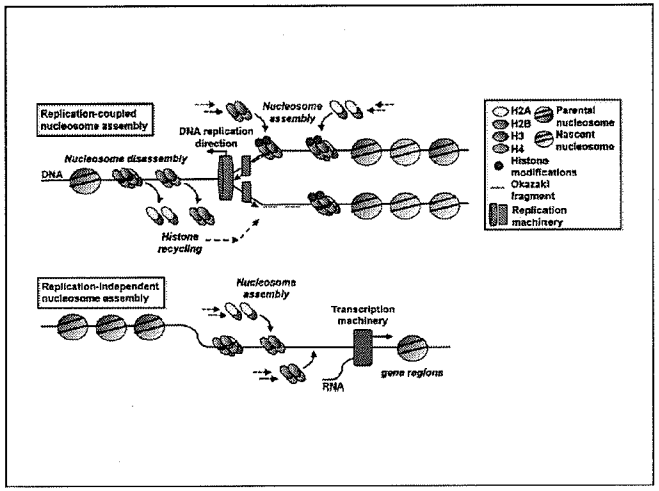

(A)
(B)
Gene expression
ON
Environmental & Intrinsic Agents
ON
ON-OFF
HDAC
ON-OFF
OFF


Genes expression
OFF
HDAC
OmaH
MeCP
OFF
HDAC
OmaH
MeCP
OFF/ON
HAT
DNA demethylase
ON
Methylated CpG site
Unmethylated CpG site
Acetylated histone
Unacetylated histone

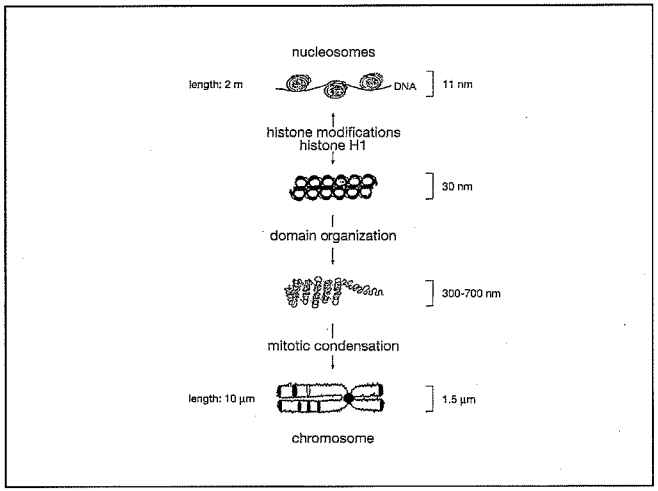

Nuclear position
Chromatin state
Promoter
Enhancer
Expressed hCD2 locus
Repressed hCD2 locus (Unstable silencing)
Repressed hCD2 locus (Stable silencing)

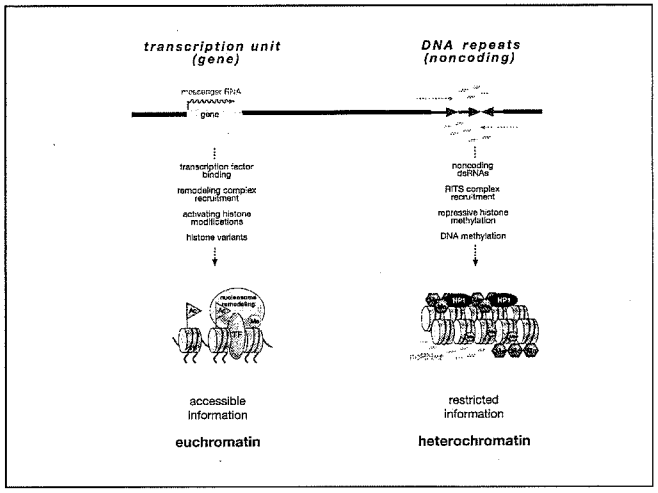

Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants.
Du J, et al. Cell. 2012 Sep 28;151(1):167-80.
A
CMT3
ZMET2
ZMET2(130-912)
B
TRD subdomain
Chromodomain
BAH
C
D
Molar Ratio
ZMET2 Chromodomain
ZMET2 BAH domain


DNA methyl-transferase
Unmethylated CpG
Unmethylated transcriptionally active DNA
Methylated DNA
MBD/HDAC complex
Stable transcriptional repression
HDAC complex
Histone deacetylation, Chromatin condensation
H3K27
CpG
MET1
KYP
H3K4
CMT3
Active chromatin
Silent chromatin



Epigenetics
Chromatin Structure



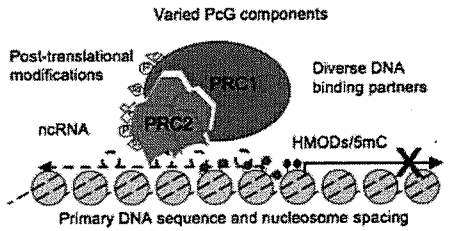


Complexity of polycomb group function: diverse mechanisms of target specificity.
Trask MC, Mager J. J Cell Physiol. 2011 Jul;226(7):1719-21.

Abstract

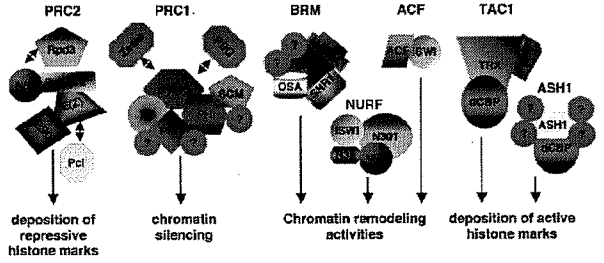
Epigenetic regulation of gene expression has become relevant to nearly all areas of biomedical research. The emergence of technologies that allow for examination of the epigenome combined with identification of key protein complexes that mediate the myriad chromatin modifications that occur have greatly enhanced the versatility and efficacy of tools with which to study normal development and disease states. The evolutionarily conserved polycomb group genes (PcG) have been identified as a predominant mechanism by which gene silencing occurs during development, differentiation, and disease. While molecular events that target PcG complexes have been well defined in some non-vertebrate models, the details of locus specificity and functional diversity of mammalian PcG proteins have not yet unresolved. Here we discuss recent findings that offer novel mechanistic events and add complexity to our understanding of PcG function in vertebrates.

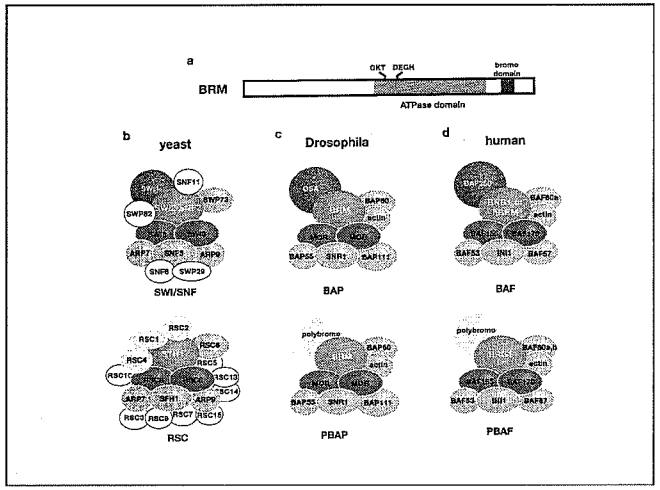
Mechanisms influencing PcG target recognition

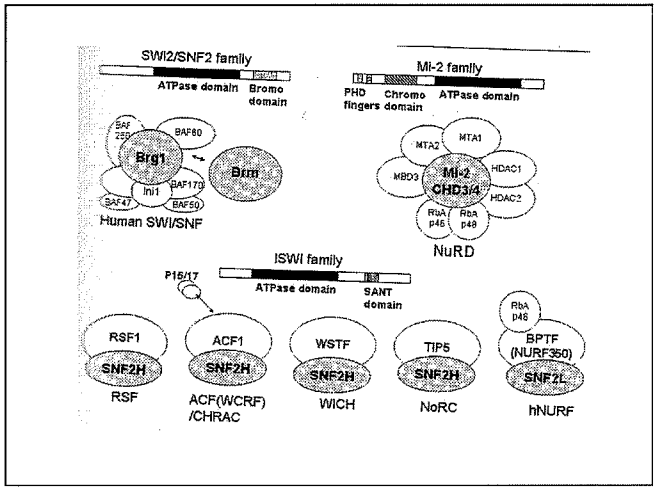


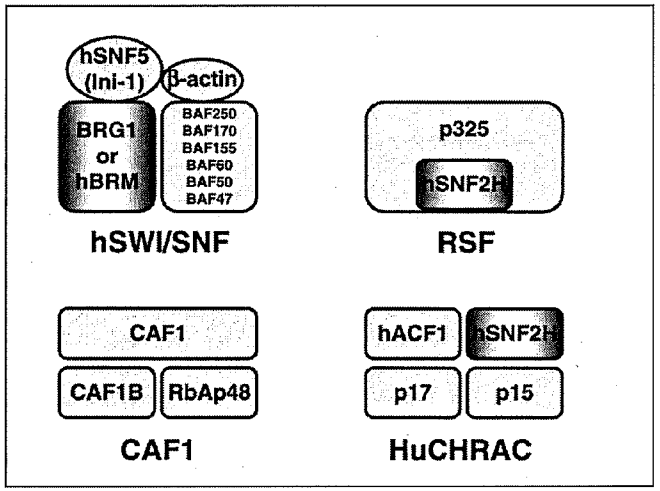
Summary of diverse mechanisms that may alter PcG binding to target DNA sequences, augment protein-protein interactions and/or modify enzymatic activity.

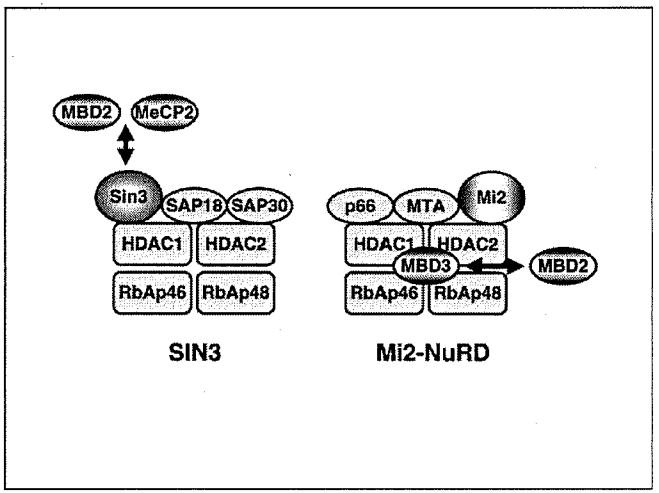
PcG complexes		trxG complexes	
PRC1	PC PH PSC dRING SCM	TAC1	TRX dCBP SBF1
		ASH1	ASH1 dCBP ...
PRC2	E(Z) ESC Su(Z)12 NURF-55	ASH2	ASH2 ...
PHO/PHOL Pipsqueak Grainyhead	DNA-binding PcG/trxG recruiters		Zeste GAF
	PcG/trxG cofactors		
Asx E(Pc) Su(Z)2 Corto Lola/Batman PCL Domino dMi2	ACF ISWI ACF	BRM BRM MOIRA OSA SNR1	Kismet Tonalli Skuld Kohtalo NURF NURF-301 ISWI NURF-55 NURF-38

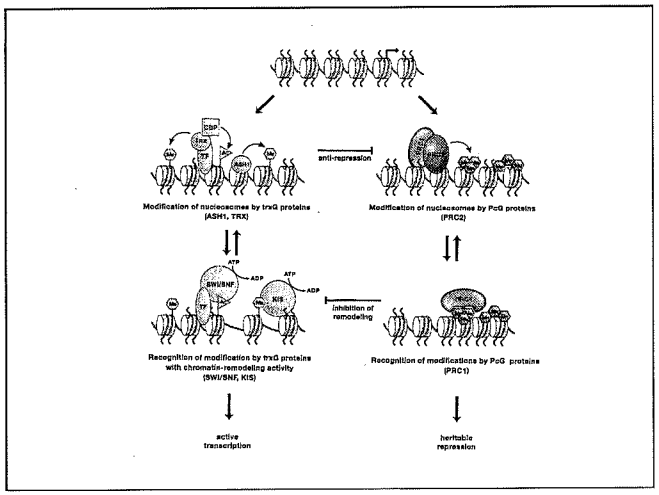


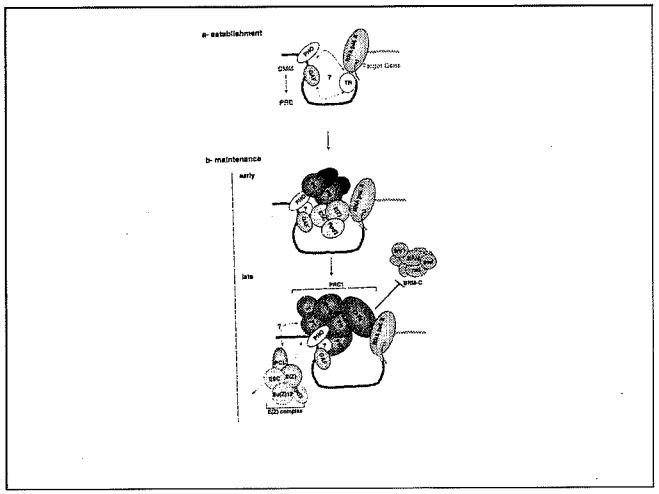

a
BRM
GKT
DEGH
bromo domain
ATPase domain
b
yeast
c
Drosophila
d
human
SWI/SNF
BAP
BAF
RSC
PBAP
PBAF

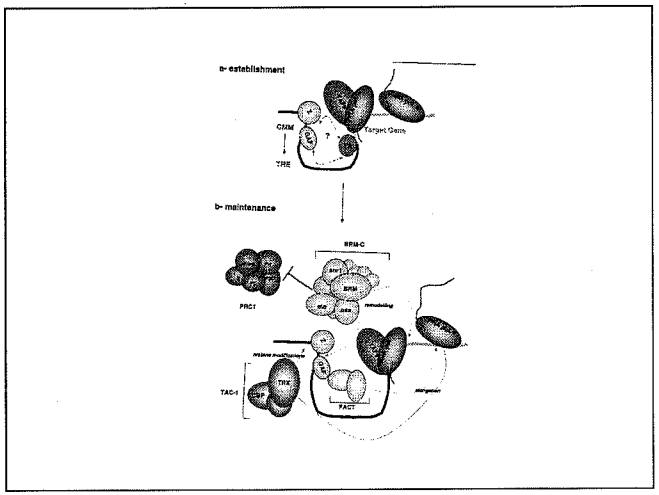

SWI2/SNF2 family
ATPase domain
Bromo domain
Mi-2 family
PHD fingers
Chromo domain
ATPase domain
Brg1
Brm
Ini1
Human SWI/SNF
Mi-2 CHD3/4
NuRD
ISWI family
ATPase domain
SANT domain
RSF1
ACF1
WSTF
TIP5
BPTF (NURF350)
SNF2H
SNF2L
RSF
ACF(WCRF) /CHRAC
WICH
NoRC
hNURF

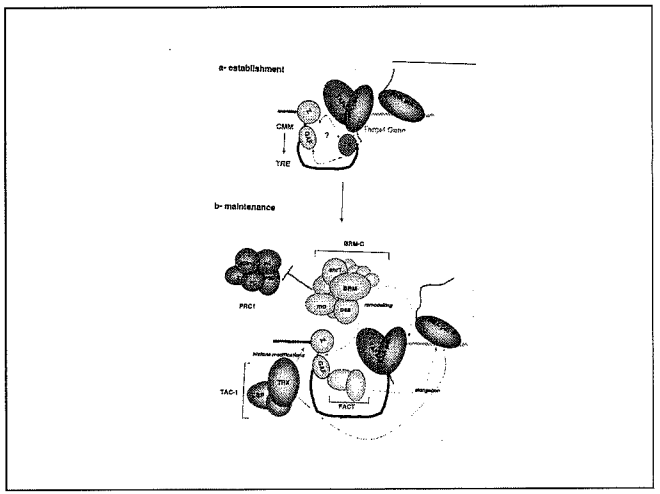

hSNF5 (Ini-1)
β-actin
BRG1 or hBRM
BAF250
BAF170
BAF155
BAF60
BAF50
BAF47
hSWI/SNF
p325
hSNF2H
RSF
CAF1
CAF1B
RbAp48
CAF1
hACF1
hSNF2H
p17
p15
HuCHRAC

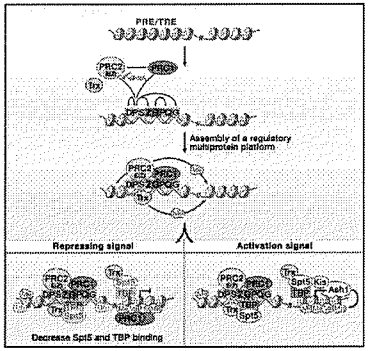

MBD2
MeCP2
Sin3
SAP18
SAP30
HDAC1
HDAC2
RbAp46
RbAp48
SIN3
p66
MTA
Mi2
MBD3
Mi2-NuRD


anti-repression
Modification of nucleosomes by trxG proteins (ASH1, TRX)
Modification of nucleosomes by PcG proteins (PRC2)
SWI/SNF
KIS
inhibition of remodeling
Recognition of modification by trxG proteins with chromatin-remodeling activity (SWI/SNF, KIS)
Recognition of modifications by PcG proteins (PRC1)
active transcription
heritable repression

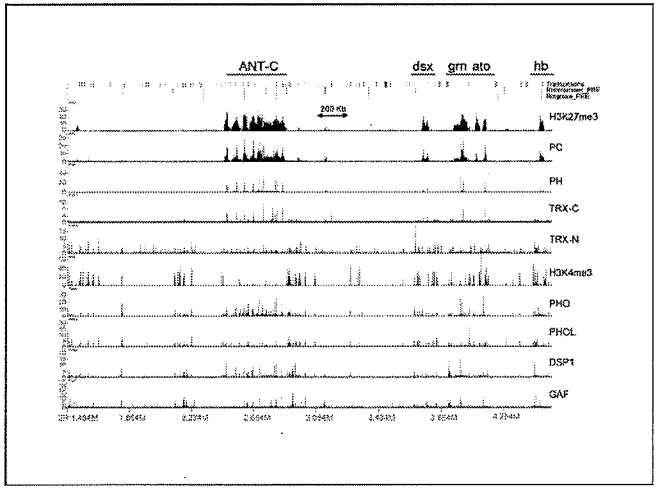

a- establishment
CNM
PRE
Target Genes
b- maintenance
early
late
PRC1
RING
E2c complex

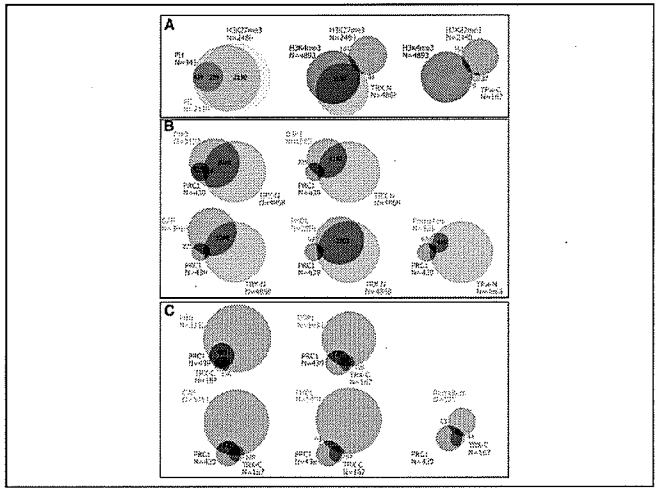

a- establishment
CNM
TRE
Target Genes
b- maintenance
RRM-C
PRC1
TAC-1
TRX
PACT

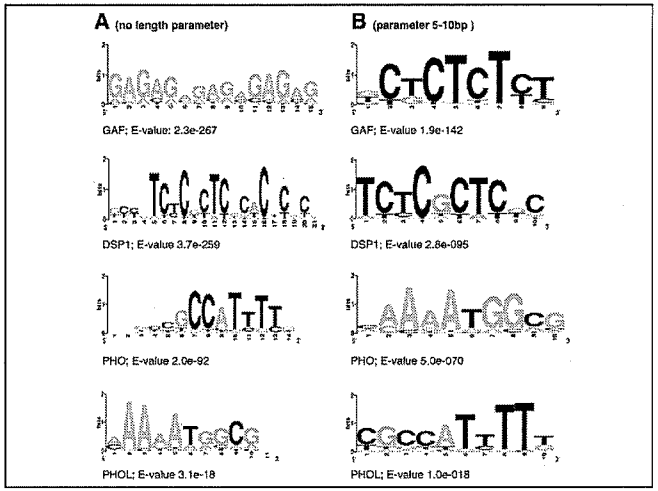

a- establishment
CNM
TRE
Target Genes
b- maintenance
RRM-C
PRC1
TAC-1
TRX
PACT

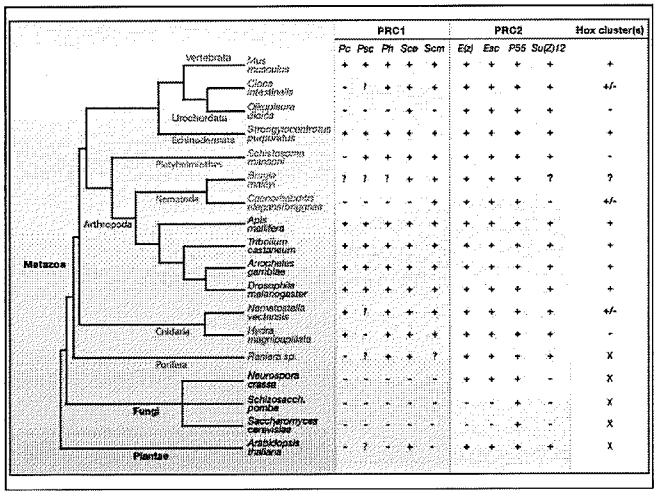


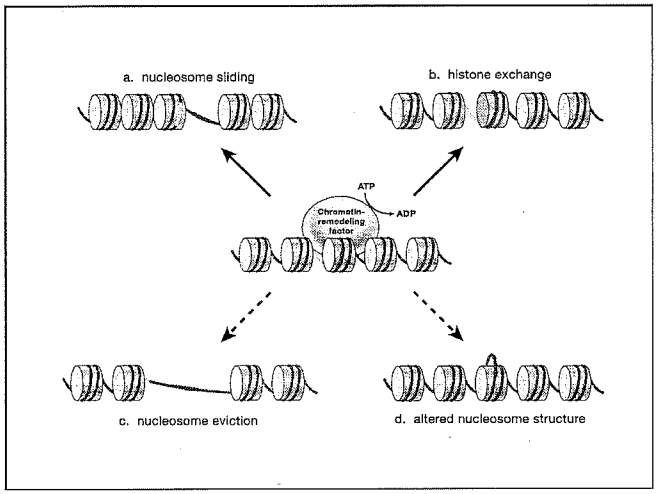
PREs and TREs as Molecular Binding PlatformsMultiple DNA binding proteins like Pho (P), Dsp1 (D), SP1/KLF (S), Zeste (Z), GAGA factor (G), Pipsqueak (PQ), and Grainyhead (G) recruit PcG complexes to PREs. The recruitment of either PcG or TrxG proteins to the PRE sequence does not depend on the activation status of the gene. Moreover, general transcription and elongation factors, such as TBP and Spt5, are constitutively bound to the PRE. A developmental signal then determines whether the PRE mediates gene activation or gene repression, which are accompanied by trimethylation (Me) of histone H3 on lysine 4 or lysine 27, respectively. Upon activation, Kismet (Kis), a protein facilitating elongation, and Ash1, leading to local H3K4 methylation, are recruited to the promoter region. Small RNAs (red ladder) may also contribute to PcG protein recruitment.

