

# Chapter 2

## The Development of Epigenetics in the Study of Disease Pathogenesis



Matlock A. Jeffries

**Abstract** The study of epigenetics has its roots in the study of organism change over time and response to environmental change, although over the past several decades the definition has been formalized to include heritable alterations in gene expression that are not a result of alterations in underlying DNA sequence. In this chapter, we discuss first the history and milestones in the 100+ years of epigenetic study, including early discoveries of DNA methylation, histone posttranslational modification, and noncoding RNA. We then discuss how epigenetics has changed the way that we think of both health and disease, offering as examples studies examining the epigenetic contributions to aging, including the recent development of an epigenetic “clock”, and explore how antiaging therapies may work through epigenetic modifications. We then discuss a nonpathogenic role for epigenetics in the clinic: epigenetic biomarkers. We conclude by offering two examples of modern state-of-the-art integrated multi-omics studies of epigenetics in disease pathogenesis, one which sought to capture shared mechanisms among multiple diseases, and another which used epigenetic big data to better understand the pathogenesis of a single tissue from one disease.

**Keywords** Epigenetics · History of epigenetics · Integrative analysis

### 2.1 The History and Discovery of the Basic Mechanics of Epigenetics

#### 2.1.1 *Experimental Observations: Setting the Stage for the Discovery of Epigenetic Mechanisms*

To set the stage for our later discussion of the epigenetic underpinnings of disease pathogenesis, we will first review the history of the discovery of epigenetic control

---

M. A. Jeffries (✉)  
University of Oklahoma Health Sciences Center, Oklahoma Medical Research Foundation,  
Oklahoma City, OK 73118, USA  
e-mail: [Matlock-Jeffries@ouhsc.edu](mailto:Matlock-Jeffries@ouhsc.edu)

© Springer Nature Singapore Pte Ltd. 2020  
C. Chang and Q. Lu (eds.), *Epigenetics in Allergy and Autoimmunity*,  
Advances in Experimental Medicine and Biology 1253,  
[https://doi.org/10.1007/978-981-15-3449-2\\_2](https://doi.org/10.1007/978-981-15-3449-2_2)

mechanisms and the field of epigenetics more generally. *Epigenetics* in its contemporary form, the focus of this book, has only emerged since the 1990s; in fact, the consensus modern definition, “stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence” was formulated at a Cold Spring Harbor meeting in 2008 (Berger et al. 2009). Prior to this, the term “epigenetic” was used with widely disparate meanings. The word itself has been in English language use since the seventeenth century, having been used by Harvey in *Exercitationes* (1651) and *Anatomical Exercitationes* (1653), with the definition “the additament of parts budding one out of another.” The subsequent history of epigenetics as we presently understand it is inevitably linked with the history of genetics, itself emerging out of historical analyses of evolution, development, and inheritance.

No discussion of the history of epigenetics is complete without reference to the theories of the French zoologist Jean-Baptiste Lamarck (Fig. 2.1). At half-past noon on the May 17, 1802, he gave the first lecture of his course on invertebrate zoology



**Fig. 2.1** Jean-Baptiste Lamarck. *Credit* Wellcome Library, London, Wellcome Images

at the Museum of Natural History in Paris, where he set forth an unprecedented (and later debunked) theory explaining the incremental development and differentiation of all life forms on earth. His theories, now known simply as “Lamarckian evolution”, were based on two major themes. First, he posited that the environment gives rise to changes in animals; he offered as examples the presence of teeth in mammals, the absence of teeth in birds, and the loss of sight in moles. The second principle argued that life was structured in an orderly manner, and that different parts of all bodies made possible the organic movement of animals (Eb 1895). Lamarck’s ideas on evolution are held as grand examples of theories of inheritance of acquired characteristics, also known as “soft inheritance”, later supplanted by the infamous descriptions of inheritance and natural selection put forth by Gregor Mendel and Charles Darwin, respectively, in the mid-1800s. Although discredited as a viable theory of evolution throughout the twentieth century, Lamarck’s underlying ideas regarding environmental contributions to phenotypic characteristics (and the heritability thereof) have seen a resurgence in modern times with recent descriptions of transgenerational epigenetic inheritance.

The first concepts we would recognize as related to modern epigenetics were formulated in the late nineteenth century with Fleming’s discovery of chromosomes and their behavior during animal cell division in 1879. Subsequent work by a variety of investigators cemented the notion of chromosomes as the quintessential substrate of inherited information. Most convincingly, in 1911 Thomas Hunt Morgan demonstrated sex–chromosome genetic linkage of several *Drosophila* genes (Morgan 1911), which was quickly followed by the first maps of individual genes arranged in a linear fashion occupying specific locations on each chromosome (Sturtevant 1913). Questions remained, however, regarding the ways in which this “hard” genetic information was used to direct the developmental programs which lead to cellular differentiation and the diversity of phenotypes seen in multicellular life forms. It quickly became apparent that chromosomal nucleic acid codes alone were inadequate to carry all developmental information, since similar amounts and organization of this material was present in widely disparate cell types.

As an example, Hermann Joseph Muller in 1930 described a class of *Drosophila* mutations which involved the displacement of large portions of genetic material from its rightful place to another chromosome (a phenomenon we would later understand as balanced translocations) (Muller 1930) resulting in unexpected phenotypic variations; particularly, the observation of mottled eyes. He initially attributed this to “genetic diversity of the different eye-forming cells,” but later research including Hannah et al. in the 1950s went on to demonstrate that this variegation arose when rearrangements inserted particular genes into regions of the chromosome with variable staining density. These regions would come to be known as heterochromatin, which we now understand to be differentiated from evenly staining chromosomal regions (euchromatin) by the density of DNA-packaging elements, directed by epigenetic modification (Hannah 1951). It was about this time (1939) when Conrad Waddington, the Buchanan Professor of Genetics at Edinburgh University, altered the Greek word *epigenesis*, defined as a theory of development in which the early embryo

existed in an undifferentiated state, to *epigenetics*, which he defined as the implementation of cellular genetic programming for development (Waddington 1939). This view was reflected in his later book *The Epigenetics of Birds*, which provides an account of the embryological development of chicks (Waddington 1952). He is also credited with the term *epigenotype*, which he defined as “the total developmental system consisting of interrelated developmental pathways through which the adult form of an organism is realized” (Waddington 1939). Although admittedly broad, this definition strikes a familiar chord when applied to our modern understanding of epigenetics, where, at least at a cellular level, epigenetic controls may indeed be well defined as the set of interrelated transcriptional control programs which dictate and ensure the proper terminal differentiation of cells in a particular tissue.

Simultaneous to and along the same lines as Waddington, seminal works were published establishing and refining the ideas behind phenotypic divergence during organism development at a cellular level. It became evident, for example, that phenotypically distinguishing features, once established, were clonally inherited to daughter cells. The mechanisms underlying these observations were unclear. Initially, it was proposed that these phenotypes were maintained by the presence or accumulation of particular biochemical reactions, suggested by the physicist Max Delbruck in 1949. In this model, self-stabilizing stimulatory and inhibitory mechanisms could theoretically lead to substantial phenotypic divergence, which could be passed on from parent to daughter cell. This theory ran into a problem, however, with the discovery that protein-free DNA could carry this information, first by Avery et al. in 1944 (the “Transforming Principle”) (Avery et al. 1944), later by Hershey and Chase in 1952 (Hershey and Chase 1952), and further reinforced with the infamous description of the structure of DNA by Crick at Cold Spring Harbor in 1953 (Watson and Crick 1953). At the same time, a pivotal study in northern leopard frogs by Robert Briggs and Thomas King was published (Briggs and King 1952) in which they demonstrated that a nuclear transplant from a frog blastula into an enucleated frog egg cell results in the normal development of an embryo, which argued that the “essential material” for complex organism development was indeed contained in the nucleus. Although discoveries since that time have cemented the concept of a fully differentiated cell having the same genetic makeup as an embryonic cell, this was not self-evident until as late as 1970. It was in this year that a pivotal study done by Laskey and Gurdon demonstrated the transfer of nuclear material from several adult somatic *Xenopus* frog cell sources (around 1–2% of transplanted nuclei, likely representing endogenous tissue pluripotent cells) into enucleated eggs resulted in embryos that developed into feeding larval stages (Laskey and Gurdon 1970). Other work soon revealed that cellular fates and phenotypic characteristics were heritable throughout cellular differentiation. For example, Ernst Hadorn in 1965 showed that imaginal disc cells from *Drosophila* (regions of embryonic tissue present in fly larvae that subsequently develop into well-defined adult structures; i.e., two per wing, etc.), were still able to develop into their requisite structures on transplantation into fly larvae, even after being passaged through multiple generations of adult flies (Hadorn 1965).

### 2.1.2 *Getting to the Substance of the Matter: Discoveries of Physical Mechanisms of Epigenetic Control*

Building on prior observations, a number of scientists made key discoveries in the mid-twentieth century that would help explain *how* these cellularly heritable gene expression patterns were made possible. First were investigations into the mechanisms underlying X-chromosome inactivation, the process by which one of two X chromosomes in females is randomly selected to more or less fully transcriptionally “shut down”. Ohno demonstrated in 1959 that the “sex chromatin”, a female-specific chromosome first discovered by Barr and Bertram (1949), and still commonly bearing the eponymous title Barr Body, was, in fact, an X chromosome which had been densely packaged (Ohno et al. 1959). Lyon and colleagues in 1961 went on to demonstrate that the selection of X chromosome for condensation into a Barr Body from the paternal or maternal genetic material was a random process, providing further evidence that the *mechanism* of inactivation of genetic material was the most important component, not any preexisting underlying difference in genetic code (Lyon 1961).

DNA-binding proteins, histones chief among them, had long been recognized as associated with nucleic acids, having first been described by Kossel (1884), which, incidentally, earned him the 1910 Nobel Prize in Physiology or Medicine. The husband-and-wife research team Steadman and Steadman in 1950 described variations in histone proteins isolated from different tissues, initially incorrectly ascribing these differences to variation in their constituent amino acids, but nonetheless correctly surmising the importance of what would later be defined as histone post-translational modifications in contributing to the regulation of gene transcription (Stedman and Stedman 1950). In their letter to *Nature* in 1950, they stated:

It has always been a puzzle to us, how the physiological functions of the cell nuclei in the same organism can differ from one cell-type to another when they all contain identical chromosomes and hence identical genes. ...The demonstration...that some of the basic proteins present in cell nuclei are certainly cell-specific leads to the hypothesis that one of their physiological functions is to act as gene suppressors. (Steadman and Steadman 1950)

Soon though, the field of gene transcription shifted toward *cis*-acting regulatory elements and transcription factors (see below) It was thereafter generally assumed that histones were passively acting packaging proteins that served only to suppress gene transcription, even though evidence existed that extensive genomic regions with an open chromatin conformation did not, in fact, exist (Clark and Felsenfeld 1971). Further disagreement came with the description of posttranslational modifications of histone proteins were in 1964 by Mirsky and Allfrey. Importantly, they demonstrated that acetylated histones were associated with more active gene transcription than their non-acetylated counterparts (Allfrey et al. 1964), suggesting that certain histones were required for gene activation, rather than a straightforward removal of histone proteins. Several additional posttranslational modifications of proteins were discovered in the ensuing years, including methylation and phosphorylation, although their associated cellular functions remained elusive. It would not be until the

mid-1990s that the binding of transcriptional regulators to post-translationally modified histone proteins would elevate their status to bona fide epigenetic modifications (Hecht et al. 1995). Among the seminal works at this time, Taunton and colleagues in 1996 demonstrated that a mammalian histone deacetylase was closely related to a protein previously described as a transcriptional repressor in yeast (Taunton et al. 1996). Furthermore, Brownell et al. in 1996 showed that a histone acetyltransferase from the ciliate protozoan *Tetrahymena* was homologous to the yeast regulatory protein Gcn5 (Brownell et al. 1996), proving definitively that histone post-translational modification enzymatic activity was correlated with gene transcription regulation.

The regulatory potential of modifications of the DNA molecule itself was put forth by Holliday and Pugh (1975) and Riggs (1975) in 1975, who proposed (correctly, in retrospect) that methylation of DNA nucleotides could at least partially account for Barr Body chromosomal inactivation. It is important to note here that methylation of cytosine DNA residues had been previously reported in the literature more than 20 years earlier in 1951, by Wyatt et al., although the function of the modification was unknown at the time (Wyatt 1951). In a remarkably prescient description of what would later come to be identified as an epigenetic control mechanism, Holliday and Pugh wrote:

Since the ultimate control of development reside in the genetic material, the actual program must be written in base sequences in the DNA. It is also clear that cytoplasmic components can have a powerful or overriding influence on genomic activity in particular cells, yet these cytoplasmic components are, of course, usually derived from the activity of genes at some earlier stage of development. A continual interaction between cytoplasmic enzymes and DNA sequences is an essential part of the model to be presented. (Holliday and Pugh 1975)

They further correctly hypothesized that the chemical modification of DNA by methylation at cytosine residues in cytosine–guanosine (CpG) pairs is added by both a set of de novo enzymes (the DNA methyltransferases DNMT3a and DNMT3b) and a separate maintenance enzyme required to copy epigenetic marks to daughter cells during division (the DNA methyltransferase DNMT1), first described some 20 years later in 1982 and 1983, respectively (Jähner et al. 1982; Bestor and Ingram 1983). Furthermore, they also correctly pointed out that these modifications are reversible, a necessary function for any organism to react in a bidirectional fashion to environmental manipulation, although the enzymes responsible for DNA demethylation were not described and validated until some three decades later (Tahiliani et al. 2009; Ito et al. 2010; Ko et al. 2010). These observations gave the first depictions of sequence-specific epigenetic mechanisms by which transcriptional patterns could be set and transmitted through cellular generations.

It is here that we first encounter evidence of the nascent field of epigenetics intersecting with studies of human disease pathogenesis. In his 1979 paper entitled “A New Theory of Carcinogenesis”, Holliday argued that repair of damage to DNA induced by known carcinogens could lead to loss of epigenetic information; specifically, that “replacement” DNA would lose any preexisting methylation marks and, if in the right gene, might partially explain tumorigenesis (Holliday 1979). Although not entirely accurate, an epigenetic basis for malignancy, and epigenetic aberrations within known oncogenes and near chronic-disease-associated susceptibility alleles

is now a well-studied phenomenon. There were two lines of evidence then uncovered that unequivocally linked methylation of cytosine DNA residues to gene expression and confirmed previous hunches about DNA methylation and X-chromosome inactivation. The first was the discovery of bacterial restriction enzyme pairs (known as isoschizomers) where one enzyme was methylation-sensitive; that is, one enzyme which would cut double-stranded DNA at a particular position regardless of the methylation status of an included cytosine, whereas another would cut at the same sequence only when cytosine residues were in a single methylation state. An example of this can be seen in the pair MspI and HpaII; both cut GCGC motifs, but Msp I requires the middle cytosine to be methylated. Several studies in the late 1970s and early 1980s soon leveraged this technique to demonstrate that genes with DNA methylation of their respective promoter regions were inactive, whereas unmethylated promoter regions were associated with active gene expression (Doerfler 1981). The second line of evidence came soon thereafter, with the discovery that the nucleoside analogue 5-azacytidine. When incorporated into cellular DNA, it inhibits the activity of DNA methyltransferases, leading to reductions in DNA methylation levels globally and subsequent increases in gene expression levels, including reactivation of genes located on the inactive X chromosome (Jones 1985).