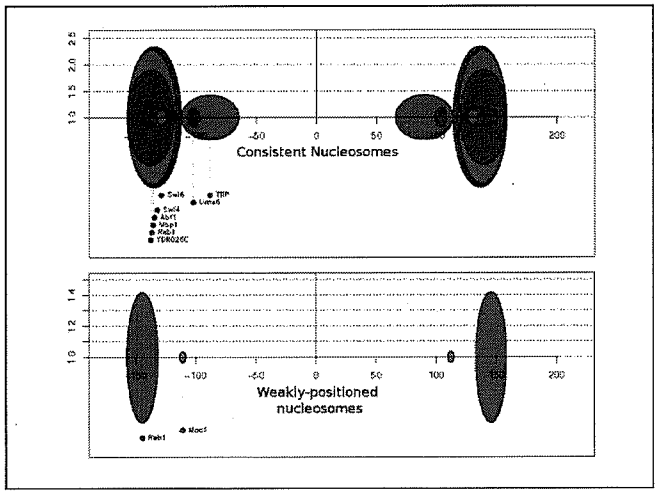


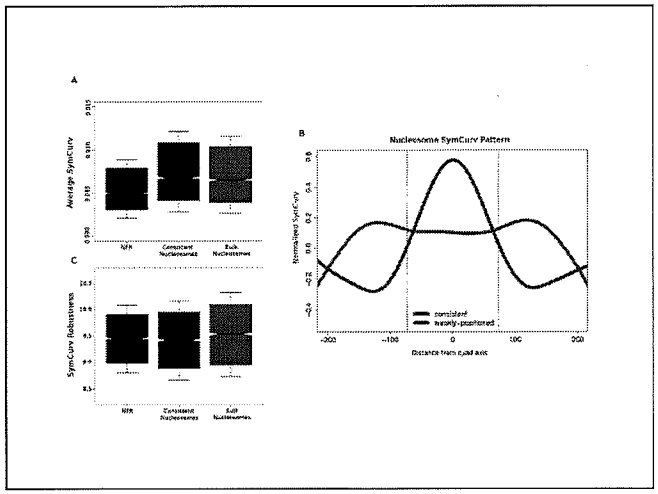


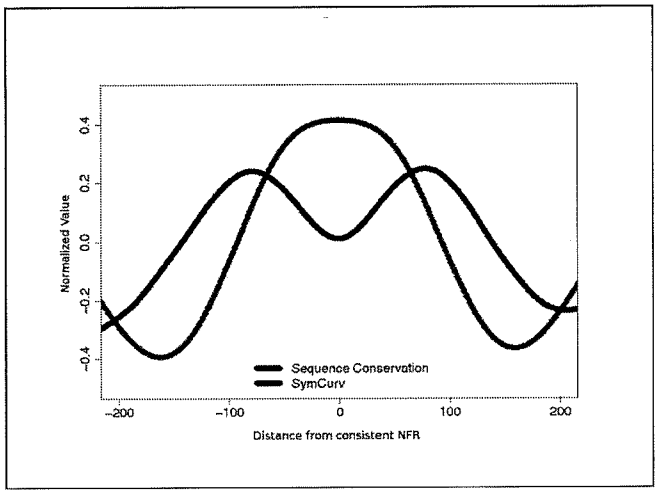


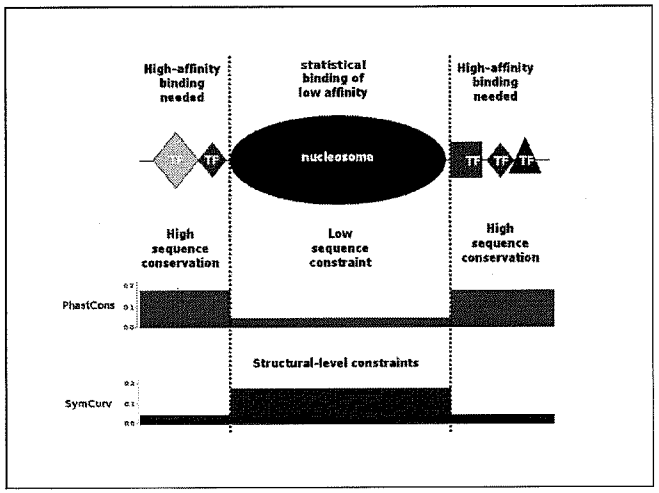

PRC1
PRC2
Hox cluster(s)
Pc Psc Ph Scm
Esc E(z) P55 Su(z)12
Vertebrata
Urochordata
Echinodermata
Platyhelminthes
Nematoda
Arthropoda
Cnidaria
Porifera
Fungi
Plantae
Metazoa

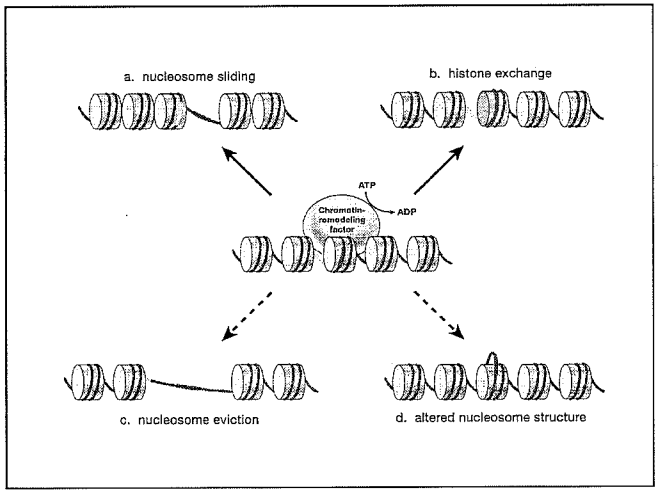

a. nucleosome sliding
b. histone exchange
ATP
ADP
Chromatin-remodeling factor
c. nucleosome eviction
d. altered nucleosome structure

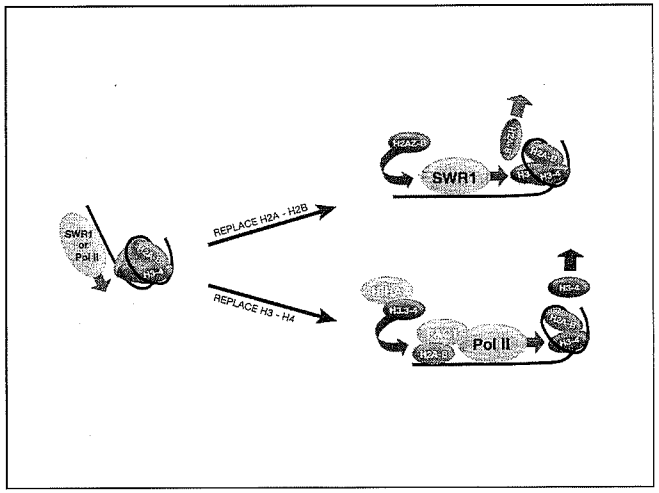

Consistent Nucleosomes
Weakly-positioned nucleosomes

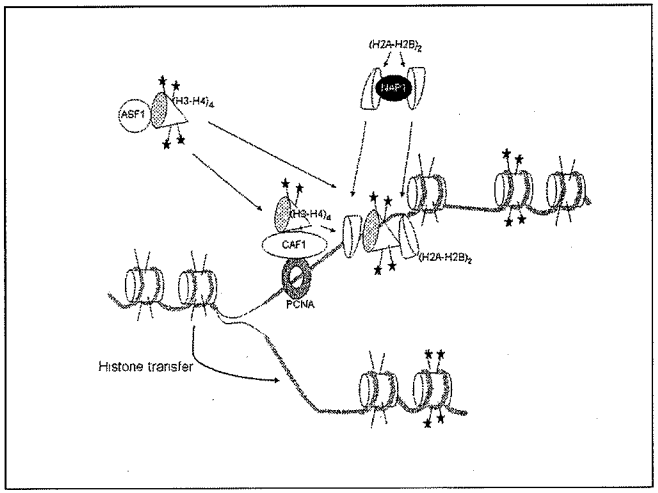

A
B
C
Nucleosome SymCurv Pattern
consistent
weakly-positioned

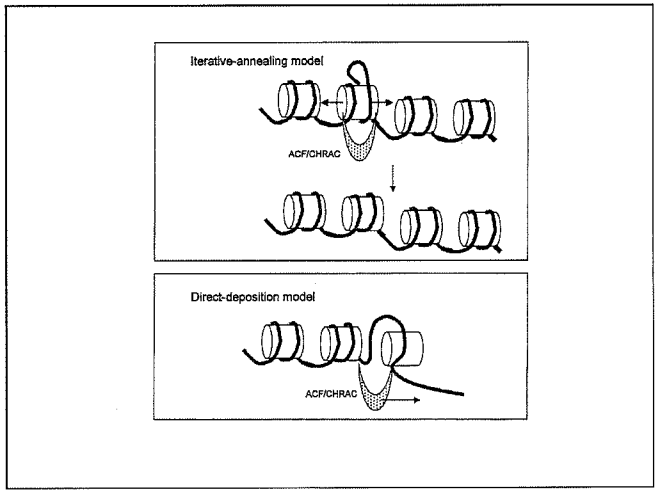

Normalized Value
0.4
0.2
0.0
−0.2
−0.4
Sequence Conservation
SymCurv
−200
−100
0
100
200
Distance from consistent NFR


High-affinity binding needed
statistical binding of low affinity
High-affinity binding needed
nucleosome
TF
TF
TF
TF
High sequence conservation
Low sequence constraint
High sequence conservation
PhastCons
Structural-level constraints
SymCurv

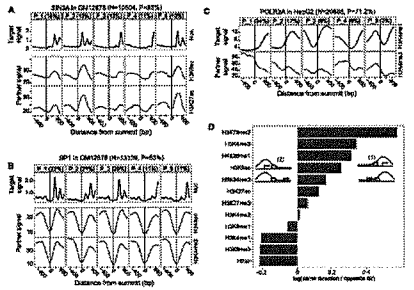

a. nucleosome sliding
b. histone exchange
ATP
ADP
Chromatin-remodeling factor
c. nucleosome eviction
d. altered nucleosome structure


SWR1 or Pol II
REPLACE H2A - H2B
REPLACE H3 - H4
SWR1
Pol II





Ubiquitous heterogeneity and asymmetry of the chromatin environment at regulatory elements.
Kundaje A, et al. Genome Res. 2012 Sep;22(9):1735-47.

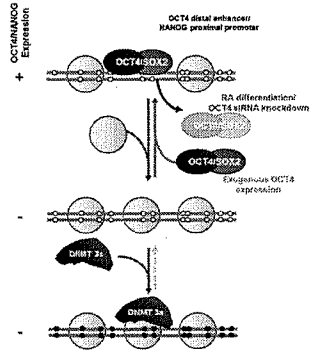


(A,B) The top row shows the most prevalent nucleosome positioning clusters around SIN3A and SP1 sites, respectively, in GM12878. The remaining rows show the signal of histone modifications, averaged over all sites in the corresponding clusters. TSS-proximal TFs, such as SIN3A, exhibit correlated nucleosome positioning and histone modification patterns. Such correlations, however, are not evident for TFs that tend to bind more distally from TSSs (e.g., SP1). (C) Clusters of H3K4me1 signal around POLR2A sites in HepG2 and the corresponding H3K4me3 signal. There is a clear anticorrelation between the two histone marks. (D) For all CAGT runs around TFBSs, we considered all TSS-proximal sites that were assigned to asymmetric clusters, and counted how many times the direction of transcription of the TSS closest to a site agreed with (configuration (1)) or opposed (configuration (2)) the direction of the asymmetry pattern (from low to high signal) of the cluster to which the site was assigned. We are showing the log10-ratio of the two counts, aggregated over all CAGT runs for the same mark. Values >0 (corresponding to ratios >1) imply that the mark tends to increase in the direction of transcription, while values <0 imply that the mark tends to increase in the opposite direction.

OCT4 establishes and maintains nucleosome-depleted regions that provide additional layers of epigenetic regulation of its target genes

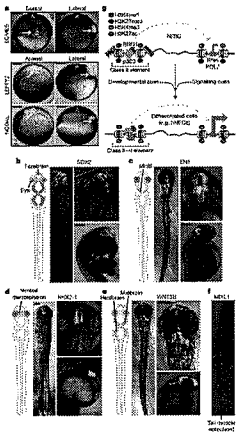
Jueng Soo You, Theresa K. Kelly, Daniel D. De Carvalho, Phillippa C. Taberlay, Gangning Liang, and Peter A. Jones[1]

PNAS | August 30, 2011 | vol. 108 | no. 35 | 14497–14502



A proposed model of the mechanism of repression and de novo DNA methylation of OCT4 and NANOG during cell differentiation. (Refer to Discussion for a detailed explanation.)

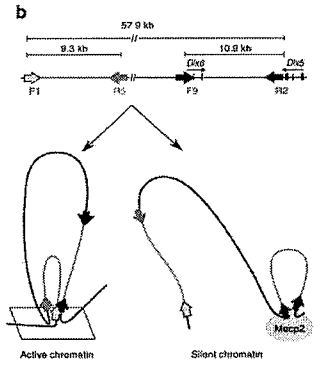
A unique chromatin signature uncovers early developmental enhancers in humans.
Rada-Iglesias A, et al. Nature. 2011 Feb 10;470(7333):279-83



Class II elements have developmental enhancer activity *in vivo*. a, Merged bright-field and GFP images are shown for representative shield stage zebrafish embryos injected with class II elements proximal to human EOMES, LEFTY2 and NODAL. For the EOMES enhancer, dorsal (anterior to top) and lateral (shield to right) views are presented in the left and right panels, respectively. For LEFTY2 and NODAL, animal pole (shield to top) and lateral (shield to right) views are presented in the left and right panels, respectively. White arrows indicate the location of the shield in each image. A, anterior; D, dorsal. Scale bar, 150 μm. b–f, Merged bright-field and GFP images are shown for representative 24–28 h.p.f. zebrafish embryos injected with class II elements proximal to SOX2 (b), EN1 (c), NKX2-1 (d), WNT8B (e) and MIXL1 (f) genes. In b–e, schematics highlighting the relevant anatomical structures where GFP expression was reproducibly observed are shown on the left, and three images correspond, from left to right and top to bottom, to whole-embryo flattened dorsal views, dorsal anterior views and lateral anterior views, respectively. In f, a lateral posterior view is shown. In b–f, scale bar = 150 μm. MHB, midbrain–hindbrain boundary. g, Proposed model for enhancer bookmarking during early embryonic development. Poised developmental enhancers (class II) are marked by a unique chromatin signature, involving occupancy of chromatin modifiers p300, BRG1 and PRC2 and nucleosomal regions marked by H3K4me1 and H3K27me3. During differentiation, appropriate developmental and signalling cues are able to rapidly transition these poised, pre-marked enhancers into an active state represented by the acquisition of H3K27ac, RNA POL2 binding, recruitment of tissue-specific transcription factors (TFs) and loss of H3K27me3, leading to the establishment of tissue-specific gene expression patterns.

- Gene expression and transcription is regulated at many levels, including at the chromatin level
- Differences in methylation can result in different chromatin loops
- Thus, CpG methylation can regulate nearby genes or genes 10-100kb away

- Take home: It's unclear whether regulated methylation events are impacting proximal or distal genes



Nat Gen 37, 31 (2005)

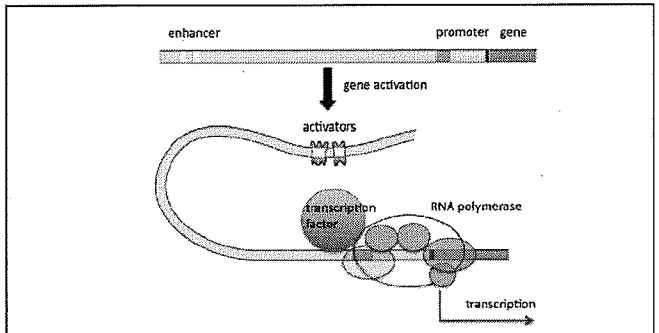


Figure 2 *Enhancers and long-range chromatin interactions*
Enhancer is a short DNA region to which activators can bind. Activators bind to an enhancer and affect the transcription of a gene that is located away from the enhancer by recruiting transcription factors and the Pol II complex to the relevant promoter. This process requires the role of three-dimensional chromatin looping

Chromatin-interaction compartment switch at developmentally regulated chromosomal domains reveals an unusual principle of chromatin folding

Shin-ichiro Takebayashi, Mahru Dileep, Tyrone Ryba, Jonathan M. Dennis, and David M. Gilbert

12574–12578 | PNAS | July 31, 2012 | vol. 109 | no. 31

Changes in chromatin interaction maps of the Dppa2/4 and Rex1 chromosomal domains parallel replication-timing changes during differentiation.