An obvious question then became the focus of the epigenetics field: how are DNA methylation marks (or histone posttranslational modifications) targeted to the right genomic locations within a cell at the right time? This final fundamental topic in epigenetic regulation began to be answered in the late 1980 and 1990s, when intermediate factors that possess both DNA localization domains as well as binding sites for chromatin- or DNA methylation-altering enzymes were described. Of particular historical importance in this context included the description of the myogenic differentiation factor MyoD by Davis et al. (1987), Weintraub et al. (1989). MyoD is a transcription factor that binds to the enhancer regions of hundreds of myogenic genes and recruits various histone posttranslational modifiers, including p300, LSD1, G9a, and HDAC1; in this way, it functions as a “bridge” between short-term gene activation or repression and long-term epigenetic modification to produce a particular transcriptional phenotype driving myogenic differentiation. In 1994, Pazin and colleagues induced gene activation in *Drosophila* embryo chromatin extract by mixing with yeast activator GAL4, fused with the transcriptional activation region of the herpes virus protein VP16 (Pazin et al. 1994). They demonstrated that gene activation occurred via an ATP-dependent mechanism, and was accompanied by reorganization of nucleosomes within the chromatin. These findings would later be clarified with the discovery of nucleosome remodeling complexes, including SWI/SNF (Peterson and Herskowitz 1992) and NURF (Mizuguchi et al. 1997), which have subsequently been shown to bind themselves to other protein sequences (like Dri ARID) which provide sequence specificity to their gene transcription modulatory activity (Sun et al. 2015).

These discoveries have become particularly important over the past decade or two as investigators have begun identifying alterations in various epigenetic-targeting effectors in association with human disease. Take, for example, the description by Sawalha and colleagues in 2008 of genetic mutations in the methylcytosine-binding protein 2 (*MECP2*) gene which increase the risk of systemic lupus erythematosus

(Sawalha et al. 2008); MECP2 functions to bind methylated DNA and direct further histone posttranslational modification to “turn off” genes, as well as recruiting the maintenance DNA methyltransferase DNMT1 during cellular division, maintaining fidelity of epigenetic marks. Koelsch, Sawalha, and colleagues went on to show that overexpression of *MECP2* in a mouse model altered all three major domains of epigenetic control: DNA methylation, noncoding RNA (discussed in detail below), and histone posttranslational modification (Koelsch et al. 2013).

### 2.1.3 A New Epigenetic Role for RNA

We have until this point focused on DNA and protein as the central players in the epigenetic regulation space; however, we must not ignore the role of RNAs, which had a bit of a late start but have nonetheless more recently risen to epigenetics fame. Coincident with the discoveries of DNA in the mid-twentieth century, the “central dogma” of molecular biology emerged; that is, the concept that DNA encodes RNA which, in turn, is “read” by ribosomes to produce proteins. The first glimpse of a potential regulatory role of RNA was found in 1966, when Warner and colleagues identified a complex group of nuclear RNAs aside from what was then termed “ribosomal precursor” RNA, which they termed “heterogeneous nuclear RNA” or hnRNA (Warner et al. 1966). At the same time, researchers identified retrotransposons, large areas of repeated elements within DNA in multiple species across several domains. Integrating this information, Roy Britten and Eric Davidson published “Gene Regulation for Higher Cells: A Theory” in *Science Magazine*, which outlined their theory that higher organisms contained extensive RNA networks which regulated gene transcription (Britten and Davidson 1969). The next big discovery occurred in 1997 with the demonstration of noncoding regions interspersed among coding regions in messenger RNA, areas they termed “introns” (Berget et al. 1977; Chow et al. 1977). Unfortunately, it was erroneously assumed that the noncoding portions of these transcripts represented “junk” RNA, and little additional work was done on these noncoding RNA regions in the ensuing decade.

This thinking radically changed, however, with the landmark discovery in 1993 by Ambros and colleagues of the regulatory nature of small (~22 nucleotides in length) bits of RNA. Working in *C. elegans*, they found that the gene *lin-4* encoded two small RNAs, one 22 and the other 61 nucleotides, which bound to a repeated sequence element in the 3' untranslated region of *lin-14* mRNA, which codes the LIN-14 protein, critical for *C. elegans* development. They further demonstrated that this antisense RNA–RNA interaction was tightly inversely correlated with LIN-14 expression, suggesting that this small RNA was indeed regulating the gene (Lee et al. 1993). The world of noncoding RNA grew much larger with the discovery in 1998 of the RNA interference pathway in both *C. elegans* (Fire et al. 1998) and plant species (Waterhouse et al. 1998), discoveries which would lead to the awarding of the Nobel Prize for Physiology or Medicine to Fire et al. (1998). This RNA interference pathway had been suggested in plants by the process of transgene silencing,



a mechanism described in the early 90s which has been subsequently identified as RNA-directed DNA methylation (see the previous discussion of this epigenetic control) (van der Krol et al. 1990; Wassenegger et al. 1994). Scientists then identified the presence of double-stranded RNA, and the fact that these were processed into short fragments (small interfering RNA, or siRNA) upon viral infection of plant cells. Subsequently, endogenous dsRNA precursors were identified, as were the cellular proteins responsible for their processing (including Argonaute, Drosha, and Dicer) (Bernstein et al. 2001; Doi et al. 2003; Lee et al. 2003; Basyuk et al. 2003) in the early 2000s. Although it was initially thought that these noncoding RNAs functioned more or less exclusively by binding messenger RNA and targeting it for degradation (thereby downregulating gene expression), later studies demonstrated that certain siRNA processing machinery, particularly Argonaute, was expressed not only in the cytoplasm, but in the nucleus (Ahlenstiel et al. 2012; Ameyar-Zazoua et al. 2012). This curious observation led to later demonstrations that these short RNAs can direct other forms of epigenetic modulation within the nucleus; incidentally, by recruiting a variety of histone posttranslational-modification enzymes, including the aforementioned polycomb complex (Kim et al. 2006).

Over the past decade a large range of noncoding RNAs have been identified, all with distinct regulatory pathways, many of which involving direct epigenetic modulation of a target gene. For example, long noncoding RNAs are intricately linked with paternal imprinting (the process by which the redundant copies of each gene, either maternal or paternal in origin, are inactivated) (Sleutels et al. 2002; Thakur et al. 2004). Interestingly, X chromosome inactivation is also partially directed by the coating of the nascent chromosome in a short noncoding RNA known as the X-inactive specific transcript (X-ist), which, in turn, recruits Polycomb repressive complex 2 (PCRC2), which trimethylates lysine 27 of histone 3 (H3K27me3), a repressive histone mark. The expression of X-ist itself is epigenetically regulated by a variety of other lncRNAs (review, Lee and Bartolomei 2013). Several novel functions of miRNAs have been proposed, and several new classes of larger noncoding RNA with important transcriptional significance have been recently described, including Piwi-interacting RNAs (piRNAs), small nucleolar RNAs (snoRNAs), and others.

## 2.2 Epigenetics Moves Mainstream and Meets the Clinic

Now that we have outlined a historical perspective of the discovery of epigenetics, we will next move to describe how our knowledge of these fundamental mechanisms have expanded and, in some cases, radically changed our understanding both of normal cellular processes and of disease pathogenesis. As an example, we will first explore the epigenetic changes underlying the normal aging process, then introduce the concept of an epigenetic “clock”, a topic that has received much attention in the recent literature. We will then discuss the epigenetic effects of antiaging interventions. Epigenetics has not only been studied in the context of disease pathogenesis however; especially from a clinical standpoint, epigenetic biomarkers hold great

promise, and will be discussed. Finally, though this chapter focuses heavily on a historical view of epigenetics, we will then close with examples of how modern epigenetic, transcriptomic, and proteomic tools are being used along with modern computational approaches in the modern age to help make sense of the massive quantities of data being generated through increasingly detailed looks at each of these “-omics” domains.

### ***2.2.1 The Epigenetics of Aging***

We will start with the epigenetics of aging, a normal process that nonetheless contributes to the development of a large number of chronic diseases. As we have previously detailed, a robust and relatively error-free epigenetic regulatory system is required for the normal development of a multicellular organism, as tissue types arise from embryonic multipotent cells, and particular transcription patterns must be set epigenetically to allow for appropriate tissue differentiation throughout organism development and cellular replenishment or maintenance throughout life. One could imagine, then, that reductions of cellular “plasticity” caused by loss of or improper epigenetic regulation, could substantially affect many fundamental biological processes. There is now a large body of literature which supports the notion that epigenetic aberrancies occur as part of normal aging, and that these may predispose to the development of a variety of diseases. In this section, we will discuss a few of the general epigenetic modifications that have been associated with aging, and a relatively new epigenetic paradigm (the epigenetic clock) that may have important implications in our understanding of epigenetic aging and disease susceptibility.