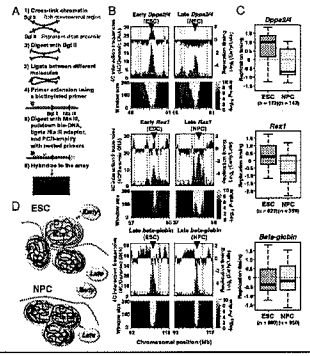


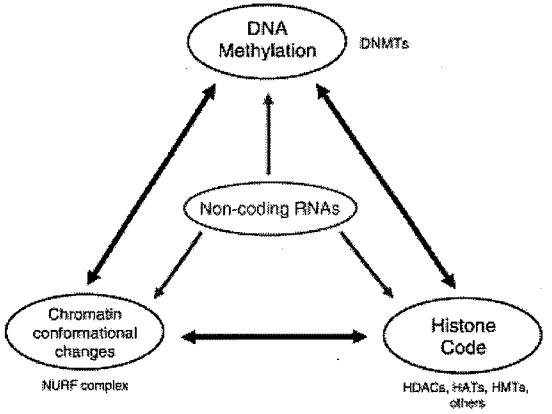
Table 2. Selected genetic disorders affecting chromatin structure in *trans*

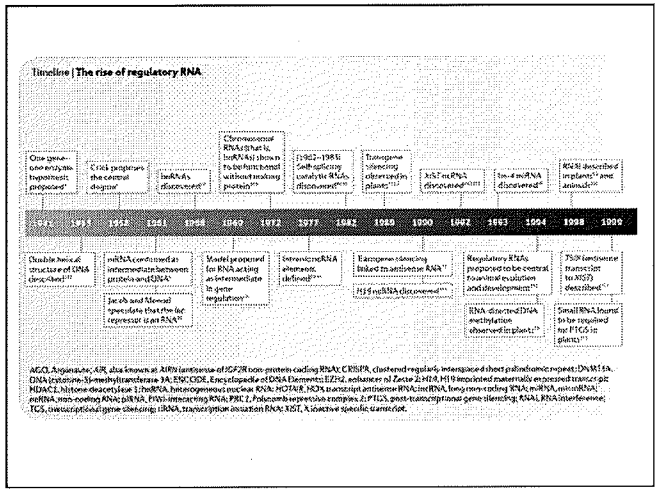
Disorder	Gene	Comments
Rubinstein-Taybi syndrome	*CREBBP, EP300*	
Rett syndrome	*MECP2*	loss of function as well as duplication causes a broad spectrum of phenotypes
α-Thalassemia and X-linked mental retardation	*ATRX*	somatic mutations cause α-thalassemia and myelodysplastic syndrome
ICF Syndrome	*DNMT3B*	
Schimke immuno-osseous dysplasia	*SMARCAL1*	
Mental retardation	*MTHFR*	

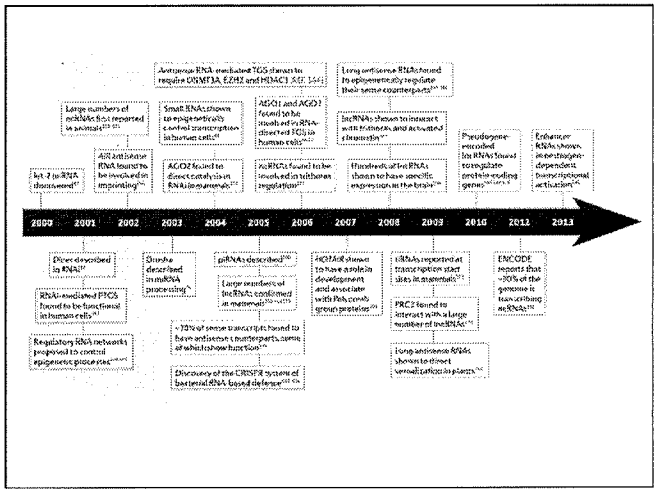
Table 3. Selected genetic disorders affecting chromatin structure in *cis*

Disorder	Gene	Comments
αδβ- and δβ-thalassemia	deletion of LCR causes decreased globin expression	
Fragile X syndrome	expansion of CCG repeat leads to abnormal methylation and silencing of *FMR1*	premutation alleles (60–200) cause a neurodegenerative disorder
FSH dystrophy	contraction of D4Z4 repeats causes less repressive chromatin	
Multiple cancers	germ-line epimutation of *MLH1*	

Epigenetics
non-coding RNAs







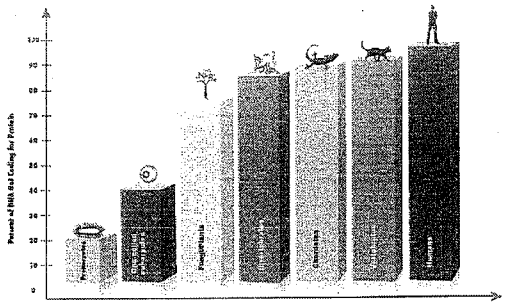
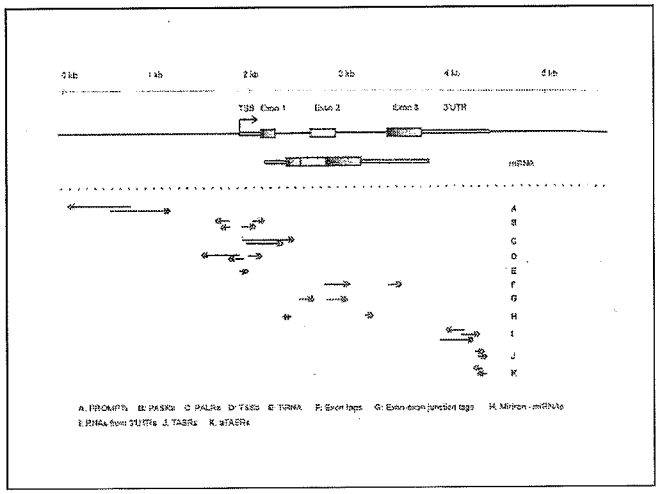


Figure 7.3 Non-coding RNAs as drivers of evolution and complexity? The histogram depicts the percentage of non-coding DNA in different organisms along the evolutionary scale. The Y axis represents the percentage of DNA not coding for protein and the X axis shows histograms for different organisms from prokaryote to humans. It is clear that in prokaryotes the percentage of DNA not coding for proteins is very small (approximately 20% or less). On the other hand, organisms such as chordates, vertebrates and human have more than 85% of non-coding DNA in their genomes. Importantly, a big proportion of this non-coding DNA is transcribed into lncRNAs, most of which still have unknown functions. Histograms and chart adapted by Kathryn A. Bischof.

Table 16.2 RNA classes emanating from transcriptional units

RNA class	Size	Location	Orientation	Reference
PROMPTs	long	0.5–2.5 kb upstream of TSS	Sense or antisense	Preker (2008)
PALRs	230 nt to 1 kb	Within 0.5 kb of TSS, can extend into 1st exon or intron	sense	Kapranov (2007), Fejes-Toth (2009)
PASRs	22–200 nt	Start downstream or upstream of TSS, within 0.5 kb of TSS	Sense or antisense	Kapranov (2007), Fejes-Toth (2009)
aPASRs	<200 nt	PASRs with poly-U tail at 5′ end. Proposed to be generated by RDRP-like copying from poly adenylated transcripts emanating from promoter regions	antisense	Kapranov (2010)
TIRNA	18 nt	−50 to +120 of TSS	Sense	Taft (2009)
TSSs	20–60 nt	250 nt upstream or 50 nt downstream of TSS	Sense or antisense	Seila (2008)
TASRs	20–200 nt	Near termination sites	Sense or antisense	Kapranov (2007)
aTASRs	<200 nt	Contain a poly-U tail at 5′ end. Proposed to be generated by an RDRP-like copying mechanism from 3′ ends of polyadenylated RNA	Antisense	Kapranov (2010)

TSS, transcription start sites; PROMPTs, promoter upstream transcripts; PALRs, promoter-associated long RNA; PASR, promoter-associated short RNA; TASRs, terminal-associated short RNA; aTASRs, antisense TASR; TSSs, TSS-associated RNAs; RDRP, RNA-dependent RNA polymerase.



A: PROMPTs B: PASRs C: PALRs D: TSSs E: TiRNA F: Exon tags G: Exon-exon junction tags H: Mirtron - miRNAs
I: RNAs from 3'UTRs J: TASRs K: uTASRs

Table 1. ncRNA classes and their function

ncRNA class	Size	Function	References
miRNAs	~23 nt	Regulation of hundreds to thousands of protein-coding and non-coding genes by mechanisms such as post-transcriptional gene silencing in animals and plants	[41–44]
piRNAs	26–31 nt	Repeat silencing in the genome and regulation of DNA methylation affecting gene expression	[45–49]
Small RNAs	~20–300 nt	Diverse functions from RNA modification to genomic imprinting in eukaryotic cells	[50–55]
LncRNAs	≥300 to thousands of nt	Diverse functions from structural mechanisms to gene regulation by epigenetic modifications	[9, 10, 56–61]

LncRNAs, long ncRNAs; miRNAs, microRNAs; nt, nucleotides; piRNAs, Piwi-interacting RNAs.
The classification of ncRNAs is mainly based in size and functionality and this categorization is very subjective. piRNAs and miRNAs are classified as small RNAs by several groups; however, they are described separately here since they differ in function.

Table 3. ncRNAs produced by eukaryotic genomes that have been recently described

ncRNA class	Size	Organism	Function	References
CUTs and SUTs	>200 nt	Yeast	Broad regulators of gene transcription in yeast by epigenetic mechanisms in bidirectional promoters	[12, 13]
Repetitive RNAs	Diverse	Mouse and human	Epigenetic regulation? Other mechanisms?	[14]
LincRNAs	≥200 nt	Mouse and human	Associated with diverse biological processes based on high conservation between different species	[9, 10]
tiRNAs	~18 nt	Human, chicken and fruit fly	Putative function in chromatin modifications and protein recruitment for transcription initiation	[11]
PASR RNAs	Produced as long ncRNAs (PALRs) and processed to small RNAs of ≤200 nt	Human	The majority has obscure function but it is speculated that they could serve as important components of regulatory circuitries	[23]
TSSa RNAs	20 to 90 nt	Human	Possible role in maintaining the structure of the chromatin and driving the transcription of nascent RNAs	[65]
NRO-RNAs	~100 nt	Human	They may play a role in gene transcription	[66]
PROMPTs and PALRs	An average of 500 nt	Human	Produced upstream of active genes and associated with chromatin changes	[23, 67]
nomtRNAs	≥200 nt	Human	Regulation of cell cycle and proliferation by unknown mechanisms	[19]
TUFs	≥200 nt to several kb	Human	Important functions in stem-cell differentiation	[70]

CUTs, cryptic unstable transcripts; kb, kilobases; LincRNAs, large intergenic ncRNAs; nomtRNA, non-coding mitochondrial RNAs; NRO-RNA, nuclear run-on RNAs; nt, nucleotides; PALR, promoter-associated long RNAs; PASR, promoter-associated short RNAs; PROMPT, promoter upstream transcripts; tiRNAs, transcription initiation RNAs; TSSa, transcription start site-associated RNA; TUF, transcripts of unknown function.

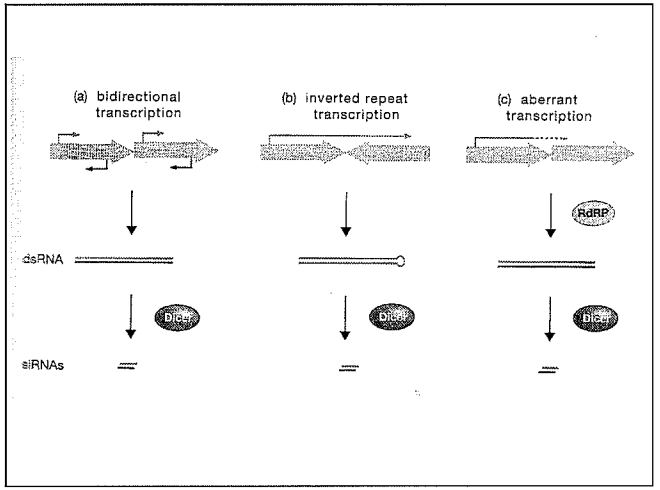
Table 4. Summary of new ncRNAs classes

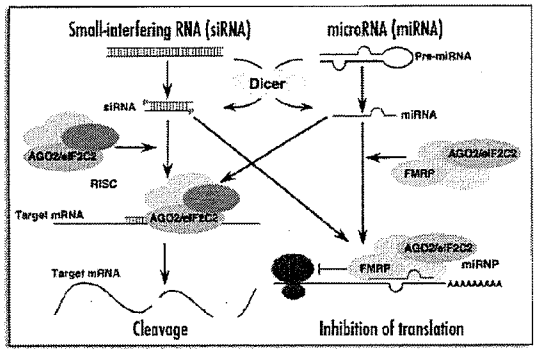
ncRNA class	Function	References
5′ region ncRNAs or paRNAs	Possible mechanism of action is to stabilize and maintain an opened chromatin for gene transcription	[11–13,23,66–68]
3′ intergenic ncRNAs	Anti-sense to protein-coding genes and putative function in gene regulation by epigenetic mechanisms	[13, 69]
Mitochondrial ncRNAs	Cell cycle regulation and other unknown mechanisms	[19]
Satellite ncRNAs	A putative role in the formation and function of centromere-associated complexes	[15]
Repeat-associated ncRNAs	Epigenetic regulation? Repeat silencing?	[14]
Other ncRNAs or TUFs	Function in important molecular mechanisms in stem cells with a role in differentiation and other unknown functions	[70]

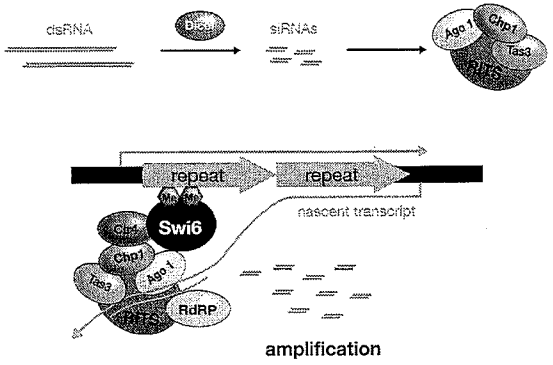
ncRNAs, non-coding RNAs; paRNA, promoter-associated ncRNAs; TUFs, transcripts of unknown function.

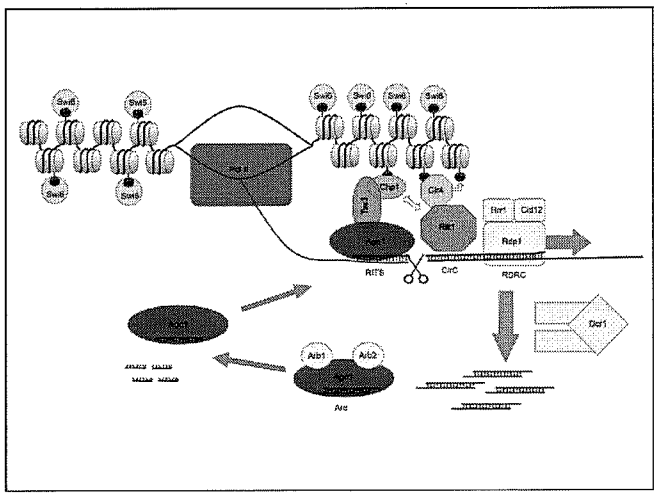
miRNA,siRNA and RNAi

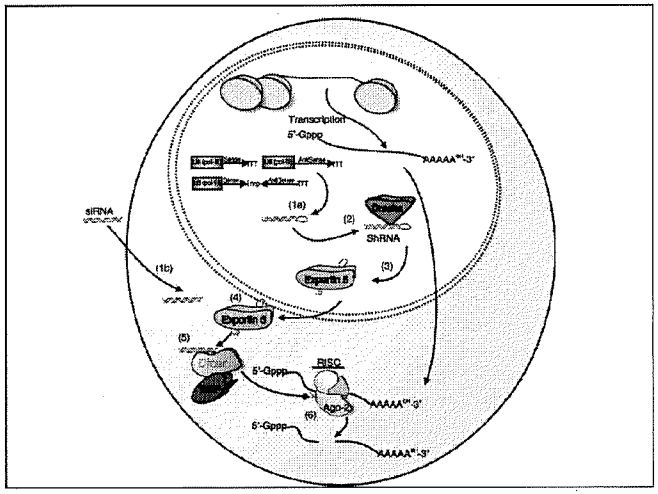
Sequence dependent mRNA stability and translation control

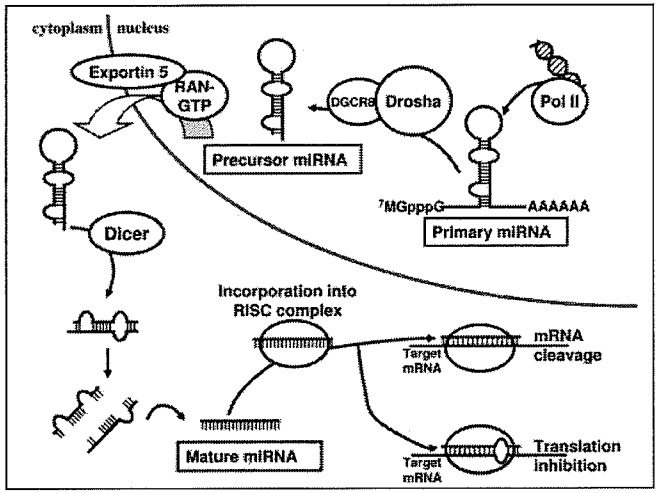












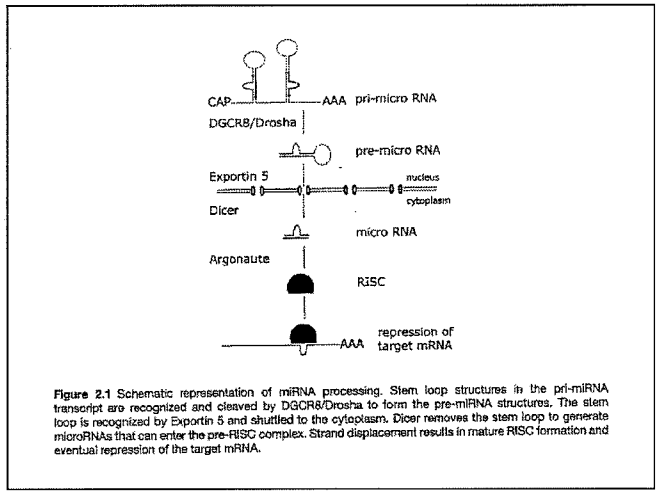


Figure 2.1 Schematic representation of miRNA processing. Stem loop structures in the pri-miRNA transcript are recognized and cleaved by DGCR8/Drosha to form the pre-miRNA structures. The stem loop is recognized by Exportin 5 and shuttled to the cytoplasm. Dicer removes the stem loop to generate microRNAs that can enter the pre-RISC complex. Strand displacement results in mature RISC formation and eventual repression of the target mRNA.

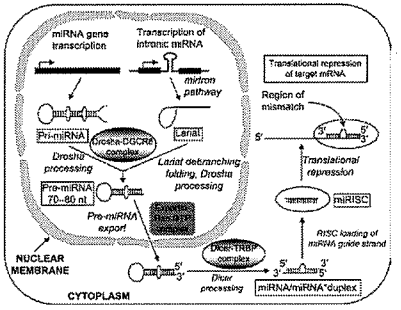


Figure 3. The pri-miRNA transcript is processed in the nucleus by Drosha-DGCR8 complex to produce a 70–80 nt long precursor miRNA (pre-miRNA), which is transported to the cytoplasm by Exportin-5 and Ran-GTP. Mirtrons are processed by spliceosomes, released as lariats that first undergo debranching followed by folding to form the pre-miRNA. In the cytoplasm, the 70–80 nt long pre-miRNA is further processed into 22 nt long mature duplex by Dicer. From this duplex, only the guide strand is loaded onto the RISC. If there is a region of mismatch between the target mRNA and the guide-strand sequences, the miRNA pathway is activated in which the miRNA binds to the 3′-UTR of the mRNA and suppresses its translation. Not shown in the figure is the siRNA pathway. If the target mRNA and the guide-strand sequences are fully complementary with no mismatch, the siRNA pathway is activated and the target mRNA is degraded (Figure reprinted from Choudhuri, S., Cui, L. and Klaassen, C.D. *Toxicol. Appl. Pharmacol.* 245: 378–393. 2010, with permission from Elsevier).

Table 1. Conservation of RNAi and heterochromatin proteins

S. pombe	*A. thaliana*	*C. elegans*	*Drosophila*	*H. sapiens*
Dcr1	DCL1 to 4	Dcr-1	Dcr1 and 2	DCR-1
Ago1	AGO1 to 10	Rde-1, Alg-1 and -2	Ago1 to 3, Piwi	AGO-1 to AGO-4
	PRG-1 and 2, and 19 others	Aubergine/Sting	Piwi-1 to Piwi-4	PIWI-1 to PIWI-4
Chp1[a]	CMT3	–	–	–
Tas3[b]	–	–	–	–
Rdp1	RDR1 to 6	Ego-1, Rrf-1 to -3	–	–
Hrr1	SGS2/SDE3	ZK1067.2	GH20028p	KIAA1404
Cid12	[c]	Rde-3; Trf-4[c]	CG11265[c]	POLS[c]
Swi6	LHP1 (TFL2)	Hpl-1, Hpl-2, F32E10.6	HP1	HP1α, β, γ
Clr4	SUVH2 to 6	[d]	Su(var)3-9	SUV39H1 and 2
Rik1[e]	DDB1	M18.5	Ddb1	DDB1
Cul4	CUL4	Cul4	Cul4	CUL4
Sir2	SIR2	Sir2-1	Sir2	SIRT1
Eri1	ERI1	Eri-1	CG6393	THEX1

[a] An obvious ortholog of the chromodomain protein, Chp1, has not been identified in the other model organisms listed here, but most eukaryotic cells contain multiple chromodomain proteins. CMT3 in *Arabidopsis* is a chromodomain DNA methyltransferase, which acts in the same pathway as AGO4 and may be analogous to Chp1.

[b] No obvious orthologs of Tas3 have been identified, but it shares weak sequence similarity with a mouse ovary testis specific protein (NP_035152).

[c] Cid12 belongs to a large family of conserved proteins that share sequence similarity with the classic poly(A) polymerase as well as 2′-5′-oligoadenylate enzymes.

[d] *C. elegans* have about 20 SET domain proteins, but an H3K9 HKMT has not yet been identified in this organism.

[e] *S. pombe* contains another Rik1-like protein, Ddb1, which is involved in DNA damage repair. Metazoans and plants appear to contain only a single Rik1-like gene, called Ddb1, which has been shown to be involved in DNA damage repair, but it is unknown whether it also participates in heterochromatin formation.