Many of the epigenetic changes previously observed during aging are outlined in Table 2.1 and an overview of epigenetic changes in aging is presented in Fig. 2.2. A global example is generalized loss of histones, a feature which has been noted in organisms from yeast to humans (Dang et al. 2009; O’Sullivan et al. 2010). This loss of histones is associated with the total number of divisions a particular cell has undergone, similar to telomere shortening. The loss can be dramatic; for example, in aged yeast, ~50% of overall histone loss can be seen (Hu et al. 2014). Unsurprisingly, this histone loss renders DNA more accessible to transcription globally, and, in this example, led to the induction of all known yeast genes. Conversely, artificially increasing histone protein production results in extended lifespan in yeast (Freser 2010). Interestingly, during senescence, a pathogenic state in which cells cease to divide but continue to be metabolically active, and produce a variety of harmful inflammatory cytokines, histone genes undergo unique alternative splicing of their requisite genes to produce senescence-specific histone proteins (Rai et al. 2014).

Another epigenetically mediated mechanism underlying aging and age-related diseases are reductions in chromatin stability, which can be pathogenic in two fundamental ways. First, age-dependent decreases in chromatin stability, linked with reductions in histone proteins and appropriate chromatin packaging, can make the genome more susceptible to mutations. Additionally, less stable chromatin can reduce

**Table 2.1** Epigenetic changes during aging. Adopted from Benayoun et al. (2015)

Epigenetic mark	Function	Aging paradigm	Change in epigenetic mark with aging	Change observed in tissue or cells	Method of detection	Species
5-mC	Transcriptional repression/activation	Organismal (Werner syndrome)	No global change	MSCs	RRBS	<i>H. sapiens</i>
		Cellular (senescence)	Global decrease and local increases	Fibroblasts	Immunostaining WGBS Chromatography	<i>H. sapiens</i> <i>M. musculus</i> <i>M. auratus</i>
		Organismal	Minor global increase, local increases, and some local decreases	HSCs	RRBS WGBS	<i>M. musculus</i>
		Organismal	Local increases and some local decreases	Small intestine, colon, lung, liver, spleen, brain, blood, kidney and muscle	Pyrosequencing MCAM Beadchip arrays	<i>H. sapiens</i> <i>M. musculus</i>
		Organismal	Increase (LINEs), local decreases and rare local increases	Sperm	Pyrosequencing (LINEs) Beadchip arrays	<i>H. sapiens</i>
		Organismal	No global change	Cortex	WGBS	<i>M. musculus</i> <i>H. sapiens</i>

(continued)

Table 2.1 (continued)

Epigenetic mark	Function	Aging paradigm	Change in epigenetic mark with aging	Change observed in tissue or cells	Method of detection	Species
5-hmC	Transcriptional activation	Organismal	Minor increase (SINEs and LTRs)	Cerebellum	Immunostaining Chemical tagging and sequencing	<i>M. musculus</i>
		Organismal	Minor global decrease	HSCs	HPLC-MS	<i>M. musculus</i>
mCH (where H represents A, C, or T)	Unclear	Organismal	Minor global decrease	Cortex	WGBS	<i>H. sapiens</i> <i>M. musculus</i>
		Cellular (senescence)	Decrease	Fibroblasts	Western blot	<i>H. sapiens</i>
Histone H2A	Core chromatin component	Organismal (replicative lifespan)	Decrease	Yeast cells	Western blot	<i>S. cerevisiae</i>
		Cellular (senescence)	Decrease	Fibroblasts	Western blot	<i>H. sapiens</i>
H2B	Core chromatin component	Organismal	Decrease	Muscle stem cells	Transcriptional profiling	<i>M. musculus</i>
		Cellular (senescence)	Decrease	Fibroblasts	Western blot	<i>H. sapiens</i>
H3	Core chromatin component	Organismal	Decrease and changes in occupancy	Yeast cells	Western blot MNase-seq	<i>S. cerevisiae</i>
		Cellular (senescence)	Decrease	Epidermis and fibroblasts	Immunostaining Western blot SILAC-MS	<i>H. sapiens</i>

(continued)

**Table 2.1** (continued)

Epigenetic mark	Function	Aging paradigm	Change in epigenetic mark with aging	Change observed in tissue or cells	Method of detection	Species
		Organismal	Decrease	Soma, whole male flies and muscle stem cells	Western blot (normalization to tubulin) Transcriptional profiling	<i>C. elegans</i> <i>D. melanogaster</i> <i>M. musculus</i>
		Organismal	No change	Head	Western blot (normalization to total protein or DNA)	<i>D. melanogaster</i>
H4	Core chromatin component	Cellular (senescence)	Decrease	Fibroblasts	Western blot SILAC-MS	<i>H. sapiens</i>
macroH2A	Component of heterochromatin	Cellular (senescence) Organismal	Increase Increase	Fibroblasts Lung and liver	Immunostaining Western blot Immunostaining Western blot	<i>H. sapiens</i> <i>M. musculus</i>
HP1	Component of heterochromatin	Organismal	Remodeling	Head, gut, and fat cells	Immunostaining ChIP-chip	<i>D. melanogaster</i>
HP1 $\alpha$	Component of heterochromatin	Organismal (HGPS and Werner syndrome) Organismal	Decrease Decrease	Fibroblasts and MSCs MSCs	Immunostaining Western blot Western blot	<i>H. sapiens</i> <i>H. sapiens</i>
HP1 $\beta$	Component of heterochromatin	Cellular (senescence)	Increase	Fibroblasts	Immunostaining Western blot	<i>H. sapiens</i>

(continued)

Table 2.1 (continued)

Epigenetic mark	Function	Aging paradigm	Change in epigenetic mark with aging	Change observed in tissue or cells	Method of detection	Species
HP1 $\gamma$	Component of heterochromatin	Organismal (HGPS)	Decrease	Fibroblasts	Immunostaining	<i>H. sapiens</i>
		Organismal	Decrease	Fibroblasts	Immunostaining	<i>H. sapiens</i>
H3K9me1	Enriched in euchromatin along gene bodies, involved in transcriptional repression	Cellular (senescence)	Increase	Fibroblasts	Western blot SILAC-MS	<i>H. sapiens</i>
		Organismal	Decrease	Fibroblasts	Western blot SILAC-MS	<i>H. sapiens</i>
H3K9me2	Enriched along gene bodies, involved in transcriptional repression	Cellular (senescence)	Decrease	Fibroblasts	Western blot SILAC-MS	<i>H. sapiens</i>
		Organismal	Decrease	Whole male flies	Western blot	<i>D. melanogaster</i>
H3K9me3	Enriched in heterochromatin regions, involved in transcriptional silencing	Organismal (HGPS and Werner syndrome)	Decrease	Fibroblasts and MSCs	Immunostaining Western blot ChIP-seq	<i>H. sapiens</i>
		Cellular (senescence)	Decrease	Fibroblasts	SILAC-MS Western blot	<i>H. sapiens</i>
		Organismal	Increase and remodeling	Head	Western blot ChIP-chip	<i>D. melanogaster</i>
		Organismal	Decrease	Fibroblasts and soma	Immunostaining Western blot	<i>H. sapiens</i> <i>C. elegans</i>

(continued)

Table 2.1 (continued)