Table 1 Compilation of features of heterochromatin and RNAi from yeast, animals and plants

	Fungi		Ciliates	Animals			Plants
	S. pombe	*N. Crassa*	*T. thermophila*	*C. elegans*	*D. melanogaster*	*H. sapiens*	*A. thaliana*
H3K9me	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HP1	Swi6, Chp1 & Chp2	HP1	Pdd1p &Pdd3p	HPL1 &HPL2	HP1, HP1b & HP1c	HP1α, β & γ	TFL2/LHP1
DNA methylation	–	Yes	Yes	–	Low levels	Yes	Yes
short RNA	siRNA	siRNA	scnRNA	miRNA & siRNA	miRNA, siRNA, piRNA	miRNA, siRNA & piRNA	miRNA & siRNA
Dicer	Dcr1	DCL1 & DCL2	DCL1	Dcr1	Dcr1 & Dcr2	DCR1	DCL1, DCL2, DCL3 & DCL4
Argonaute	Ago1	QDE2 & SMS2	TWI1	27AGO genes	Ago1-3, Piwi1-4	AGO1-4, PIWI1-4	AGO1-10
RDP	Rdp1	QDE1, SAD1 & RRP3	–	Ego1, Rrf1-3	–	–	RDR1-6
ds siRNA ribonuclease	Eri1	n.d.	–	Eri1	CG6393	THEX1	–
pol IV	–	–	–	–	–	–	NRPD1a & b

n.d.: Not determined

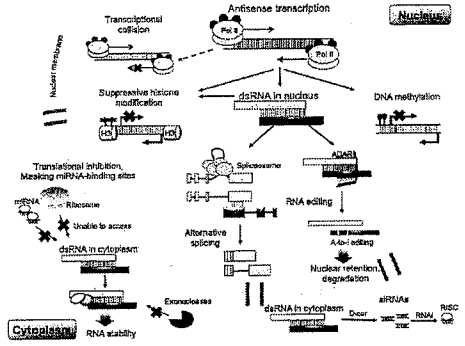
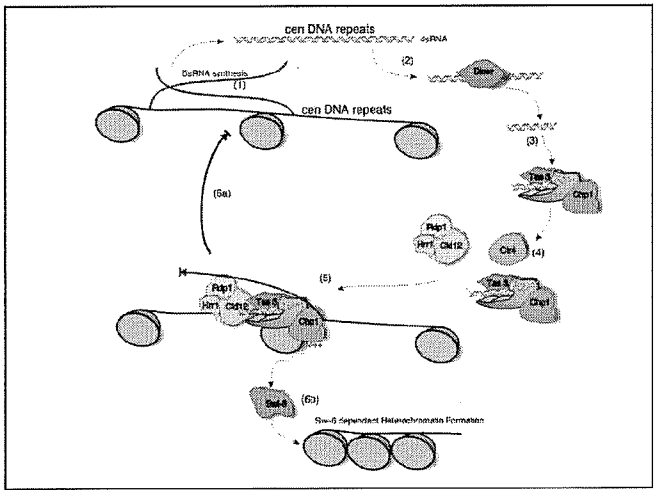
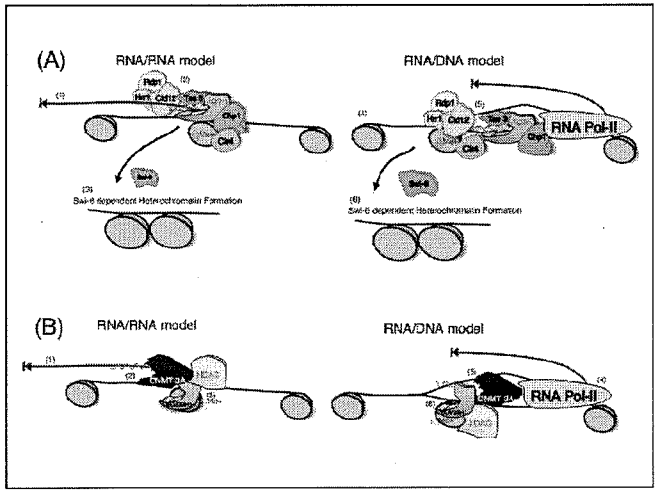
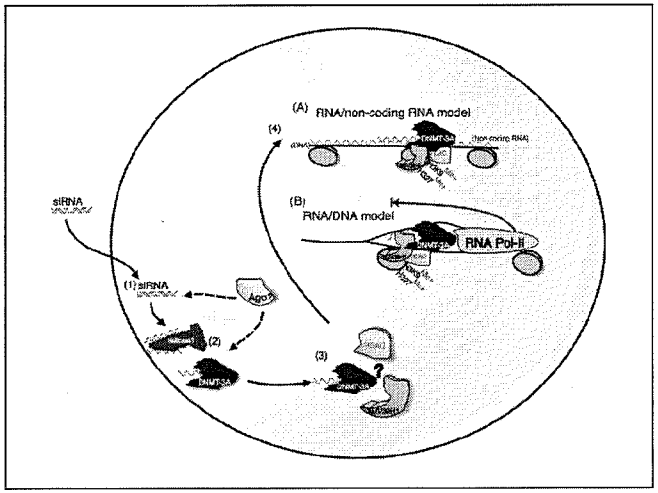
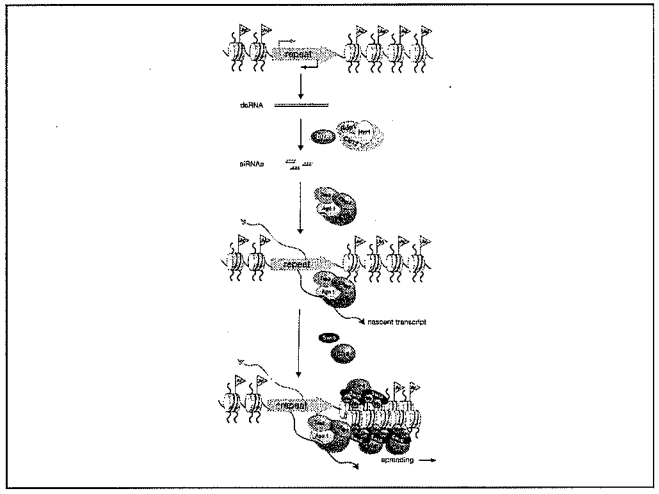


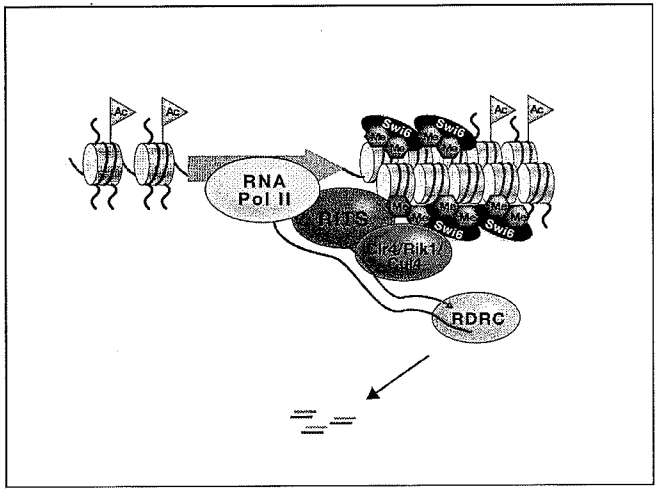
Figure 1.6 Summary of types of gene expression regulation starting with antisense transcription in the cell. Antisense transcription from SAT genes can lead to transcriptional collision, suppressive histone modification and methylation of DNA on the opposite strand. Antisense transcripts generate dsRNAs in the nucleus and cytoplasm. Some nuclear dsRNAs trigger alternative splicing and RNA editing. dsRNA production in the cytoplasm can lead to RNA stability by translational silencing, an impediment to the binding of miRNAs and exonucleases. The RNAi machinery produces siRNAs from dsRNAs. Pol II, RNA polymerase II; ADAR, adenosine deaminase acting on RNA; H3, histone 3; RISC, RNA-induced silencing complex.

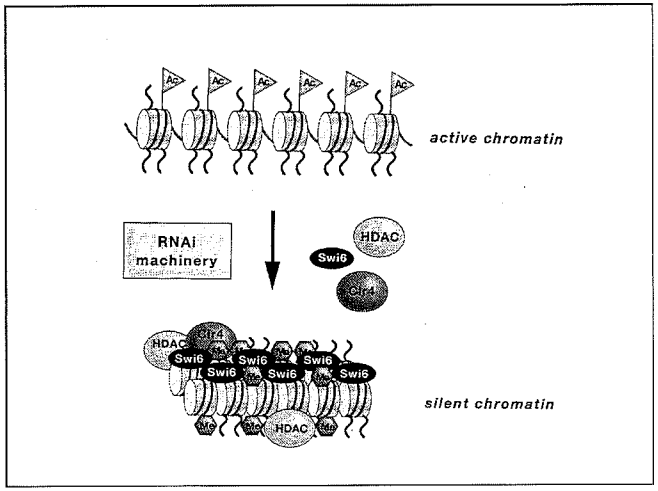

cen DNA repeats
dsRNA
dsRNA synthesis
(1)
(2)
Dcr1
cen DNA repeats
(3)
Tas3
Chp1
(6a)
Rdp1
Hrr1
Cid12
Ctr1
Ctr2
(4)
(5)
Swi6
(6b)
Swi6-dependent Heterochromatin Formation


(A)
RNA/RNA model
RNA/DNA model
RNA Pol-II
(B)
RNA/RNA model
RNA/DNA model
RNA Pol-II


(A) RNA/non-coding RNA model
(4)
(B)
RNA/DNA model
RNA Pol-II
siRNA
(1) siRNA
Ago1
(2)
(3)


dsRNA
siRNAs
nascent transcript
spliceosome





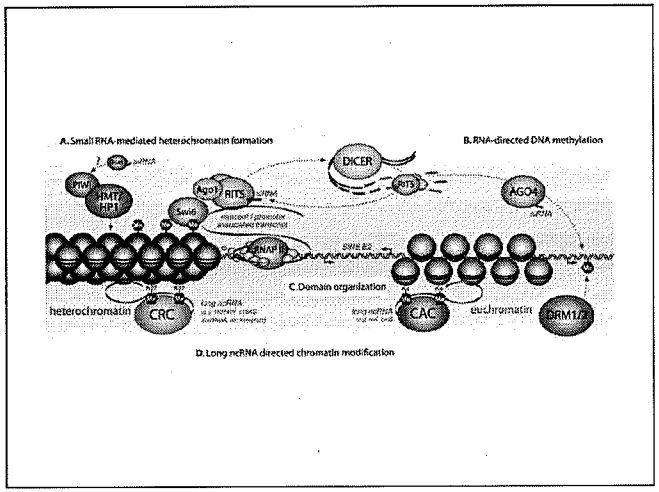


Table 1. *Oncogenic MiRNAs*

miRNA	Regulation	Role in cancer	References
miR-17-92 cluster	c-MYC induces the expression	Overexpressed in B-cell lymphomas and can accelerate tumor progression in a mouse B-cell lymphoma model. Plays a role in tumor angiogenesis. Overexpressed in lung cancers.	(46,47,62,63)
miR-372, miR-373		Overexpressed in testicular germ cell tumors. Permit proliferation and tumorigenesis of primary human cells that have both oncogenic RAS and active wild-type p53.	(54)
miR-21		Expressed in glioblastomas and breast cancers. An anti-apoptotic factor.	(34,36,37)
miR-155		Overexpressed in breast, colon, and lung cancers, as well as in B-cell lymphomas. Overexpression correlates with poor survival in lung cancers and poor prognosis in DLCBL. High expression in children with Burkitt's lymphoma.	(34,35,56,63,64–66)
miR-146	NF-κB	Overexpressed in breast, pancreas, and prostate cancers.	(66,67)

Table 2. *Tumor-suppressor MiRNAs*

miRNA	Regulation	Role in cancer	References
miR-127	DNA methylation and histone modifications	Decreased expression in bladder and prostate cancers. Targets proto-oncogene *BCL6*.	(49)
miR-15a, miR-16-1		Often down-regulated in B-cell chronic lymphocytic leukemias. Targets *BCL2*.	(32,33)
let-7		Targets RAS. Reduced expression in human lung cancer associates with poor prognosis.	(57–60)
miR-145		Reduced expression in colon and breast cancers.	(34,61)

Spring 2017 – Epigenetics and Systems Biology
Lecture Outline (Systems Biology)
Michael K. Skinner – Biol 476/576
Weeks 5, 6 and 7 (February 7, 14 and 21)

Epigenetics (History / Molecular Processes/ Genomics)

- Definitions and History
- Molecular Factors (DNA Methylation, Histone Modification, Chromatin Structure, ncRNA)
- Epigenetics Technology and Genomics

Required Reading

Holliday R. Epigenetics: a historical overview. Epigenetics. 2006 Apr-Jun;1(2):76-80.

Books (Reserve in Library)

Kevin V. Morris (2012) Non-coding RNAs and Epigenetic Regulation of Gene Expression: Drivers of Natural Selection. Caister Academic Press.

Russo, V.E.A., Martienssen, A.and Riggs, A.D. (eds.). 1996. Epigenetic Mechanisms of Gene Regulation. Cold Spring Harbor Press. Cold Spring Harbor.

Allis, C.D., Jenuwein, T. & Reinberg, D., Eds. (2007). Epigenetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Jeanteur, P. EPIGENETICS AND CHROMATIN. Progress in Molecular and Subcellular Biology, 2005, Volume 38, 151-167, DOI: 10.1007/3-540-27310-7_6

lncRNA and lincRNA

Sequence independent contol of transcription and genome activity

Table 2. ncRNAs and epigenetic modifications

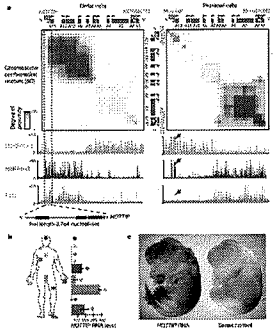
ncRNA(s)	Organism	Function in epigenetics	References
GAL10 ncRNA	Yeast	Influence in the post-translational modification of histones facilitating the expression of overlapping genes	[81, 82]
ICR1 and PWR1	Yeast	Cis-acting ncRNAs that modulate the expression of the *FLO11* gene in association to epigenetic mechanisms	[83]
Kcnq1ot1	Mice	Cis-acting silencer that might create a repressive environment for the epigenetic silencing of adjacent genes	[84]
~1600 lincRNAs	Mice	Highly conserved lncRNAs that can associate with chromatin complexes and histone modifications regulating gene expression	[9]
HOTAIR	Human	Regulation of the expression of genes in trans by selecting histone modifications and proteins that bind to DNA	[58]
~4000 lincRNAs	Human	Conserved lncRNAs that are able to modulate gene expression by associating to chromatin complexes with proteins and histone modifications	[10]

LncRNA, long non-coding RNAs; LincRNAs, large or long intergenic ncRNAs.

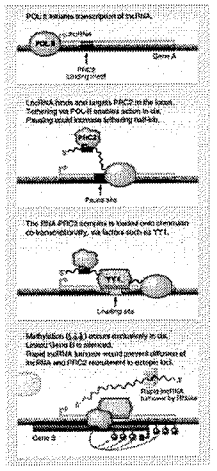
A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression

Kevin C. Wang[1–3], Yul W. Yang[2*], Bo Liu[1*], Amartya Sanyal[2], Ryan Corces-Zimmerman[1], Yong Chen[2], Bryan R. Lajoie[4], Angeline Protacio[1], Ryan A. Flynn[1], Rajnish A. Gupta[1], Joanna Wysocka[1], Ming Lei[2], Job Dekker[4], Jill A. Helms[3] & Howard Y. Chang[1]

126 | NATURE | VOL 472 | 7 APRIL 2011



a, Chromatin state map of distal versus proximal cells. Top panels, chromosome conformation capture-carbon copy (5C) analysis of distal (foreskin) and proximal (lung) human fibroblasts. Heat map representations (generated by my5C, ref. 30) of 5C data (bin size 30 kb, step size 3 kb) for *HOXA* in foreskin and lung fibroblasts. Red intensity of each pixel indicates relative interaction between the two points on the genomic coordinates. The diagonal represents frequent cis interactions between regions located in close proximity along the linear genome. 5C signals that are away from the diagonal represent long-range looping interactions. Bottom panels, chromatin occupancy across *HOXA*. x-axis is genomic coordinate; y-axis depicts occupancy of the indicated histone marks or protein (ChIP/input). Box and arrows highlight chromatin states of *HOTTIP* gene. b, *HOTTIP* RNA expression in primary human fibroblasts from 11 anatomic sites. Means ± s.d. are shown (n = 2). c, In situ hybridization of *HOTTIP* RNA in E13.5 mouse embryo.



LncRNAs tether epigenetic complexes to chromatin, enabling allele- and locus-specific regulation. LncRNA that is synthesized binds to an epigenetic complex (such as PRC2) and, together, are loaded onto chromatin cotranscriptionally through DNA-bound factors (such as YY1 for Xist RNA). Epigenetic modifications then silence the gene, and rapid lncRNA turnover prevents its diffusion to other loci.

Table 1. Emerging themes in lncRNA regulation. Potential groupings of lncRNA based on proposed interactions, functions, and mechanisms. Representative lncRNAs of each group are shown.

Functions	Examples	Action	Hypothesized mechanisms
Cis-tether cis-targeting	RepA, Xist, Tsix, Gtl2, Kcnq1ot1, Airn, Hottip, ANRIL, Oct4-ps5, COLDAIR, Evf2, BDNF-AS	Cis	Co-transcriptional targeting of chromatin factors: allelic and locus-specific action; repressive or activating
Trans-targeting	pRNA, asOct4-ps5	Trans	RNA:DNA triplex via Hoogsteen base pairing or sense:antisense Watson:Crick base pairing
Enhancer	Xite, ncRNA-a7, eRNAs	Cis	Mediated by RNA, transcriptional force, or chromatin intermediaries
Decoy	PTEN-ps Tsix	Trans Cis	Competitive inhibition of target protein or RNA
Scaffold	MALAT1, TUG1, NEAT1, HOTAIR, roX1, roX2	Trans	RNA subunit of effector complex
Allosteric modification	TLS	Cis (trans also possible)	Alters conformation and activity of protein partners
Coactivator or co-repressor	SRA, SINE B2, Jpx, pRNA	Trans (cis also possible)	Accessory factor for transcriptional activation or repression

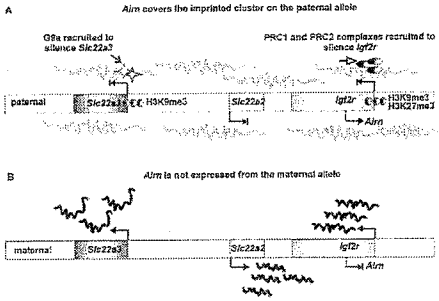


Figure 16.3 Graphical representation of mechanism of imprinting induced by *Air* macro ncRNAs at the *Igf2r* locus in placenta. *Air* overlaps partially with the *Igf2r* gene on the cluster containing *Igf2r*, Slc22a2 and Slc22a3 genes. *Air* RNA is shown as grey wavy lines in (A). The genes in this cluster are expressed from the maternal locus on which *Air* is silent. The black wavy lines in (B) depict RNA expressed from the three protein coding genes. *Air* is expressed in the antisense direction to *Igf2r* from the paternal cluster and covers the imprinted region on the paternal locus. *Air* recruits G9a (stars, open arrow) to mediate H3K9 trimethylation (shown as circles) at the *Slc22a3* locus resulting in silencing. At the *Igf2r* site, *Air* recruits Polycomb repressive complex 2 (PRC2) (ellipses, closed arrow) to induce trimethylation at H3K9 and H3K27 (shown as circles) to repress *Igf2r*. *Air* also recruits PRC1 to facilitate H2AK119 mono ubiquitinylation at the *Igf2r* locus and create a repressive environment devoid of RNA polymerase (not shown).

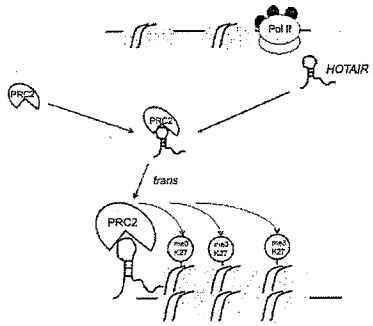


Figure 1.10 Model of regulation of chromatin domains via histone-modification enzymes with *HOTAIR*. *HOTAIR* transcribed from the *HOXC* locus recruits PRC2 to the *HOXD* locus *in trans*. PRC2 recruitment leads to H3K27 methylation and transcriptional silencing of neighbouring *HOX* genes.

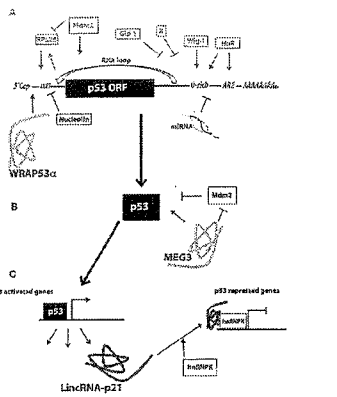


Figure 11.1 (A) Factors regulating mRNA stability and translation efficiency. X denotes a p53-regulating protein of unknown identity. (B) Regulation of p53 protein levels and transcriptional activity by the long ncRNA MEG3. (C) p53 can in turn activate the transcription of the long non-coding lincRNA-p21, which together with hnRNPK mediates transcriptional silencing of p53 transrepressed promoters.