Epigenetic mark	Function	Aging paradigm	Change in epigenetic mark with aging	Change observed in tissue or cells	Method of detection	Species
H3K27me3	Enriched in euchromatin along gene bodies, involved in transcriptional repression	Organismal (HGPS and Werner syndrome)	Decrease and remodeling	Fibroblasts and MSCs	Immunostaining ChIP-seq	<i>H. sapiens</i>
		Cellular (senescence)	Remodeling	Fibroblasts	ChIP-seq	<i>H. sapiens</i>
		Organismal	Decrease	Soma	Western blot	<i>C. elegans</i>
H4K20me2	DNA repair and genomic stability	Organismal	Increase and remodeling	Brain, muscle stem cells and HSCs	Immunostaining ChIP-seq	<i>N. furzeri</i> <i>M. musculus</i>
		Cellular (senescence)	Increase	Fibroblasts	Western blot SILAC-MS	<i>H. sapiens</i>
		Organismal (HGPS)	Increase	Fibroblasts	Western blot Immunostaining	<i>H. sapiens</i>
H4K20me3	Enriched in pericentric heterochromatin	Cellular (senescence)	Decrease	Fibroblasts	Western blot SILAC-MS	<i>H. sapiens</i>
		Organismal	Increase	Liver and kidney	Liquid chromatography	<i>R. norvegicus</i>
		Organismal	Global increase and remodeling	Cortex	ChIP-seq	<i>M. mulatta</i>

(continued)

Table 2.1 (continued)

Epigenetic mark	Function	Aging paradigm	Change in epigenetic mark with aging	Change observed in tissue or cells	Method of detection	Species
H3K4me3	Enriched in euchromatin at promoters, involved in transcriptional activation	Organismal (Werner syndrome)	No global change	MSCs	ChIP-seq	<i>H. sapiens</i>
		Cellular (senescence)	Remodeling	Fibroblasts	ChIP-seq	<i>H. sapiens</i>
		Organismal	No change	Soma	Western blot	<i>C. elegans</i>
		Organismal	Decrease and remodeling	Head	ChIP-chip	<i>D. melanogaster</i>
H3K36me3	Enriched in euchromatin along gene body, involved in transcriptional elongation	Organismal	Minor global increase and remodeling	Neurons, HSCs and muscle stem cells	ChIP-seq	<i>H. sapiens</i> <i>M. musculus</i>
		Organismal (replicative lifespan)	No global change and remodeling	Yeast cells	ChIP-seq	<i>S. cerevisiae</i>
		Organismal	Minor global decrease, remodeling	Soma and head	Western blot ChIP-chip ChIP-seq	<i>C. elegans</i> <i>D. melanogaster</i>
H3K56ac	DNA replication, DNA damage response, and nucleosome assembly	Cellular (senescence)	Decrease	Fibroblasts	Immunostaining Western blot	<i>H. sapiens</i>
		Organismal (replicative lifespan)	Decrease	Yeast cells	Western blot	<i>S. cerevisiae</i>

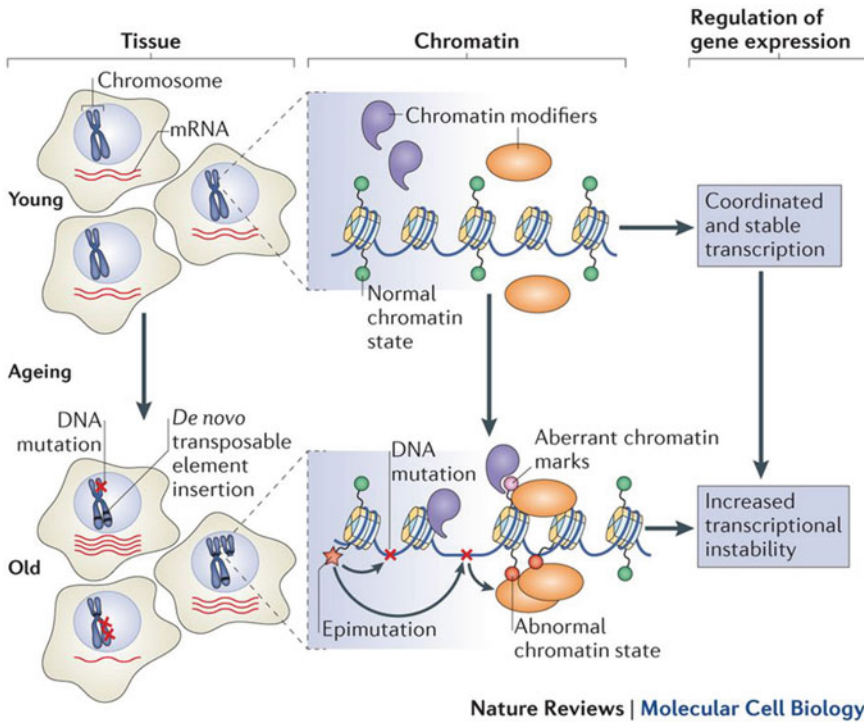
(continued)



Table 2.1 (continued)

Epigenetic mark	Function	Aging paradigm	Change in epigenetic mark with aging	Change observed in tissue or cells	Method of detection	Species
H4K16ac	Regulation of telomere silencing and regulation of chromatin compaction (?)	Organismal (HGPS)	Decrease	Fibroblasts and liver	Immunostaining Western blot	<i>M. musculus</i>
		Cellular (senescence)	Decrease	Fibroblasts	Western blot	<i>H. sapiens</i>
		Organismal (replicative lifespan)	Increase	Yeast cells	Western blot	<i>S. cerevisiae</i>
		Organismal	Decrease	Liver and kidney	Western blot	<i>H. sapiens</i> <i>M. musculus</i>
H4K12ac	Enriched in euchromatin along gene body, involved in transcriptional elongation	Organismal	Decrease upon contextual fear conditioning	Hippocampus	Western blot	<i>M. musculus</i>

RRBS reduced representation bisulfite sequencing, WGBS whole genome bisulfite sequencing, HPLC high-performance liquid chromatography, SILAC stable isotope labeling by amino acids, ChIP chromatin immunoprecipitation



**Fig. 2.2** A model for epigenetic changes observed in aging. Adopted from Benayoun et al. (2015)

the precision of transcription. The second major mechanism whereby epigenetically mediated mechanisms can affect chromatin stability is by making the epigenetic state of chromatin itself unstable, a process known as *epimutation* (Chong et al. 2007). Much like mutations in underlying genomic material, epimutations are passed on to daughter cells and increases in epimutation rate are linked to the number of cellular divisions; therefore, epimutations increase throughout life, as first described by Holliday (1987). These are driven by a number of factors, the first being errors in DNA and epigenome repair (Burgess et al. 2012). Furthermore, much like chronic inflammatory signaling seen in autoimmune diseases, chronic DNA damage signaling can cause pathogenic recruitment of a variety of histone- and DNA-methylation modifiers and induce inappropriate chromatin remodeling (Burgess et al. 2012), creating a feedback loop, resulting in ever-increasing local chromatin instability. This is modeled well in a genetic disease known as Werner syndrome, a deficiency in DNA helicase (the DNA “unwinding” enzyme) in mesenchymal stem cells, which leads to permanent aberrations in chromatin arrangements including global decreases in histone 3 lysine 9 trimethylation (H3K9me3), histone 3 lysine 27 trimethylation (H3K27me3), among others (Zhang et al. 2015). Age-associated epigenetic changes can also lead to increases in DNA transposition, which has potentially pathogenic effects. Up to 80% of the genome consists of transposable elements; themselves not encoding proteins,

but possessing the ability to regulate proximate genes. It has been previously demonstrated that transposable elements induce instability of the surrounding genome when actively transcribed; this transcription is normally reduced with suppressive epigenetic marks, especially histone 3 lysine 9 trimethylation; however, the expression of transposable elements increases with age and cellular senescence, and has been linked with human disease, particularly, diseases of the central nervous system (De Cecco et al. 2013; Reilly et al. 2013; Wood and Helfand 2013).

At a more granular level, global reductions in DNA methylation are seen in with aging, a feature noted decades ago by Wilson and Jones in 1983, who demonstrated that primary human, mouse, and hamster cells had substantially decreased global methylation levels when cultured in vitro over time compared to early passages and immortalized cell lines (Wilson and Jones 1983). Global hypomethylation induces genomic instability as well as leading to increases in gene expression of a variety of genes thought to promote tumorigenesis (Cruickshanks et al. 2013). Substantial changes in DNA methylation patterns associated with age are seen, however, when comparing gene islands, where aging tends to result in decreased methylation, whereas CpG sites within promoters tend to become hypermethylated with age, particularly in genes associated with development and differentiation (Maegawa et al. 2010; Day et al. 2013).