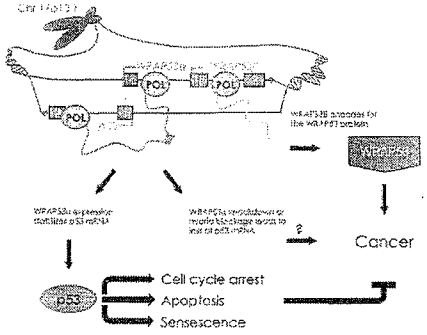
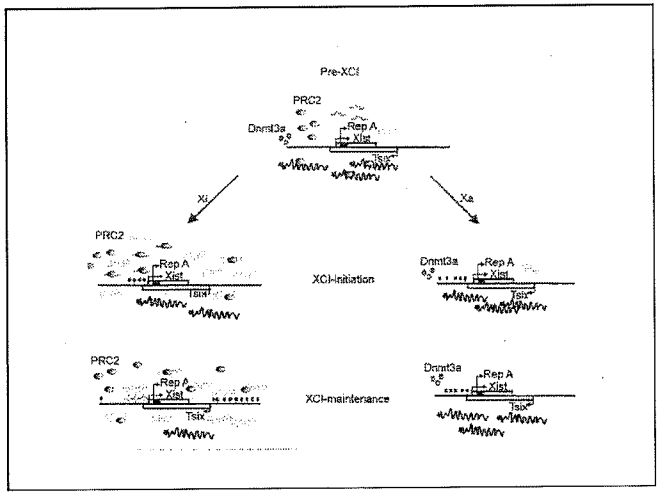
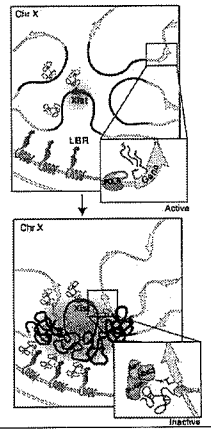


Figure 11.2 Model illustrating the *WRAP53* gene locus. WRAP53α transcripts originating from exon 1α overlaps p53 exon 1 and can via the complementarity region protect p53mRNA from degradation. Knockdown of WRAP53α or blockage of WRAP53α/p53 hybrid formation reduces p53 levels. Conversely, overexpression of WRAP53α increases p53 levels and sensitizes cells to p53-induced apoptosis. Transcripts mainly originating from exon 1β gives rise to the WRAP53 protein with oncogenic abilities.



Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing
Chen CK, Blanco M, Jackson C, et al.
Science. 2016 Oct 28;354(6311):468-472.



Non-canonical RNA-directed DNA methylation.

Cuerda-Gil D, Slotkin RK.
Nat Plants. 2016 Nov 3;2(11):16163.

Abstract

Small RNA-directed DNA methylation (RdDM) has been extensively studied in plants, resulting in a deep understanding of a major 'canonical RdDM' mechanism. However, current models of canonical RdDM cannot explain how this self-perpetuating mechanism is initiated. Recent investigations into the initiation of epigenetic silencing have determined that there are several alternative 'non-canonical RdDM' pathways that function through distinct mechanisms to modify chromatin. This Review aims to illustrate the diversity of non-canonical RdDM mechanisms described to date, recognize common themes within this dizzying array of interconnected pathways, and identify the key unanswered questions remaining in this field.

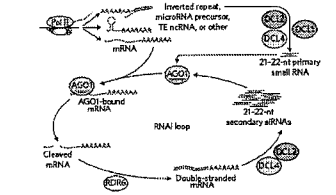
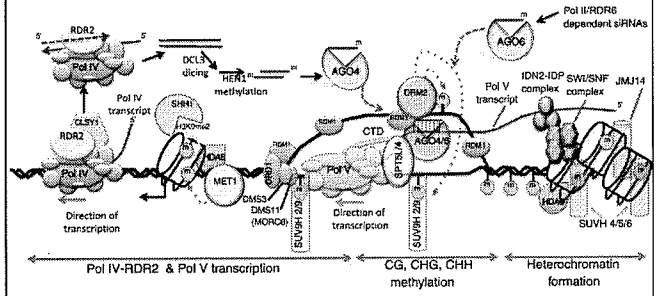


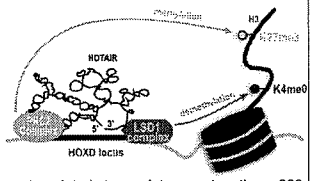
Figure 1 | The canonical RdDM pathway. In the upstream siRNA-generating phase, Pol IV and RDR2 function to produce dsRNA, which is cleaved into 24-nt siRNAs by DCL3. These siRNAs function to direct AGO4 or AGO6 to chromatin-bound transcripts produced by Pol V. In the downstream phase, the interaction of AGO4/6 with the Pol V transcript results in the recruitment and activity of the DNA methyltransferase DRM2. This pathway is reviewed in greater detail in ref. 7. Black strands represent DNA, red strands represent RNA. DNA methylation colours refer to the CG (red), CHG (blue) and CHH (green) sequence contexts, where H = A, C or T.

Figure 2 | PTGS in plants. Pol II transcription of regions such as palindromic TEs or microRNA precursors generates primary small RNAs[103], which are produced independently of RDR activity. Primary small RNAs target Pol II mRNAs for cleavage, and some of these cleaved mRNAs produce secondary siRNAs through the formation of dsRNA via RDR6 and cleavage by DCL2 and DCL4. Secondary siRNAs can target additional copies of the mRNA for cleavage and production of more siRNAs through the cycle of RNAi. ncRNA, non-coding RNA.

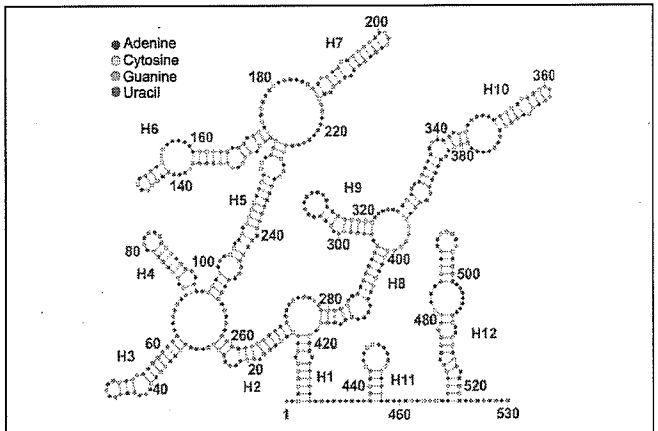
The RNAs of RNA-directed DNA methylation.
Wendte JM, Pikaard CS.
Biochim Biophys Acta. 2017 Jan;1860(1):140-148.



Biochemical Methods To Investigate lncRNA and the Influence of lncRNA:Protein Complexes on Chromatin.
McFadden EJ, Hargrove AE.
Biochemistry. 2016 Mar 22;55(11):1615-30.

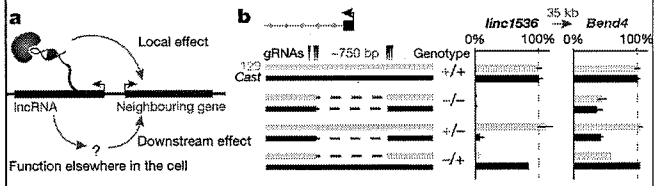


Long noncoding RNAs (lncRNAs), defined as nontranslated transcripts greater than 200 nucleotides in length, are often differentially expressed throughout developmental stages, tissue types, and disease states. The identification, visualization, and suppression/overexpression of these sequences have revealed impacts on a wide range of biological processes, including epigenetic regulation. Biochemical investigations on select systems have revealed striking insight into the biological roles of lncRNAs and lncRNA:protein complexes, which in turn prompt even more unanswered questions. To begin, multiple protein- and RNA-centric technologies have been employed to isolate lncRNA:protein and lncRNA:chromatin complexes. LncRNA interactions with the multi-subunit protein complex PRC2, which acts as a transcriptional silencer, represent some of the few cases where the binding affinity, selectivity, and activity of a lncRNA:protein complex have been investigated. At the same time, recent reports of full-length lncRNA secondary structures suggest the formation of complex structures with multiple independent folding domains and pave the way for more detailed structural investigations and predictions of lncRNA three-dimensional structure. This review will provide an overview of the methods and progress made to date as well as highlight new methods that promise to further inform the molecular recognition, specificity, and function of lncRNAs.



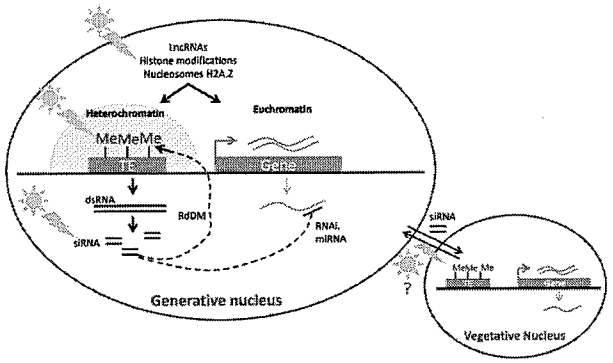
Secondary structure of human *HOTAIR* Domain 1 (1–530 n.t.; GI: 383286742). Pyle and co-workers performed chemical probing on the full-length *HOTAIR* transcript. Domain 1 contains the binding-domain of PRC2. Secondary structure rendered in VARNA.

Local regulation of gene expression by lncRNA promoters, transcription and splicing
Engreitz JM, Haines JE, Perez EM, et al (2016)
Nature. 17;539(7629):452-455.

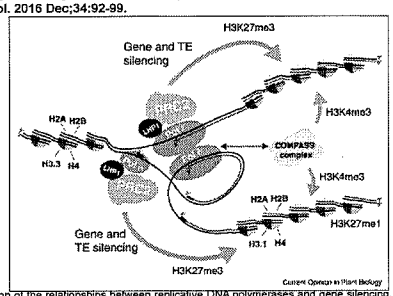


Many lncRNA and mRNA loci influence the expression of neighbouring genes.
a, Knocking out a promoter (black) could affect a neighbouring gene (blue) directly (local) or indirectly (downstream). b, Knockout of the linc1536 promoter. Left, genotypes; right, allele-specific RNA expression for 129 and castaneus (Cast) alleles normalized to 81 control clones (+/+). Error bars, 95% confidence interval for the mean (n = 2 for −/−, 3 for +/−, 1 for −/+).

Epigenetic events in plant male germ cell heat stress responses.
Chen Y, Müller F, Rieu I, Winter P.
Plant Reprod. 2016 Jun;29(1-2):21-9.

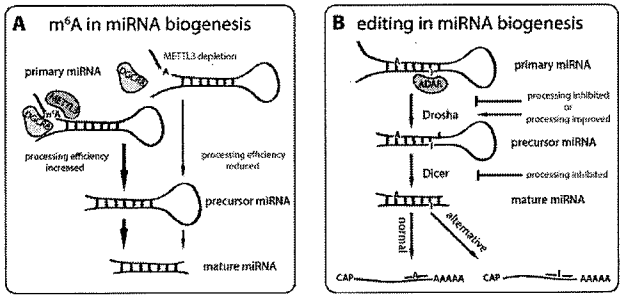


Links of genome replication, transcriptional silencing and chromatin dynamics.
Gutierrez C, Desvoyes B, Vergara Z, Otero S, Sequeira-Mendes J.
Curr Opin Plant Biol. 2016 Dec;34:92-99.



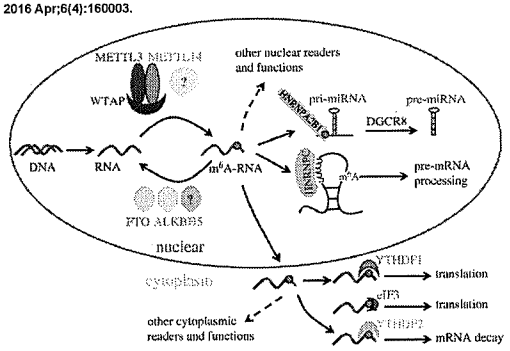
Schematic representation of the relationships between replicative DNA polymerases and gene silencing. During genome replication parental nucleosomes are disassembled and reassembled again past the replication fork. Each nucleosome contains two copies of each histone H2A, H2B, H3 and H4, as indicated. Newly assembled nucleosomes contain exclusively the canonical histone H3.1 (dark blue), which can be exchanged by the H3.3 (red). Histone chaperones responsible for depositing or exchanging histones H3.1 or H3.3 at the replication fork, as well chaperones for other histones, have been omitted for simplicity. Newly synthesized DNA strands are colored in red and the 4–6 nt long RNA primers appear as short thick lines (black). The three replicative DNA polymerases represent the holoenzyme complexes, DNA polymerases Pol δ and Pol ε, and DNA polymerase α-primase (Pol α). They are responsible for the elongation on the lagging DNA strand, on the leading strand and for the synthesis of RNA primers, respectively. Accessory replication proteins, RFC, RPA, PCNA, among others, have been omitted for simplicity. PRC2 and the LHP1 protein associate with Pol α and Pol ε to modify the chromatin past the fork, silencing genes and transposable elements (TEs) in a highly specific manner, as discussed in the text. The COMPASS complex interacts with Pol δ to modulate the H3K4me3 levels of newly formed chromatin.

Epigenetics
RNA Methylation



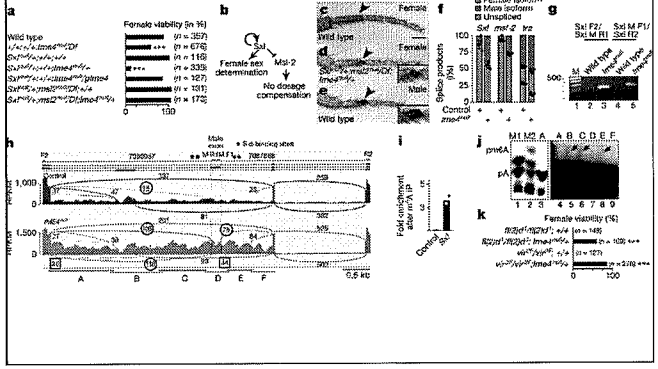
The proposed function of post-transcriptional RNA modifications in miRNA biogenesis. A) m⁶A methylation at the base of the pri-miRNA stem-loop is proposed to increase Drosha processing. This was confirmed by METTL3 depletion which results in decreased pri-miRNA processing and consequently reduced accumulation of mature miRNA. B) A-to-I editing effects on miRNA biogenesis are specific to each pri-miRNA transcript. Editing can inhibit either Drosha or Dicer processing, but may also enhance Drosha processing of the pri-miRNA. Editing sites present within the mature miRNA can lead to the targeting of a distinct alternative set of transcripts.

Recent advances in dynamic m6A RNA modification.
Cao G, Li HB, Yin Z, Flavell RA.
Open Biol. 2016 Apr;6(4):160003.

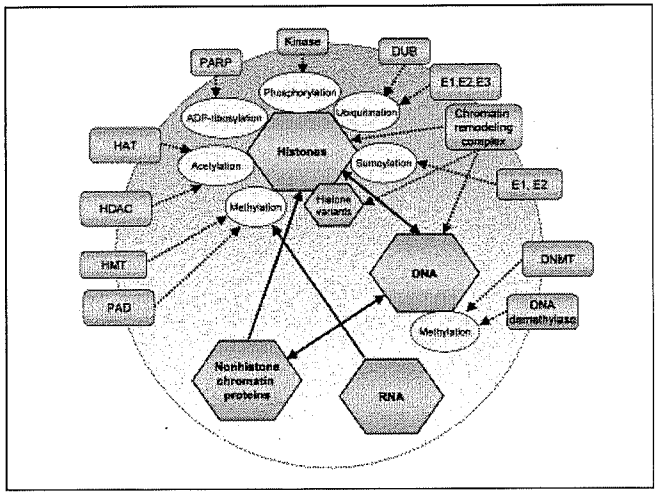


Dynamic m^{6}A RNA modifications and mediated functions. m^{6}A mRNA methylation is mediated by a multiprotein complex that includes METTL3, METTL14 and WTAP, whereas demethylases, such as FTO and ALKBH5, erase m^{6}A. Recognition of m^{6}A by HNRNPC in the nucleus mediates alternative splicing of pre-mRNA, and HNRNPA2B1 promotes pri-miRNA processing to pre-miRNA. In cytoplasm, binding of m^{6}A sites with different readers mediates divergent functions. YTHDF1 binds m^{6}A-modified mRNAs through interactions with initiation factors and ribosomes to increase translational output, and eIF3 can also directly bind to 5′UTR m^{6}A to initiate translation, whereas m^{6}A recognition by YTHDF2 leads to mRNA decay. More nuclear and cytoplasmic readers need to be defined to illuminate the functions of m^{6}A in mRNA export, translation and storage.

m^{6}A potentiates *Sxl* alternative pre-mRNA splicing for robust *Drosophila* sex determination.
Haussmann IU, Bodi Z, Sanchez-Moran E, et al, (2016)
Nature. 8;540(7632):301-304.



Epigenetics
Molecular Integration



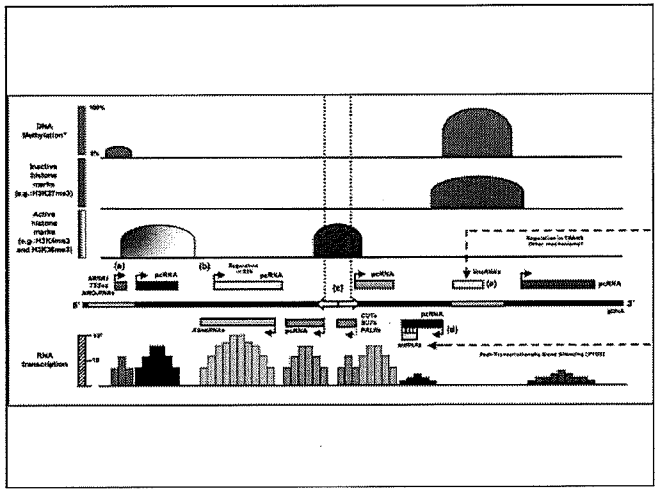
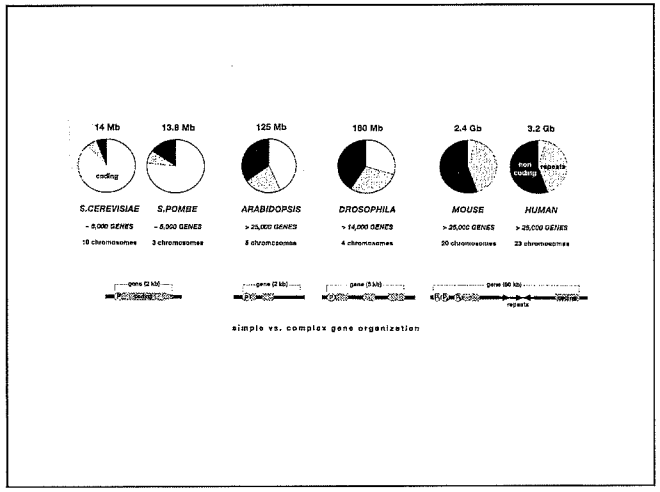
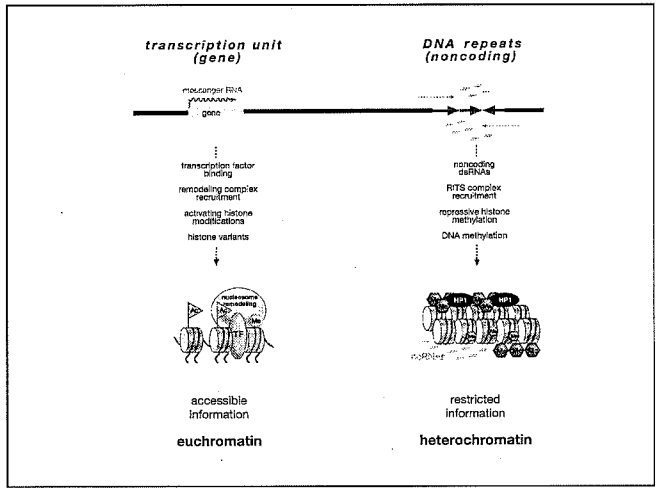


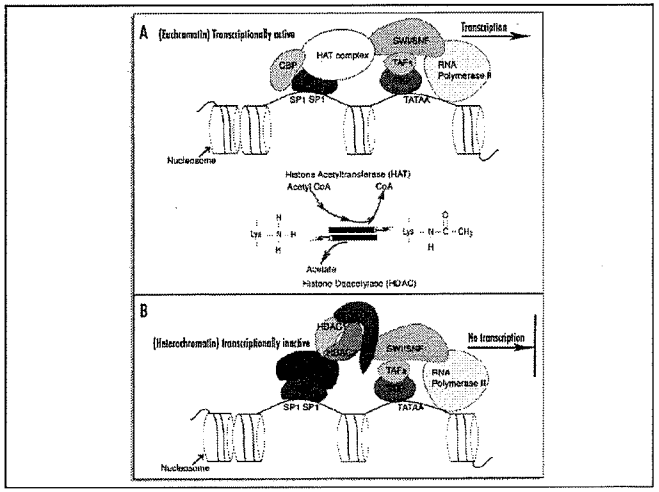
Table 16.1 Genome size increases with complexity of organisms but the gene density is not proportional to genome size (http://www.ornl.gov/sci/techresources/Human_Genome/faq/compgen.shtml)

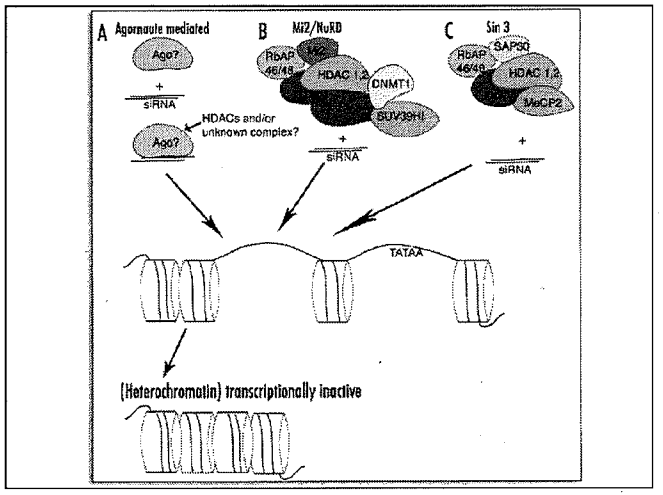
Organism	Genome size (million bps)	Number of chromosomes	Number of genes	Genes per 100,000 bp	Reference
Homo sapiens	3200	46	~25,000	~1	International Human Genome Sequencing Consortium (2004)
Mus musculus	2600	40	~25,000	~1	Waterston *et al.* (2002)
D. melanogaster	137	8	~13,000	~10	Adams *et al.* (2000)
A. thaliana	100	10	~25,000	~25	The *Arabidopsis* Genome Initiative (2000)
C. elegans	97	12	~19,000	~20	The *C. elegans* Sequencing Consortium (1998)
S. cerevisiae	12.1	32	~6000	~50	Goffeau *et al.* (1996)
E. coli	4.6	1	~3200	~70	Blattner *et al.* (1997)
H. influenzae	1.8	1	~1700	~95	Fleischmann *et al.* (1995)

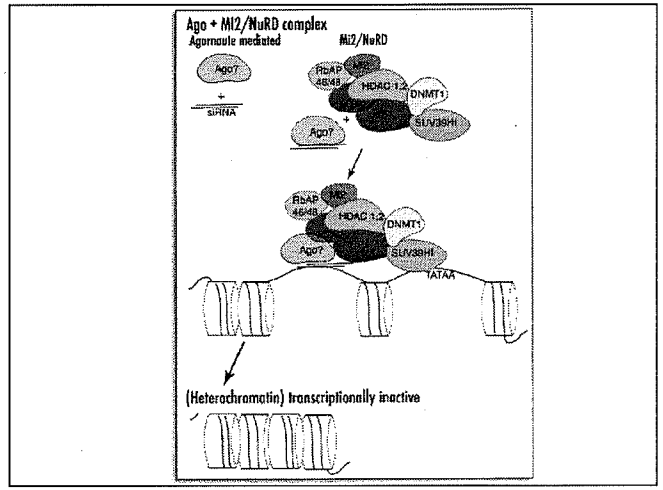


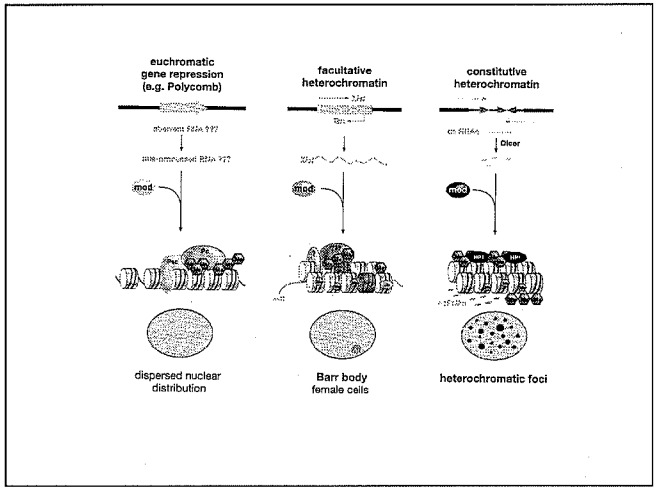
simple vs. complex gene organization

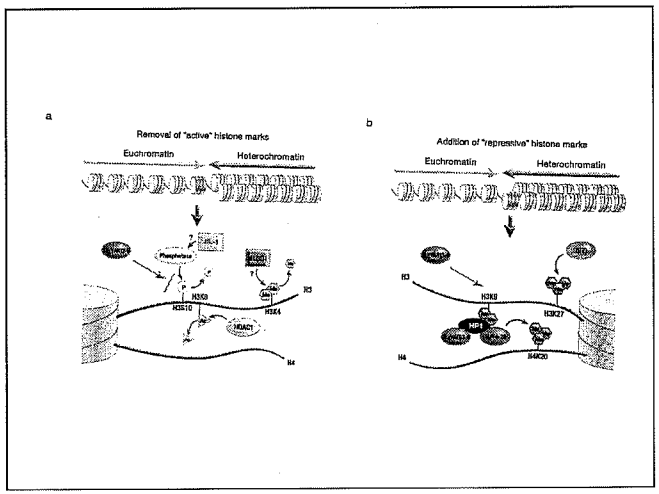


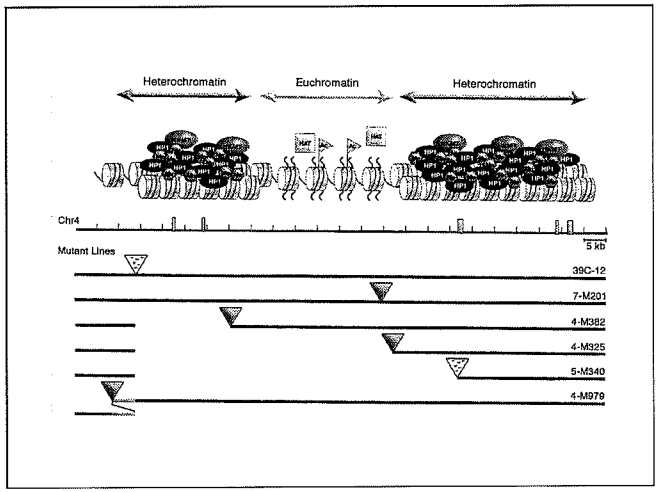

A (Euchromatin) Transcriptionally active
Transcription
SWI/SNF
HAT complex
CBP
TAFs
RNA Polymerase II
SP1 SP1
TATAA
Nucleosome
Histone Acetyltransferase (HAT)
Acetyl CoA
CoA
Acetate
Histone Deacetylase (HDAC)
B
(Heterochromatin) transcriptionally inactive
HDAC
No transcription
TAFs
RNA Polymerase II
SP1 SP1
TATAA
Nucleosome

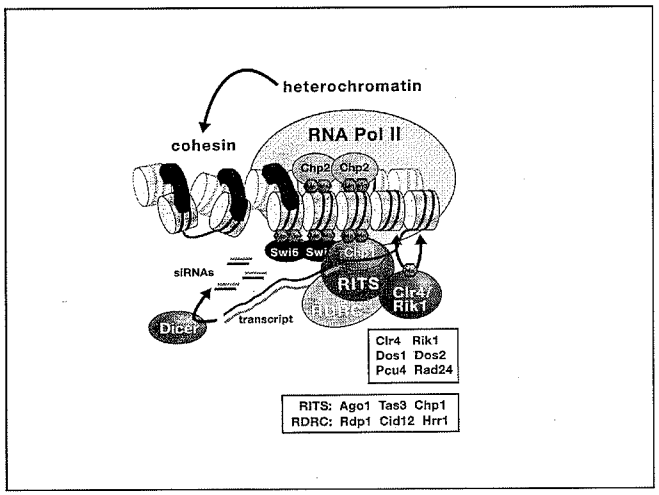

A Agornaute mediated
B Mi2/NuRD
C Sin 3
Ago?
siRNA
RbAP 46/48
HDAC 1,2
DNMT1
SUV39HI
SAP30
MeCP2
HDACs and/or unknown complex?
TATAA
(Heterochromatin) transcriptionally inactive

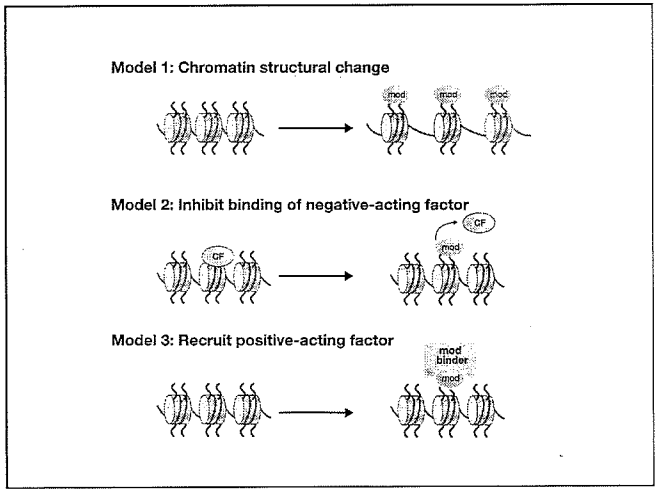

Ago + Mi2/NuRD complex
Agornaute mediated
Mi2/NuRD
Ago?
siRNA
RbAP 46/48
HDAC 1,2
DNMT1
SUV39HI
TATAA
(Heterochromatin) transcriptionally inactive

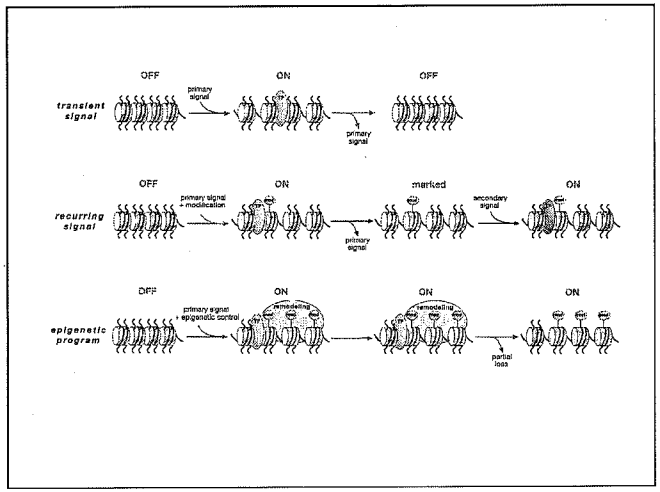

euchromatic gene repression (e.g. Polycomb)
facultative heterochromatin
constitutive heterochromatin
Dicer
dispersed nuclear distribution
Barr body female cells
heterochromatic foci

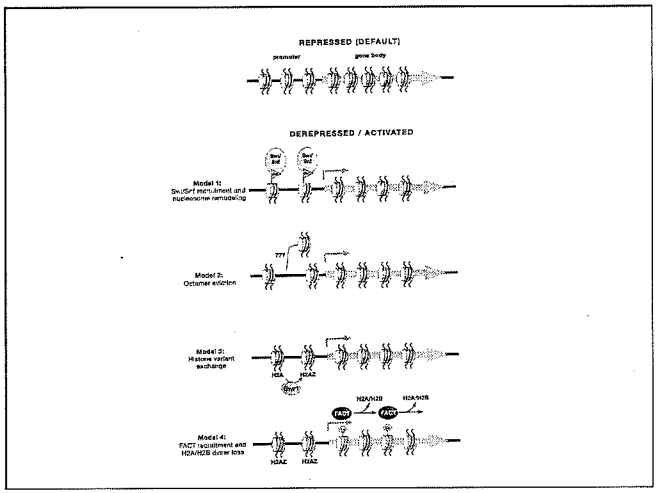

a
Removal of "active" histone marks
Euchromatin
Heterochromatin
b
Addition of "repressive" histone marks
Euchromatin
Heterochromatin

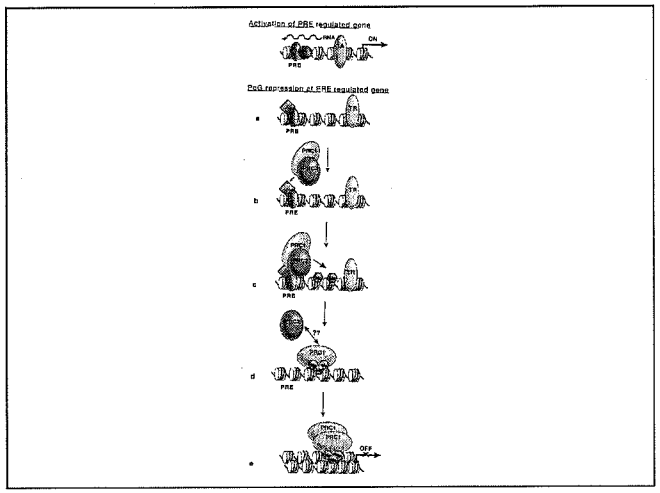

Heterochromatin
Euchromatin
Heterochromatin
Chr4
5 kb
Mutant Lines
39C-12
7-M201
4-M382
4-M325
5-M340
4-M979

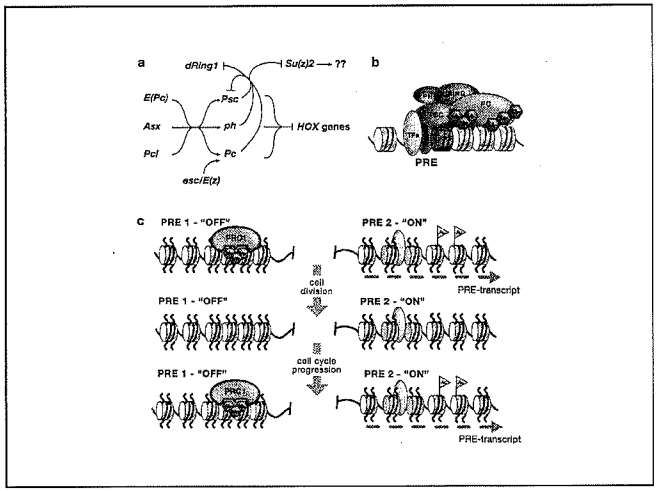

heterochromatin
cohesin
RNA Pol II
Chp2
Swi6
Swi6
Chp1
siRNAs
RITS
Clr4
Rik1
Dicer
transcript
Clr4 Rik1
Dos1 Dos2
Pcu4 Rad24
RITS: Ago1 Tas3 Chp1
RDRC: Rdp1 Cid12 Hrr1

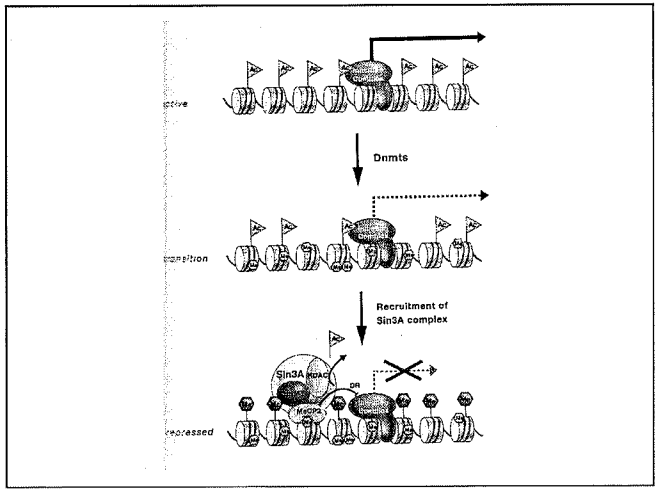

Model 1: Chromatin structural change
mod
mod
mod
Model 2: Inhibit binding of negative-acting factor
CF
CF
mod
Model 3: Recruit positive-acting factor
mod binder
mod

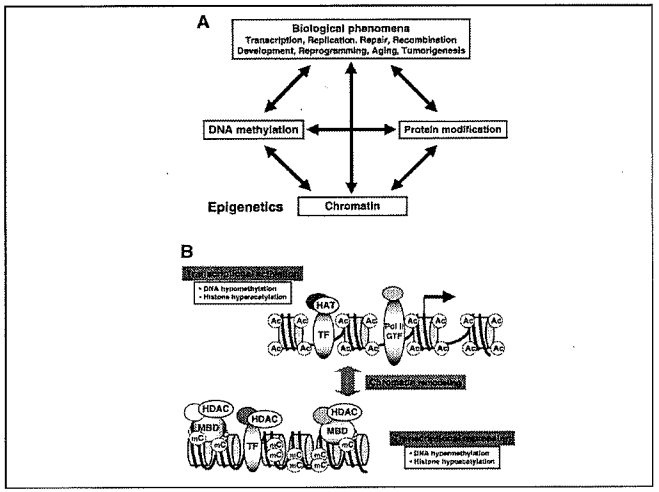

OFF
ON
OFF
transient signal
primary signal
primary signal
OFF
ON
marked
ON
recurring signal
primary signal + modification
primary signal
secondary signal
OFF
ON
ON
ON
epigenetic program
primary signal + epigenetic control
remodelling
remodelling
partial loss


REPRESSED (DEFAULT)
promoter
gene body
DEREPRESSED / ACTIVATED
Model 1: Su(z)2 recruitment and nucleosome remodelling
Model 2: Octamer eviction
Model 3: Histone variant exchange
H2A
H2AZ
Model 4: FACT recruitment and H2A/H2B dimer loss
H2A/H2B
FACT
H2AZ

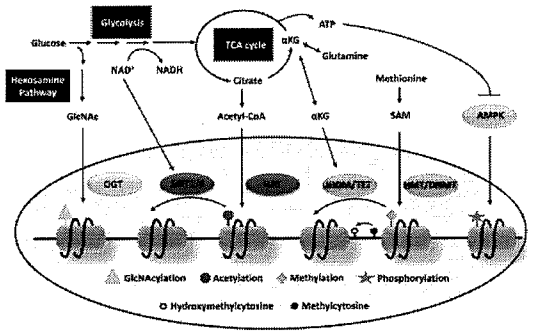

Activation of PRE regulated gene
ON
PRE
PcG repression at PRE regulated gene
a
PRE
b
PRE
c
PRE
d
PRE
OFF


a
dRing1
Su(z)2 — ??
E(Pc)
Psc
Asx
ph
HOX genes
Pcl
Pc
esc/E(z)
b
PRE
c
PRE 1 - "OFF"
PRE 2 - "ON"
cell division
PRE-transcript
PRE 1 - "OFF"
PRE 2 - "ON"
cell cycle progression
PRE 1 - "OFF"
PRE 2 - "ON"
PRE-transcript



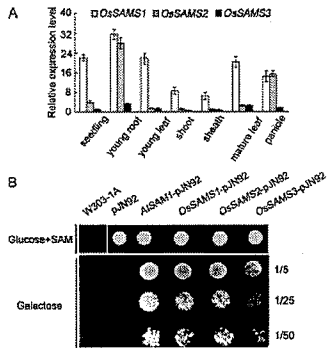


Metabolic regulation of epigenetics.
Lu C, Thompson CB. Cell Metab. 2012 Jul 3;16(1):9-17.

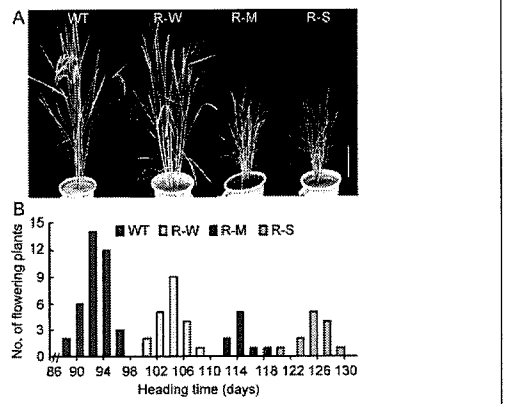


Crosstalk between Metabolism and Epigenetics As glucose enters the glycolytic pathway, a minor portion is branched to hexosamine biosynthetic pathway to produce GlcNAc which can be used as substrate for histone GlcNAcylation by OGT. Flux through glycolysis determines the NAD+/NADH ratio which is important for the activities of sirtuin histone deacetylases. Several TCA cycle intermediates can be exported out of mitochondria including citrate and αKG. Cytosolic citrate is converted to acetyl-CoA which is used as a donor for HAT-mediated histone acetylation. αKG is used as cofactor for histone and DNA demethylation reactions by JHDM and TET, respectively. The substrate for HMT and DNMT is SAM, which is synthesized from essential amino acid methionine. Finally, a low ATP/AMP ratio can activate AMPK, a kinase that phosphorylates histones.

Knockdown of SAMS genes encoding S-adenosyl-l-methionine synthetases causes methylation alterations of DNAs and histones and leads to late flowering in rice.
Li W, et al. J Plant Physiol. 2011 Oct 15;168(15):1837-43.



Biochemical characterization of Oryza sativa S-adenosylmethionine synthetase (OsSAMS1, 2, 3) and expression patterns. (A) Expression profile of OsSAMS1, 2, 3 in different ZH 10 organs detected by real-time RT-PCR. Error bars represent ±SD. (B) OsSAMS1, 2, 3 complemented yeast sam1 sam2 double mutant W744-1A. Yeast transformants were grown as patches on uracil-less synthetic glucose media (top panel), and then were diluted and replicated to uracil-less media containing galactose (bottom panel). 1/5, 1/25 and 1/50 meant dilution ratio. W303-1A was used as WT yeast, AtSAM1 was used as a positive control.