Importantly, the *rate* of age-associated epigenetic changes can vary among individuals. This concept was perhaps best illustrated by Fraga and colleagues (2005). They examined the epigenomes of lymphocytes (including DNA methylation and histone acetylation patterns) of a large cohort of monozygotic twins in Spain. They found that the epigenetic patterns of twins diverged according to their age; 50-year-old twins had substantial differences, whereas 3-year-old twins were nearly indistinguishable. These differences were, as one would expect, correlated strongly with lifestyle; twins who were younger, had similar lifestyles, and had spent most of their lifetimes together had the least DNA methylation changes throughout their genome, whereas twins who were older, had divergent lifestyles, and spent less of their lifetimes together had the most-divergent DNA methylation patterns. Importantly, these epigenetic differences were correlated with changes in gene expression patterns, and the differences persisted across a variety of tissue types. Although the full consequences of altered DNA methylation states associated with aging have not yet been worked out, there are suggestions that epigenetic alterations are associated to specific transcription factors' binding sites, and that this may contribute to gene misregulation during aging, as highlighted by Sun et al. in hematopoietic stem cells (Sun et al. 2014).

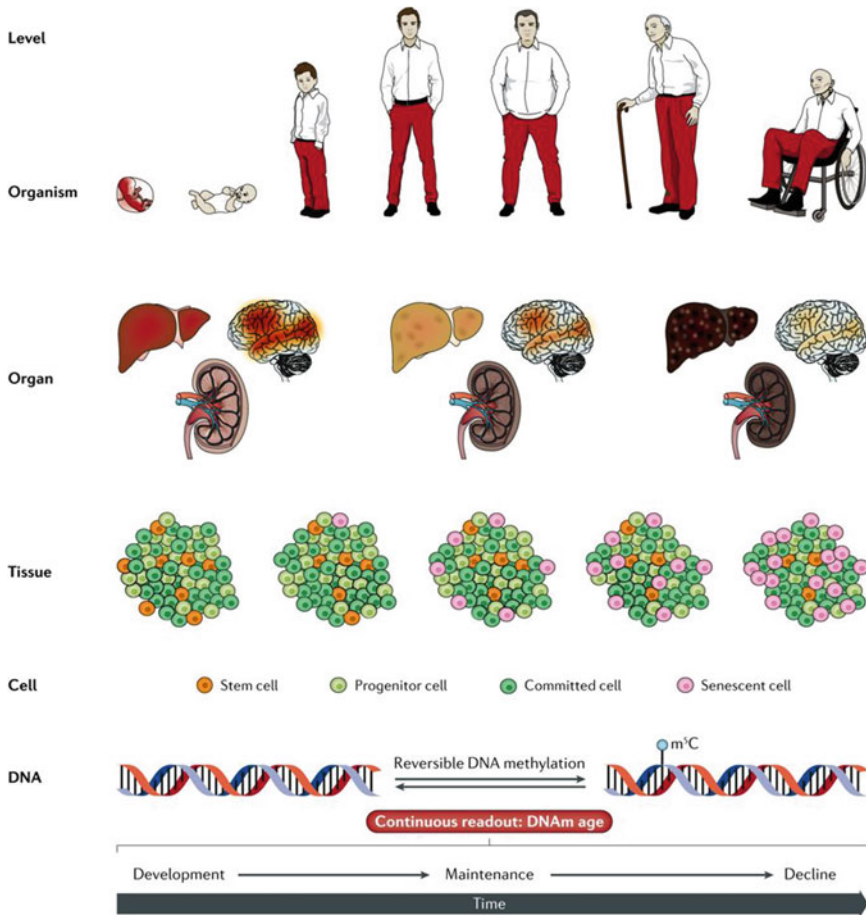
This can have direct implications in the development of age-related diseases; cancer, for instance. An article published in 2010 by Teschendorff et al. (2010) examined targets of polycomb group proteins (PCGTs), which are epigenetically repressed in multipotent cells including human embryonic and adult stem cells (Lee et al. 2006). DNA hypermethylation of PCGTs occurs at a higher rate (~12-fold) than at other genomic locations in the development of cancer. Teschendorff first identified significant linear correlations between age and DNA methylation of PCGTs (increases in methylation associated with age) irrespective of cancer status, in multiple tissues,

including whole blood, lung, cervix, and mesenchymal stem cells, among others. Interestingly though, the DNA methylation levels of these age-related PCGTs were even higher in cancer tissue, suggesting an exacerbation of the underlying DNA methylation changes caused by aging. This was bolstered by this age-related PCGT hypermethylation signature's ability to discriminate preneoplastic from normal cells.

### 2.2.2 *The Concept of an Epigenetic Clock*

In recent years, many additional investigators have sought to define the pattern of epigenetic (and particularly, DNA methylation) changes that occur during aging, and how these may be associated with or predictive of pathogenic disease states. Chief among them is Steve Horvath of the University of California—Los Angeles. His seminal article in 2013 defined a multi-tissue algorithmic predictor of biological age based on a calculated “methylation age”, defined from ~8000 healthy human tissues and cell types (Horvath 2013) (Fig. 2.3). This initial algorithm was an attempt by Horvath to produce an algorithm that could accurately predict a patient's biologic age based on DNA methylation information quantified using Illumina's genome-wide DNA methylation arrays (see subsequent chapters on epigenetic measurement methods), an important point to remember when considering later works seeking to define *differences* in this DNA methylation “clock” associated with chronic diseases. Using a machine learning approach, his final algorithm leveraged the DNA methylation values of 353 CpG sites to predict an epigenetic age, which was highly accurate with an average Pearson correlation coefficient 0.97 and median absolute error of 2.9 years; importantly, accurate correlations were made using this same algorithm across a variety of tissue types, including whole blood, a variety of brain samples, normal breast tissue, buccal cells, cartilage, dermal fibroblasts, epidermis, head and neck tissue, kidney, lung, mesenchymal stromal cells, stomach, thyroid, and others. There are a couple of notes of particular importance in his study. First and most interestingly, his algorithm was quite well validated in mixed peripheral blood samples. This is surprising given that this tissue is composed of a variety of cell types which have remarkably different lifespans; for example, monocytes live several weeks, whereas CD4<sup>+</sup> memory T cells can live decades. DNA methylation age values were concordant even when examining these cell lines individually. The second observation is that his algorithm is applicable to nonhuman primates; interestingly, it correlated best with chronological age in chimpanzees, but somewhat less well in gorillas, likely reflecting a larger evolutionary distance. Finally, as one would predict, the DNA methylation age of induced pluripotent- and embryonic stem cells are very close to 0, indicating that they do not display any substantial epigenetic aging.

This work has subsequently been extended by both Horvath and others to include the study of alterations of the epigenetic clock in a plethora of human diseases (Table 2.2). The aforementioned Werner syndrome, a disease whose character-



**Fig. 2.3** The DNA methylation-based epigenetic ‘clock’ is a continuous readout of cellular, tissue, and organismal aging. Adopted by Chen et al. (2016)

istics include many of the clinical signs of accelerated aging, is indeed associated with accelerated epigenetic aging of peripheral blood cells (Maierhofer et al. 2017). A number of central nervous system disorders exhibit accelerated epigenetic aging. Among these are Parkinson’s disease patients (Horvath and Ritz 2015) and Alzheimer’s disease (Levine et al. 2015b). In the latter case, accelerated epigenetic aging was found in the pathogenically affected organ itself (measured on dorsolateral prefrontal cortex samples), and was individually associated with a variety of neuropathological measurements, including the presence of diffuse plaques, neuritic plaques, amyloid load, cognitive functioning, and memory. In another fascinating study, Levine, Horvath, and colleagues examined methylation aging in blood samples drawn at baseline in a group of roughly 2000 patients as part of the Women’s