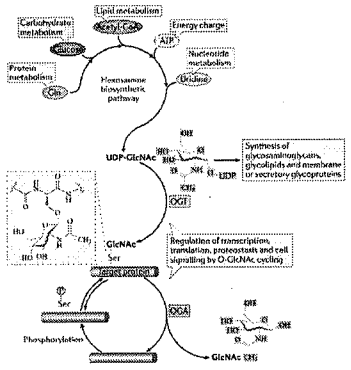


Phenotypic comparison of the OsSAMS1-RNAi transgenic plants was shown during the heading period. (A) Transgenic plants showed lato-flowering phenotype in the field. Rice plants were photographed at 114 DAG. Bar = 20 cm (B) Heading time distributions of the transgenic rice plants grown in the field. OsSAMS1-RNAi-weak, medium-strength and strong transgenic plants were abbreviated to R-W, M, S, respectively.

Bittersweet memories: linking metabolism to epigenetics through O-GlcNAcylation.
Hanover JA, et al. Nat Rev Mol Cell Biol. 2012 23;13(5):312-21

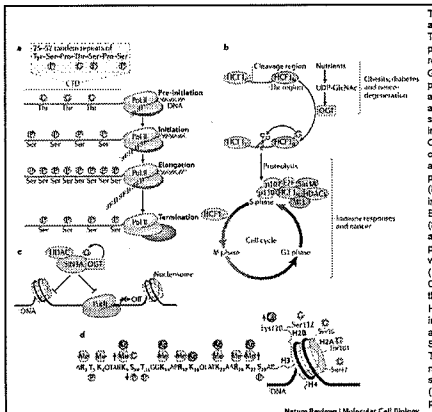
Abstract

O-GlcNAcylation, which is a nutrient-sensitive sugar modification, participates in the epigenetic regulation of gene expression. The enzymes involved in O-linked β-D-N-acetylglucosamine (O-GlcNAc) cycling - O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA) - target key transcriptional and epigenetic regulators including RNA polymerase II, histones, histone deacetylase complexes and members of the Polycomb and Trithorax groups. Thus, O-GlcNAc cycling may serve as a homeostatic mechanism linking nutrient availability to higher-order chromatin organization. In response to nutrient availability, O-GlcNAcylation is poised to influence X chromosome inactivation and genetic imprinting, as well as embryonic development. The wide range of physiological functions regulated by O-GlcNAc cycling suggests an unexplored nexus between epigenetic regulation in disease and nutrient availability.



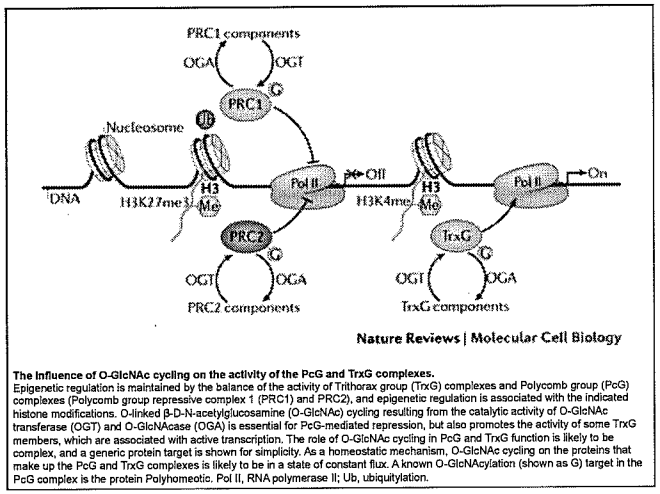
Nature Reviews | Molecular Cell Biology

The synthesis of UDP-GlcNAc via the hexosamine biosynthetic pathway is nutrient responsive.
The hexosamine biosynthetic pathway (HBP) integrates the metabolism of carbohydrates, amino acids, fat and nucleotides in the synthesis of uridine diphosphate N-acetylglucosamine (UDP-GlcNAc). UDP-GlcNAc is utilized for the synthesis of glycosaminoglycans, glycolipids and membrane and secretory glycoproteins. In addition, it is utilized by O-GlcNAc transferase (OGT) in the nucleus and the cytoplasm for the addition of GlcNAc at Ser or Thr residues (not shown) of target proteins (O-GlcNAcylation). Addition and removal of O-GlcNAc by OGT and O-GlcNAcase (OGA), respectively, contributes to continuous O-GlcNAc cycling, which dynamically modifies numerous intracellular proteins. O-GlcNAcylation often occurs at sites that are also targeted by phosphorylation. Thus, the resulting diversity of O-GlcNAcylated and phosphorylated proteins might contribute to the regulation of several complex cellular processes, including transcription, translation, proteostasis and cell signalling.

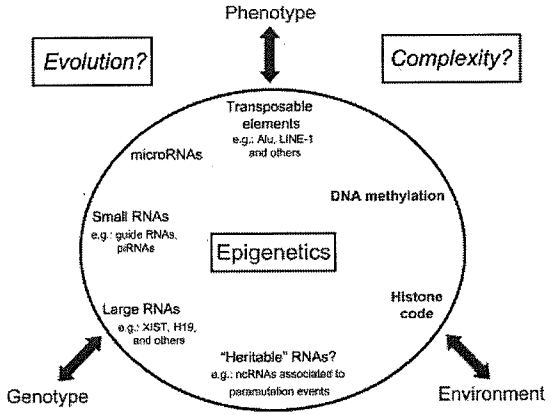


Nature Reviews | Molecular Cell Biology

The enzymes of O-GlcNAc cycling interact with and modify known epigenetic regulators. a | The carboxy-terminal domain (CTD) of RNA polymerase II (Pol II) consists of tandem heptad repeats that are phosphorylated and O-GlcNAcylated. Changes in the sites of phosphorylation (P, shown in orange) are associated with the indicated transcriptional activity of Pol II. An O-GlcNAcylation site (G, shown in blue) has been mapped to Thr4 and may inhibit Ser5 and Ser2 phosphorylation of Pol II. **b |** O-GlcNAc transferase (OGT) interacts with host cell factor 1 (HCF1) and promotes cleavage and activation of HCF1, which allows cell cycle progression. The HCF1 amino-terminal subunit (HCF1N) acts in complex with other factors, including the transcription factors p130, p107 and E2F, the co-repressor switch-independent 3A (SIN3A)–histone deacetylase (HDAC) complex and the Trithorax group-related protein MLL, to promote transition from G1 phase to S phase, whereas the HCF1 carboxy-terminal subunit (HCF1C) promotes entry to M phase. O-GlcNAcylation sites have been mapped throughout the HCF1 protein. Disruption of the regulation of HCF1 by O-GlcNAcylation may have important implications for human disease. **c |** OGT associates with and modifies the co-repressor SIN3A–HDAC complex to repress transcription. **d |** The core histones H2A, H2B, H3 and H4 are modified by O-GlcNAcylation. The mapped target sites for O-GlcNAcylation in H4 (Ser47), H2A (Thr101) and H2B (Ser36) are also phosphorylation sites that are important for nucleosome assembly. The histone tails of H3 and H2B contain O-GlcNAcylation sites at Ser10 and Ser112, respectively. Arrows indicate the effect of O-GlcNAcylation on the other histone post-translational modifications. Me, methylation; Ub, ubiquitylation.



The influence of O-GlcNAc cycling on the activity of the PcG and TrxG complexes.
Epigenetic regulation is maintained by the balance of the activity of Trithorax group (TrxG) complexes and Polycomb group (PcG) complexes (Polycomb group repressive complex 1 (PRC1) and PRC2), and epigenetic regulation is associated with the indicated histone modifications. O-linked β-D-N-acetylglucosamine (O-GlcNAc) cycling resulting from the catalytic activity of O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA) is essential for PcG-mediated repression, but also promotes the activity of some TrxG members, which are associated with active transcription. The role of O-GlcNAc cycling in PcG and TrxG function is likely to be complex, and a generic protein target is shown for simplicity. As a homeostatic mechanism, O-GlcNAc cycling on the proteins that make up the PcG and TrxG complexes is likely to be in a state of constant flux. A known O-GlcNAcylation (shown as G) target in the PcG complex is the protein Polyhomeotic. Pol II, RNA polymerase II; Ub, ubiquitylation.



Epigenetics
Genomics Technology

General Approach to Method

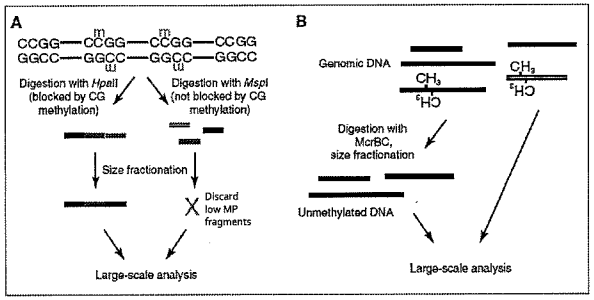
- Sample collection
- Sample enrichment to reduce genome complexity
- Biochemical modification of DNA
- Genomic analyses
  - Array-based
  - Sequencing-based
- Bioinformatic analysis

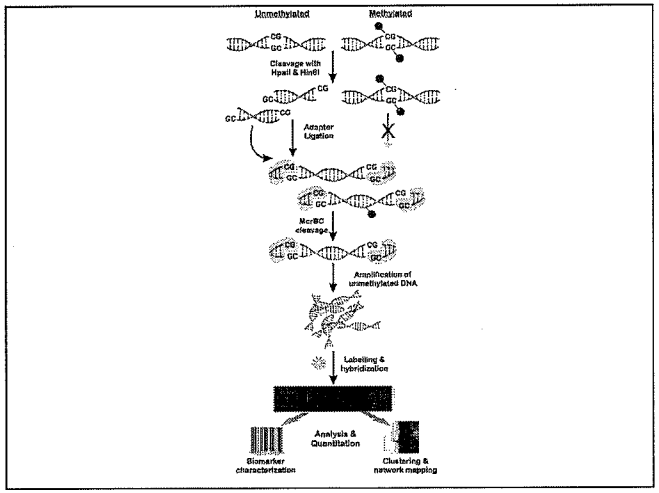
Reduction of genome complexity

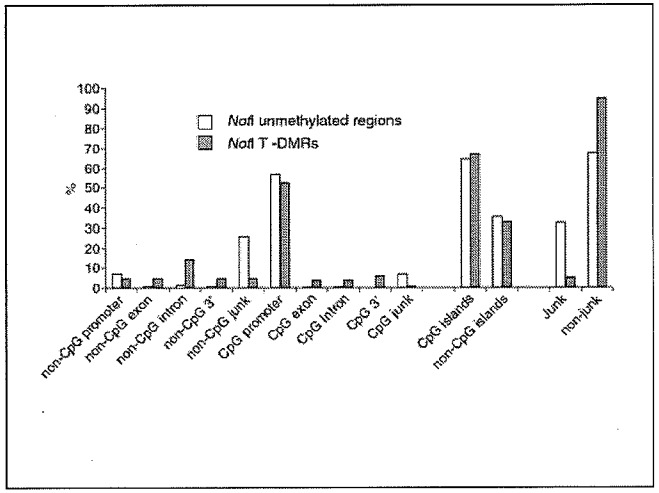
- Restriction enzymes followed by size fractionation
  - *Hpa*II, *Msp*I treatment; discard or use low MW fragments
- Affinity purification
  - E.g., Me-DIP or MBD protein
  - Possibly followed by bisulfite conversion
- Hybridization to microarray
  - E.g., capture targeted portion of genome on a microarray > elute from array > RE or bisulfite conversion > down stream assay (eg, sequencing)
- Solution-based capture
  - E.g., capture targeted portion of genome using "liquid array" > elute from oligos > bisulfate sequencing

Sample Preparation With Restriction Enzymes

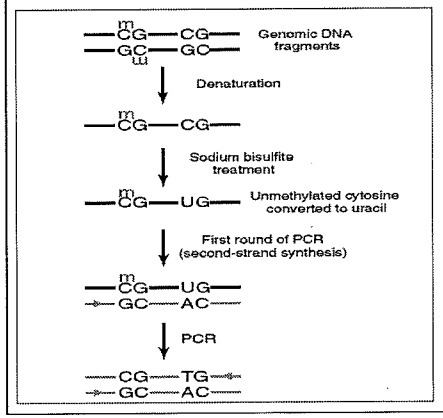
- Reduction of complexity using size fractionation







Sample Preparation with Sodium Bisulfite



- Most popular method
- Less sample input
- Suitable for many down-stream assay platforms.

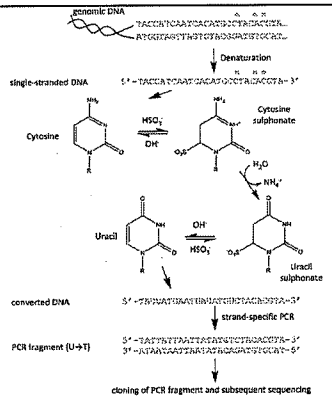


Fig. 29.1. Scheme of bisulphite conversion. Unmethylated cytosine residues of single-stranded DNA are sulphonated, followed by deamination to uracil sulphonate and desulphonation to uracil. Methylated cytosine residues are protected from this reaction. As the converted DNA is not complementary any more, the PCR has to be strand specific.

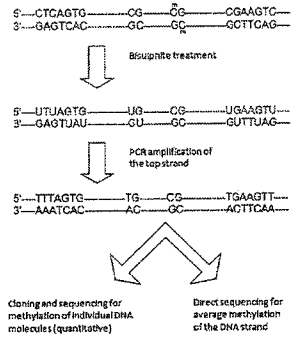


Figure 17.2 Bisulfite sequencing. Bisulfite treatment transforms the information of CpG methylation into nucleotide sequence information (C to T changes), which is easy to detect. After bisulfite conversion, the top and bottom DNA strands have different nucleotide sequence and will be amplified independently in the subsequent polymerase chain reaction (PCR) (amplification of the top strand is shown here). During the PCR all uracils are amplified as thymine by DNA polymerase. Amplified products can be sequenced directly or cloned prior to sequencing.

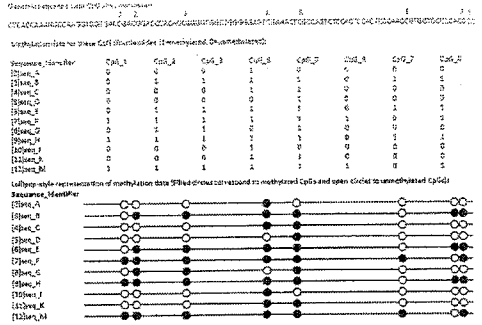
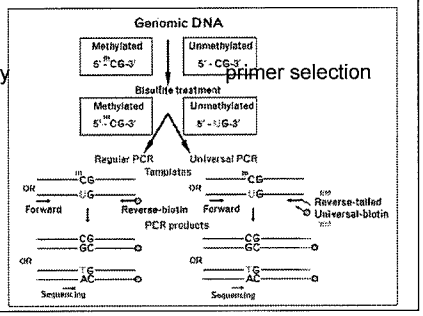
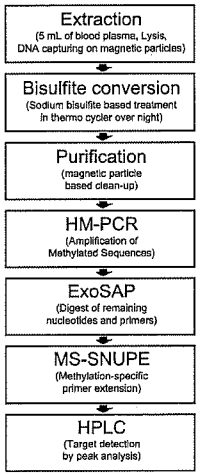


Figure 17.3 Clonal bisulfite sequencing data presentation by BiQ Analyzer. Genomic sequence of the analysed DNA fragment is given with the target CpG sites numbered. Derived methylation data are presented in table format where each row corresponds to each individual DNA sequence and each column to CpG site numbered according to the above sequence. Each data point is given either a value of 1 if methylated or 0 if unmethylated. Lower part of the figure shows a visual representation of methylation as lollipop diagrams. Each row corresponds to each individual DNA sequence, and black (methylated) and white (unmethylated) circles represent each CpG site showing the methylation status of the corresponding CpG site.

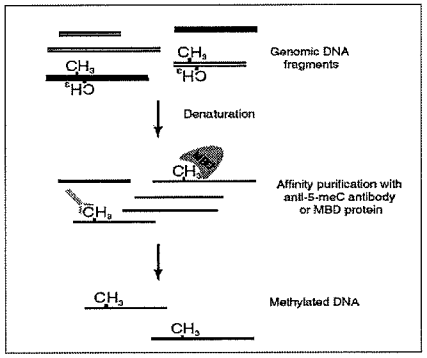
DNA Methylation Profile by Pyrosequencing

- Bisulfite DNA treatment
- Amplification of selected genes by PCR
- Measure methylation by sequencing-by-synthesis
- + Cheap
- + High resolution
- - Regions limited by primer selection





Sample Preparation with Anti- ^{5}mC antibodies

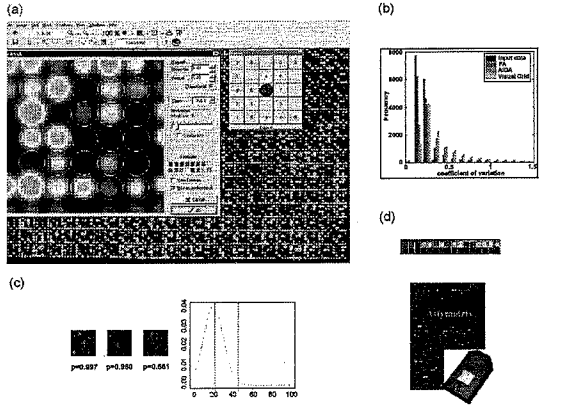


The most effective way to enrich methylated DNA and reduce complexity

Table 1 | **Main principles of DNA methylation analysis**

Pretreatment	Analytical step			
	Locus-specific analysis	Gel-based analysis	Array-based analysis	NGS-based analysis
Enzyme digestion	• *Hpa*II-PCR	• Southern blot • RLGS • MS-AP-PCR • AIMS	• DMH • MCAM • HELP • MethylScope • CHARM • MMASS	• Methyl-seq • MCA-seq • HELP-seq • MSCC
Affinity enrichment	• MeDIP-PCR		• MeDIP • mDIP • mCIP • MIRA	• MeDIP-seq • MIRA-seq
Sodium bisulphite	• MethyLight • EpiTYPER • Pyrosequencing	• Sanger BS • MSP • MS-SNuPE • COBRA	• BiMP • GoldenGate • Infinium	• RRBS • BC-seq • BSPP • WGSBS

AIMS, amplification of inter-methylated sites; BC-seq, bisulphite conversion followed by capture and sequencing; BiMP, bisulphite methylation profiling; BS, bisulphite sequencing; BSPP, bisulphite padlock probes; CHARM, comprehensive high-throughput arrays for relative methylation; COBRA, combined bisulphite restriction analysis; DMH, differential methylation hybridization; HELP, *Hpa*II tiny fragment enrichment by ligation-mediated PCR; MCA, methylated CpG island amplification; MCAM, MCA with microarray hybridization; MeDIP, mDIP and mCIP, methylated DNA immunoprecipitation; MIRA, methylated CpG island recovery assay; MMASS, microarray-based methylation assessment of single samples; MS-AP-PCR, methylation-sensitive arbitrarily primed PCR; MSCC, methylation-sensitive cut counting; MSP, methylation-specific PCR; MS-SNuPE, methylation-sensitive single nucleotide primer extension; NGS, next-generation sequencing; RLGS, restriction landmark genome scanning; RRBS, reduced representation bisulphite sequencing; -seq, followed by sequencing; WGSBS, whole-genome shotgun bisulphite sequencing.



Microarray-based Methylation assays: Array type is a key consideration

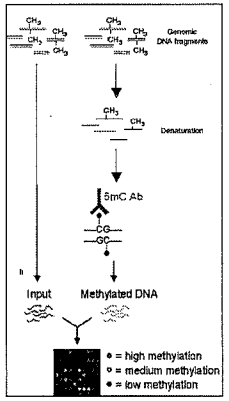
- **SNP genotyping arrays**
  - Screening assay uses methylation-sensitive enzymes > hyb to Affy SNP chips sets.. FFS at SeqWright. Cancer Res 66, 3443 (2006)
- **Promoter arrays**
  - Illumina
    - Cancer Panel I: 1,505 CpG loci from 807 genes; Golden Gate assay
    - Humanmethylation27 array: 27,578 promoter CpG sites
  - Affymetrix
    - GeneChip® Human Promoter 1.0R Array - designed for ChIP experiments
    - >4.6 million probes tiled through over 25,500 human promoter regions.
    - Average resolution of 35 bp. Approx 7.5 kb upstream through 2.45 kb downstream of 5' transcription start sites
  - NimbleGen
    - RefSeq Promoters; single array; probes with ~ 100bp spacing
    - Two-Array Sets; for splice variants and alternative transcription start sites.
    - RefSeq XM Promoters – with predicted transcripts
    - CpG Island-Plus-Promoter Arrays
- **Tiling arrays**
  - Affymetrix
    - GeneChip® Human Tiling 1.0R Array Set --designed for transcript mapping
    - >45 million probes, 14-array set covers entire genome.
    - Average resolution of 35 bp.
  - NimbleGen
    - 7-10 array set tiling the genome at an average probe spacing of 100bp or less

ChIP-chip

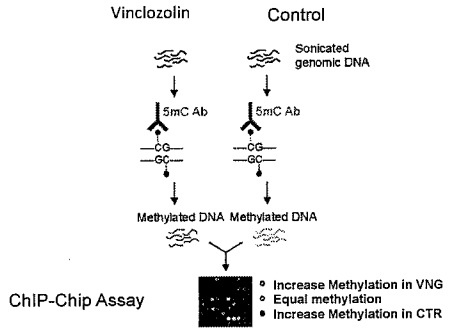
A high-throughput experimental technique that combines chromatin immunoprecipitation (ChIP) and microarray technology (chip) that directly identifies protein-DNA interactions.