**Table 2.2** Conditions linked to alterations in the epigenetic clock. Adapted from Chen et al. (2016)

Condition	Source of DNA	Age estimator
Alzheimer disease	Prefrontal cortex	Horvath's clock
Amyloid load and neuropathology	Prefrontal cortex	Horvath's clock
Blood pressure (systolic)	Blood	Hannum's clock
Body mass index	Liver	Horvath's clock
Cancer	Blood	All clocks
Cardiovascular disease	Blood	DNAm PhenoAge
Coronary heart disease	Blood	DNAm PhenoAge
Cellular senescence (oncogene-induced)	Various	Horvath's clock
Centenarian (offspring status)	Blood	Horvath's clock
Cholesterol, HDL (not LDL)	Blood	Hannum's clock and DNAm PhenoAge
Cognitive performance	Blood and brain	Horvath's clock and DNAm PhenoAge
C-reactive protein	Blood	All
Diet (carotenoids)	Blood	Hannum's clock and DNAm PhenoAge
Dementia	Blood	DNAm PhenoAge
Down syndrome	Blood and brain	Horvath's clock
Education	Blood	Hannum's clock and DNAm PhenoAge
Exercise (recreational)	Blood	Hannum's clock and DNAm PhenoAge
Frailty	Blood	Horvath's clock and DNAm PhenoAge
Gender	Blood and brain	All
Gestational week	Blood and brain	Horvath's clock
Glucose	Blood	All
Huntington disease	Blood and brain	Horvath's clock
Income	Blood	Hannum's clock and DNAm PhenoAge
Insulin levels	Blood	All
Menopause	Blood and saliva	Horvath's clock
Mortality (all-cause)	Blood	All
Obesity	Liver and blood	All clocks
Osteoarthritis	Cartilage	Horvath's clock
Parkinson disease	Blood	All
Pubertal development	Blood	Horvath's clock

(continued)

**Table 2.2** (continued)

Condition	Source of DNA	Age estimator
Sleep	Blood	Hannum's clock
Smoking	Blood	DNAm PhenoAge
<i>TERT</i> expression	Blood and fibroblasts	Horvath's clock
Triglycerides	Blood	All
Walking speed	Blood	DNAm PhenoAge
Werner syndrome	Blood	Hannum's clock and Horvath's clock

HDL, high-density lipoprotein; LDL, low-density lipoprotein. 'Age estimator' indicates whether the effect was measured using the multi-tissue 353 CpG methylation-based age estimator, also known as Horvath's clock (Horvath 2013), the single-tissue (blood) 71 CpG methylation-based age estimator, also known as Hannum's clock (Hannum et al. 2013), or the DNA methylation-based phenotypic age (PhenoAge) measure based on 513 CpGs (Levine et al. 2018)

Health Initiative, 43 of which went on to develop lung cancer during an almost 20-year follow-up period, their hypothesis being that individuals with naturally occurring "slowing" of the epigenetic clock would be protected against the oncogenic effects of carcinogens like cigarette smoking. They found that, indeed, baseline acceleration of the epigenetic clock was associated with the development of cancer; surprisingly, in all groups, including patients who were former- or never-smokers (Levine et al. 2015a). This effect was even stronger in chronologically older individuals (above 70 years of age).

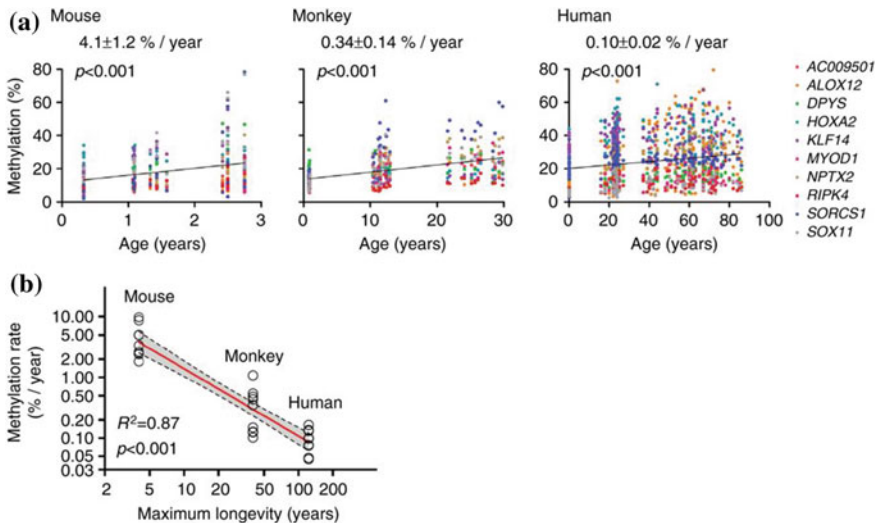
These studies offer the intriguing possibility of a *generally* increased susceptibility to a variety of chronic diseases based on epigenetic mechanisms, above and beyond those associated with specific diseases, genes, and pathways, which will be discussed in subsequent chapters. Horvath and colleagues have gone on to demonstrate premature epigenetic aging in a variety of chronic diseases and factors previously associated with chronic disease, including prenatally acquired HIV infection (Horvath et al. 2018), coronary heart disease (Horvath et al. 2016), menopause (Levine et al. 2016), cartilage of osteoarthritis patients (Vidal-Bralo et al. 2016), increases in body mass index, waist circumference, and fasting glucose (Grant et al. 2017). Furthermore, and of particular importance to the theme of this book, Irvin et al. have identified associations between accelerated epigenetic aging and a variety of inflammatory cytokines, including interleukin-6, C-reactive protein, and tumor necrosis factor-alpha (Irvin et al. 2018). In essentially all of the epigenetic aging literature to date, it appears that *accelerations* in preexisting aging-associated DNA methylation changes are associated with the onset of disease, indicating that potential future therapies directed at slowing or reducing epigenetic aging may be of benefit for prevention of chronic disease. Finally, this concept of an epigenetic clock (and of accelerations or decelerations of epigenetic aging) have been suggested as the most

promising molecular estimator of biologic age in a recent review, which pitted epigenetic clocks against transcriptomic-based, proteomic-based, metabolomic-based, and composite biomarkers, as well as telomere length (Jylhävä et al. 2017).

### 2.2.3 Epigenetic Effects of Antiaging Interventions

If aging, one of the strongest risk factors for most chronic diseases, is mediated to a large extent by widespread epigenetic changes, it stands to reason that interventions that slow or reverse the biological effects of aging might be acting through epigenetic mechanisms. A few studies have been published recently which do indeed indicate that this is the case, and serve to reinforce the notion that epigenetic states generally are plastic and potentially reversible. This concept that will no doubt be increasingly important in the near future as we continue developing the ability to modify epigenetic states in a site-specific manner.

As an example of antiaging epigenetic modification, we will highlight a 2017 article by Maegawa and colleagues (2017). In this study, they examined whole-blood DNA methylation patterns in a genome-wide fashion from young and old individuals from three species with widely disparate maximum lifespans: mice and rhesus monkeys fed either a “normal” diet or a calorie-restricted one, and human samples (Fig. 2.4). They describe several interesting findings. First, they quantify the



**Fig. 2.4** Cross-species comparison of DNA methylation drift and aging. **a** Correlations of methylation % and age across the lifespan of mouse, monkey, and human. **b** Association of normalized (% per year) DNA methylation drift and maximum species longevity. Adopted from Maegawa et al. (2017)



rate of epigenetic “drift”; that is, the average percent change in methylation rates per year of life. There was a very strong inverse correlation between the rate of epigenetic drift and the species maximum longevity, suggesting that age-related drift in DNA methylation may have a role to play in limiting the lifespan of a particular species. Next, they confirm that these changes in DNA methylation were indeed correlated with change in gene expression, suggesting that these changes are functional, rather than innocent bystanders. Finally, they examined in great detail what occurs to DNA methylation patterns when mice and monkeys have their food calories reduced by 40% and 30%, respectively. Caloric restriction has been previously demonstrated in many studies to have antiaging (or, at least, life-extending) properties (review, Fernández-Ruiz 2017). In the case of mice, caloric restriction (CR) was started in young adulthood (0.3 years of age) and continued throughout their 2–3-year lifespan, whereas in monkeys, CR began in middle age, and data analyzed at 22–30 years of age. DNA methylation analysis of whole-blood samples showed that old-CR animals were much more like young animals, and regular-diet animals were in a separate cluster. This effect was most pronounced at CpG sites which were unmethylated in young animals, and were more pronounced in mice, which had caloric restriction begun earlier on relative to their total lifespan than in monkeys. These epigenetic effects were best explained by a substantial reduction in epigenetic drift. Finally, they went on to examine the tissue specificity of these changes using samples from a variety of organs, including spleen, bone marrow, liver, kidney, small intestine, and large intestine, and found that this phenomenon of epigenetic drift reduction with calorie restriction was present in most tissues, exceptions being kidney and liver. This study has several direct implications on the study of epigenetics in chronic diseases and, perhaps most importantly, offers evidence that at least some of the disease-reducing benefits seen in interventions that increase lifespan may occur through epigenetic modulation.