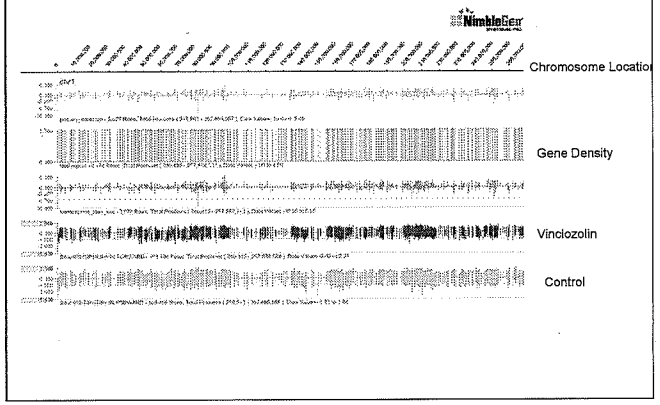
DNA Methylation Assay on NimbleGen or Aglient Microarrays

- Ratio-based Me-DIP method
- One channel: untreated; other channel IP' ed with anti-^{5}mC Ab
- Array can be a promoter or tiling array
- Arrays are 60-mers; relatively inexpensive custom arrays
- Assay and array platform has several hundred bp resolution

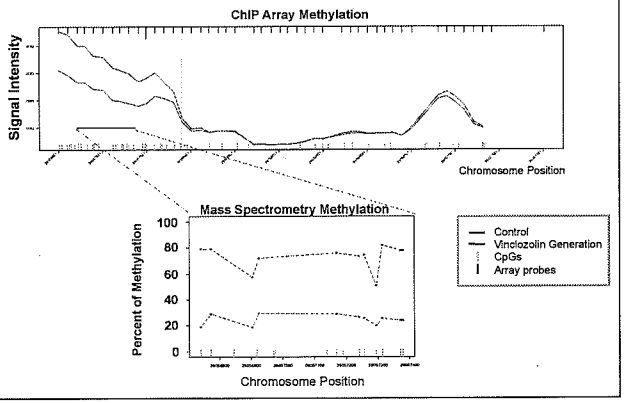


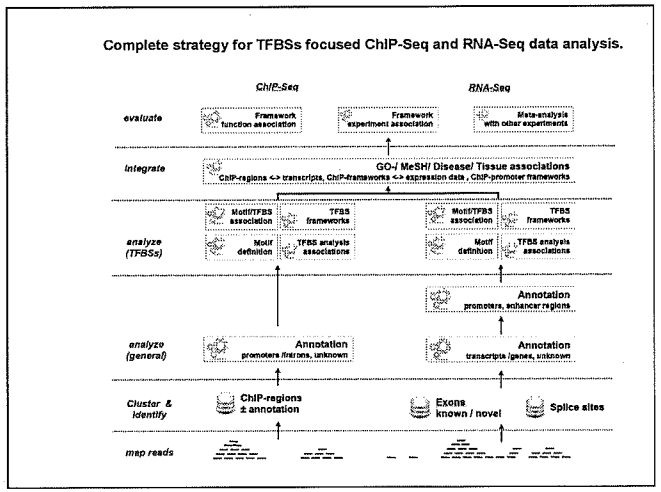
Comparative Methylation, MeDIP Chip, F3 Generation Sperm DNA pools





Methylation Olr735

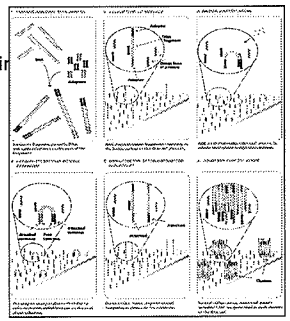


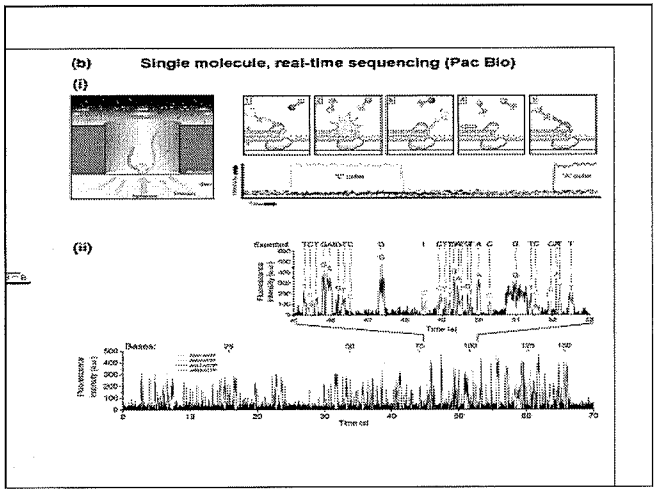


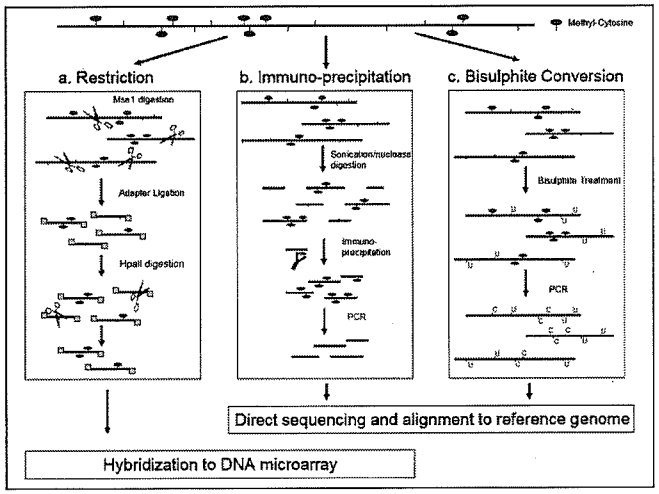
DNA Methylation Assay by Solexa Sequencing

- Bisulfite DNA treatment
- Enrichment by methylation, size selection, SBC-like approaches
- Possible targeted sequences
- Measure methylation by sequencir
- Likely direction of RNA profiling, profiling

Genomic DNA → Bisulfite treatment







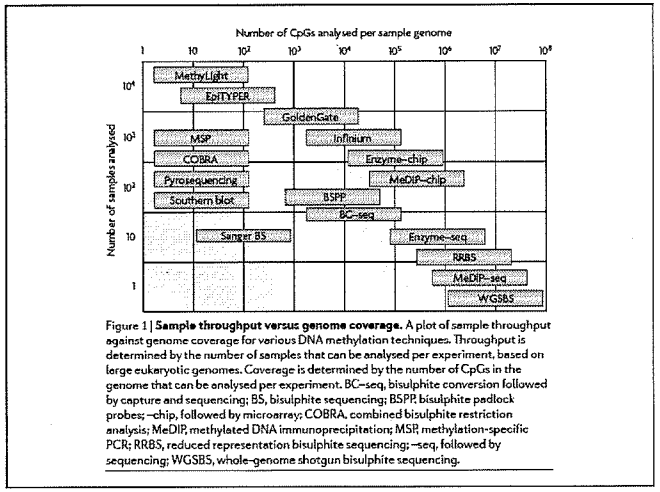
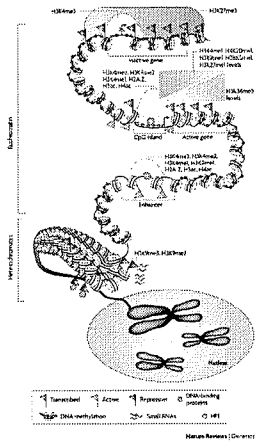


Figure 1 | **Sample throughput versus genome coverage.** A plot of sample throughput against genome coverage for various DNA methylation techniques. Throughput is determined by the number of samples that can be analysed per experiment, based on large eukaryotic genomes. Coverage is determined by the number of CpGs in the genome that can be analysed per experiment. BC–seq, bisulphite conversion followed by capture and sequencing; BS, bisulphite sequencing; BSPP, bisulphite padlock probes; –chip, followed by microarray; COBRA, combined bisulphite restriction analysis; MeDIP, methylated DNA immunoprecipitation; MSP, methylation-specific PCR; RRBS, reduced representation bisulphite sequencing; –seq, followed by sequencing; WGSBS, whole-genome shotgun bisulphite sequencing.

Epigenetic Mechanisms of Gene Regulation

- -DNA Methylation
- -Histone Modification
- -Chromatin Structure
- -DNA Organization into Domains (eg Loops)
- -Nuclear Compartmentalization (eg nuclear matrix)
- -Noncoding functional RNAs
- -RNA Methylation

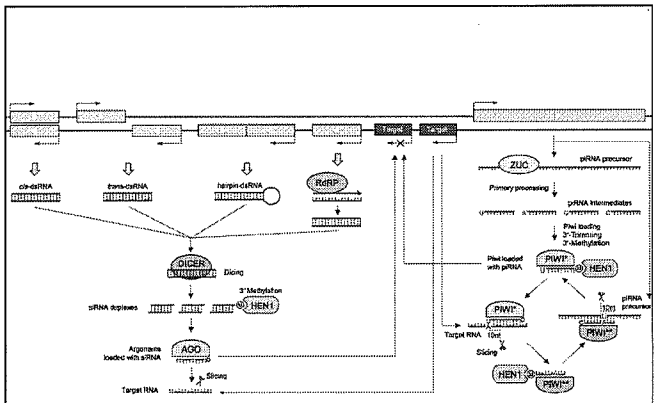


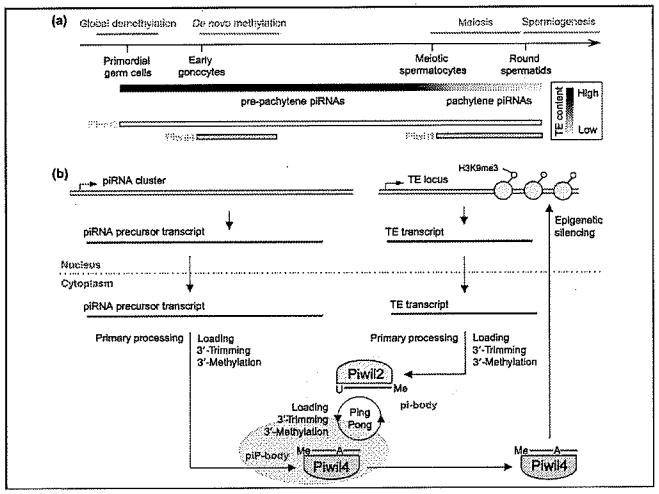
CRISPR-Mediated Epigenome Editing.
Enríquez P.
Yale J Biol Med. 2016 Dec 23;89(4):471-486.

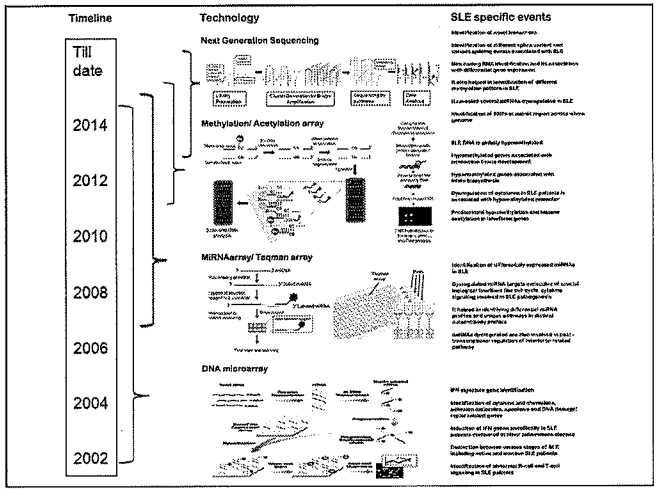
Mounting evidence has called into question our understanding of the role that the central dogma of molecular biology plays in human pathology. The conventional view that elucidating the mechanisms for translating genes into proteins can account for a panoply of diseases has proven incomplete. Landmark studies point to epigenetics as a missing piece of the puzzle. However, technological limitations have hindered the study of specific roles for histone post-translational modifications, DNA modifications, and non-coding RNAs in regulation of the epigenome and chromatin structure. This feature highlights CRISPR systems, including CRISPR-Cas9, as novel tools for targeted epigenome editing. It summarizes recent developments in the field, including integration of optogenetic and functional genomic approaches to explore new therapeutic opportunities, and underscores the importance of mitigating current limitations in the field. This comprehensive, analytical assessment identifies current research gaps, forecasts future research opportunities, and argues that as epigenome editing technologies mature, overcoming critical challenges in delivery, specificity, and fidelity should clear the path to bring these technologies into the clinic.

Strategies for precision modulation of gene expression by epigenome editing: an overview.
Laufer BI, Singh SM.
Epigenetics Chromatin. 2015 Sep 17;8:34.

Genome editing technology has evolved rather quickly and become accessible to most researchers. It has resulted in far reaching implications and a number of novel designer systems including epigenome editing. Epigenome editing utilizes a combination of nuclease-null genome editing systems and effector domains to modulate gene expression. In particular, Zinc Finger, Transcription-Activator-Like Effector, and CRISPR/Cas9 have emerged as modular systems that can be modified to allow for precision manipulation of epigenetic marks without altering underlying DNA sequence. This review contains a comprehensive catalog of effector domains that can be used with components of genome editing systems to achieve epigenome editing. Ultimately, the evidence-based design of epigenome editing offers a novel improvement to the limited attenuation strategies. There is much potential for editing and/or correcting gene expression in somatic cells toward a new era of functional genomics and personalized medicine.







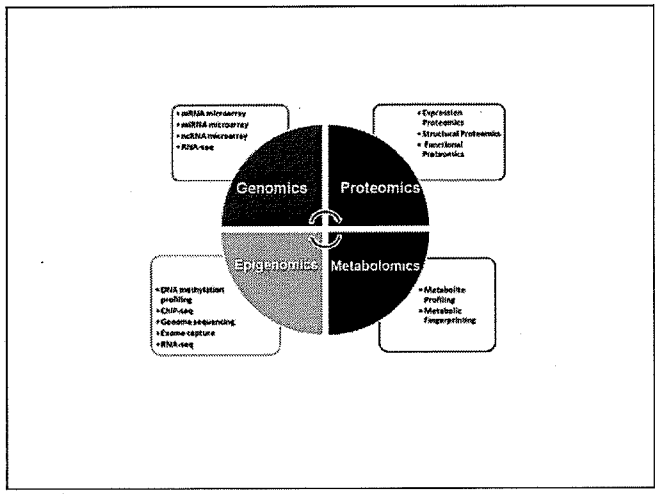


Table 1. Resources for processing and analyzing genomics data

	Software tools	Microarray Analysis
1.	(DAVID) v.6.7	http://david.abcc.ncifcrf.gov/
2.	GenePattern 2.0	http://www.broadinstitute.org/cancer/software/genepattern/
3.	Galaxy Project	http://galaxyproject.org/
4.	GenomeSpace	http://www.genomespace.org
5.	Bioconductor	http://www.bioconductor.org/
6.	Cytoscape	http://www.cytoscape.org/
7.	Ingenuity Pathways Analysis	www.Ingenuity.com
8.	GeneGO, MetaCore	http://portal.genego.com/
9.	Partek 6.3	www.partek.com
	Software tools	RNA-seq Analysis
1.	Galaxy Project	http://galaxyproject.org/
2.	GenePattern 2.0	http://www.broadinstitute.org/cancer/software/genepattern/
3.	R-workbench	http://www.ebi.ac.uk/Tools/rcloud/
4.	Palmapper	http://raetschlab.org/suppl/palmapper
5.	mTim	http://raetschlab.org/suppl/mtim
6.	rQuant	http://raetschlab.org/suppl/rquant
7.	rDiff	http://cbio.mskcc.org/public/raetschlab/user/drewe/rdiff/

Table 2. List of databases and tools for analyzing miRNA

Databases	Web Links
1. miRBase	http://www.mirbase.org/
2. miRNAmap	http://mirnamap.mbc.nctu.edu.tw/
3. miRWalk	http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/index.html
4. miRDB	http://mirdb.org/miRDB/
5. Tarbase V.5c	http://diana.cslab.ece.ntua.gr/tarbase/
6. mirGen	http://www.diana.pcbi.upenn.edu/miRGen.html
7. miROrtho	http://cegg.unige.ch/mirortho
8. mir2disease	http://www.mir2disease.org/
9. miRanda	http://www.microrna.org/microrna/home.do
10.TargetScan 6.2	http://www.targetscan.org/
11.miRanda	http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/
12.DIANA microT	www.microrna.gr/microT
13.PicTar	http://pictar.mdc-berlin.de/
14.miRecords	http://mirecords.umn.edu/miRecords/
15.mirTarBase	http://mirtarbase.mbc.nctu.edu.tw/
16.miRWalk	http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/
17.miRNAMap	http://mirnamap.mbc.nctu.edu.tw/

Table 3. Protein-Protein Interactions Databases

Databases	Web Links
1.STRING 9.05	http://string-db.org
2.Molecular Interaction Database (MINT)	http://mint.bio.uniroma2.it/mint/Welcome.do
3.Database of Interacting Proteins (DIP)	http://dip.doe-mbi.ucla.edu/dip/Main.cgi
4.IntAct	http://www.ebi.ac.uk/intact/
5.HPRD	http://www.hprd.org/
6.BioGRID 3.2	http://thebiogrid.org/
7.PrePPT	http://bhapp.c2b2.columbia.edu/PrePPI/

Table 5. Software tools for epigenomics data processing and analyses

Software Tools	Methylation Analysis
1. Batman	http://td-blade.gordon.cam.ac.uk/software/batman
2. BDPC	http://bioinfo.iicb.res.in/BDPC
3. Epinexus	http://epinexus.net/home.html
4. MetDB	http://www.methdb.de
5. methPrimerDB	http://medgen.ugent.be/methprimerdb
6. MethyCancer Database	http://methycancer.psych.ac.cn
7. Methyluml	http://www.bioconductor.org/packages/bioc/html/methyluml.html
8. mPod	http://www.compbio.group.cam.ac.uk/Projects/pMeth.html
9. PubMeth	http://www.pubmeth.org
10.QUMA	http://quma.cdb.riken.jp
11.TCGA Data Portal	http://cancergenome.nih.gov/dataportal
Software Tools	**Bisulphite Sequencing Analysis**
1. BSMAP	http://code.google.com/p/bsmap
2. MethBLAST	http://medgen.ugent.be/methBLAST
3. methPrimer	http://www.urogene.org/methprimer
4. Methyl Primer Express	http://www.appliedbiosystems.com/methylprimerexpress
5. CpGviewer	http://dna.leeds.ac.uk/cpgviewer
Software Tools	**CpG Island Analysis**
1. CpGcluster	http://bioinfo2.ugr.es/CpGcluster
2. CpGFinder	http://linux1.softberry.com
3. CpG Island Explorer	http://bioinfo.hku.hk/cpgiexplorer.html
4. CpG Island Explorer	http://cpgislands.usc.edu
5. CpG Promoter	http://www.cshl.edu/OTT/html/cpg_promoter.html
6. CpG ratio/GC content Plotter	http://mwsross.bmsad.ac.uk/public/cgi-bin/cmpsl
7. EMBOSS CpGPlot/CpGReport/Isochore	http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html

Table 6. Software tools for analyzing and interpreting metabolomics data

Software Tools	Web Links
1. BioSpider	http://www.biospider.ca/
2. COLMAR	http://spinportal.magnet.fsu.edu/
3. FID	http://www.cs.helsinki.fi/group/sysfys/software/fraqid/
4. HORA suite	http://www.pasteurnetlab.org/
5. MeRDB 2.0	https://merdb.cebitec.uni-bielefeld.de/cgi-bin/login.cgi
6. MetaboAnalyst 2.0	http://www.metabolanalyst.ca/MetaboAnalyst/faces/Home.jsp
7. MetaboMiner	http://wishart.biology.ualberta.ca/metabominer/
8. MolFind	http://metabolomics.pharm.uconn.edu/Software.html
9. OpenMS 1.11.1	http://open-ms.sourceforge.net/openms-1-11-1-released/
10. SetupX	http://fiehnlab.ucdavis.edu/projects/binbase_setupx/
11. Seven Golden Rules	http://fiehnlab.ucdavis.edu/projects/Seven_Golden_Rules/Software/
12. XCMS	http://metlin.scripps.edu/xcms/
13. MSEA	http://www.msea.ca
14. MBRole	http://csbg.cnb.csic.es/mbrole
15. MPEA	http://ekhidna.biocenter.helsinki.fi/poxo/mpea/
16. IMPaLA	http://impala.molgen.mpg.de/
17. BioCyc - Omics Viewer	http://biocyc.org
18. MetPA	http://metpa.metabolomics.ca
19. Reactome	http://www.reactome.org
20. Cytoscape	http://www.cytoscape.org/