## 2.3 Epigenetic Modifications as Biomarkers of Disease

So far in this chapter, we have focused on epigenetic modifications and their association with the disease from a pathogenesis standpoint. Another application of epigenetic study that has garnered more attention in the past few years is the potential of epigenetic modifications as biomarkers of the presence or progression of a particular disease. In this section, we will offer several examples of recently described epigenetic biomarkers of some of several important human diseases, and, when appropriate, discuss the advantages of using epigenetic biomarkers over other traditional methods of diagnosing and monitoring disease progression.

The field of oncology has seen the most interest in epigenetic biomarkers, where markers fall into one of three major categories: early or initial diagnosis, risk stratification, and prediction of treatment response. Take, for example, colorectal cancer, the second leading cause of cancer-related deaths in the US behind lung and bronchial cancer (Siegel et al. 2011). The gold standard method for colon cancer screening

at present is a colonoscopy procedure, whereby precancerous polyps can be identified and removed. This method has indeed reduced significantly colorectal cancer incidence and mortality (Schoen et al. 2012); however, an estimated 6% of colonoscopies detect advanced adenomas, and around 1% detect frank adenocarcinoma (Ferlitsch et al. 2011). Furthermore, the colonoscopy procedure itself is associated with nontrivial morbidity and is both expensive and burdensome for patients, who must undergo significant preparatory work the day before the procedure, receive conscious sedation during the procedure, etc. Clearly, a less invasive, more accurate, and earlier diagnostic biomarker for colon cancer would be desirable. Colon cancer occupies a unique position from an epigenetic standpoint in that much epigenetic analysis has already been done on cancerous and precancerous colon lesions (polyps, which are removed as part of screening colonoscopies) and the distinguishing epigenetic features have been well described (Lao and Grady 2011). Building on these data, several studies have suggested epigenetic biomarkers for both diagnosis and evaluation of the future risk of colon cancer. For example, hypermethylation of the *MGMT* gene has been strongly associated with the future likelihood of colon cancer development, even in grossly normal colon tissue, and has been suggested as a key early factor in carcinogenesis (Menigatti et al. 2009; Lang 2011). More recently, Luo and colleagues have reported that the DNA methylation pattern of a panel of six genes (*AOX-1*, *RARB2*, *RERG*, *ADAMTS9*, *IRF4*, and *FOXE1*) in a much more accessible biomarker fluid (peripheral blood cells) was highly associated with colorectal cancer (Luo et al. 2016). To date, at least two epigenetic biomarker panels have been FDA approved for colorectal cancer diagnosis including ColoVantage (a predictor using methylation of *SEPT9* in peripheral blood) and ColoSure (a predictor using methylation of the vimentin gene in fecal samples).

Another example of recent epigenetic biomarker development can be found in prostate cancer, a disease that has seen much controversy in recent years. This mostly stems from the prostate-specific antigen (or PSA) test, a previously widely used biomarker for the presence of prostate cancer, which led to what many experts feel is an overly aggressive biopsy- and treatment-paradigm in a disease that is frequently indolent in nature. The PSA test is a poor marker for a number of reasons, including a non-definitive cutoff for positivity, a nontrivial rate of PSA elevations without detectable prostate cancer, and a suboptimal rate of false-negative results (Castle 2015). Like colon cancer, much has already been described regarding the specific epigenetic changes that are associated with the transition from precancerous to malignant prostate tissue (Ruggero et al. 2018). A number of noninvasive prostate cancer epigenetic screening methods have been published, including *APC* gene methylation screening in urine (Jatkoe et al. 2015), *CHD13* in serum (Wang et al. 2014), and *ERBeta* in serum (Brait et al. 2017). Unlike colorectal cancer, however, there have to date been no FDA-approved epigenetic biomarkers for the diagnosis of prostate cancer.

Germane to this book on rheumatic disease are recent studies examining epigenetic biomarkers in systemic lupus erythematosus (SLE). One would expect, given that autoimmune diseases are driven to a substantial degree by alterations in circulating inflammatory cells, that peripheral blood-based epigenetic biomarkers might

be both related to the underlying biological pathogenesis and easily accessible for clinical application; indeed, the search for epigenetic biomarkers in peripheral blood cells has been quite fruitful in several autoimmune diseases. Lupus is an interesting and important target for novel biomarker development; it is a quite heterogeneous disease characterized by autoantibody production against a variety of nuclear targets, affecting almost every body system. The diagnosis of lupus is notoriously difficult, but very important, as delayed diagnosis can lead to substantial irreversible organ damage (Fortin et al. 1998). Currently available laboratory markers for lupus have substantial limitations, including mismatches in sensitivity and specificity.

An article by Zhao and colleagues in 2015 offers a good example of diagnostic biomarker development in lupus (Zhao et al. 2016). First, they screened peripheral blood mononuclear cell DNA methylation patterns from lupus patients, healthy controls, and non-lupus autoimmune rheumatoid arthritis and Sjogren's syndrome patients, and an independent validation cohort of lupus patients and healthy controls for potential disease-associated biomarkers using an Illumina genome-wide DNA methylation array (see subsequent chapter on methods for epigenetic quantitation). They identified differentially methylated sites within the *IFI44L* gene as highly associated with the presence of lupus; then, they examined two specific locations within this region in a much larger group of patients as part of a discovery cohort, using a different method (bisulfite pyrosequencing). They went on to confirm the sensitivity and specificity of the DNA methylation values of these two CpG sites in multiple validation cohorts, including among the same ethnic group (Chinese) and among a different ethnic group (Europeans). Remarkably, these methylation sites had quite high sensitivity and specificity for the diagnosis of lupus, in the 90% range for both among Chinese, and in the 70–80% range for both in Europeans. Furthermore, they were able to accurately differentiate lupus from both rheumatoid arthritis and Sjogren's syndrome patients. A subsequent study by Coit et al. (2015) identified a single CpG within the *CHST12* gene of naive T cells from lupus patients as highly associated with the presence of lupus nephritis, a manifestation of SLE which is difficult to detect without an invasive biopsy, with a sensitivity of 86% and specificity of 71%.

Other studies have similarly noted strong associations with certain clinical manifestations and disease indices with alterations in DNA methylation of easily accessible tissues. These include *IL10* and *IL12* hypomethylation correlation with lupus disease activity scores (Lin et al. 2012), *IL6* methylation correlation with lupus disease activity, prediction of flare, and serum complement levels (Mi and Zeng 2008; Tang et al. 2014), *FOXP3* methylation association with disease activity (Horwitz 2008), and retroviral element *HERV-E* and *HERV-K* methylation association with both disease activity and the presence of a variety of autoantibody specificities (Okada et al. 2002; Piotrowski et al. 2005).

Another area of biomarker research where epigenetics can play a strong role is in predicting the response to particular therapies. In rheumatoid arthritis (RA), for example, we are fortunate to have dozens of effective “traditional” and biologic medications with more being approved on an almost yearly basis; unfortunately, however, experience has shown that these are not universally effective. Additionally, most of these new drugs take some time to reach peak effectiveness; consequently,

when a patient has a suboptimal response within the first weeks of treatment, it is nearly impossible to determine whether this represents a failure of the therapy or simply a delayed response. Given the substantial cost of these treatments, a need clearly exists for biomarkers that can help stratify RA patients and predict whether or not they will respond to a particular drug or class of drugs.

Traditional non-biologic synthetic disease-modifying drugs (DMARDs) may actually partially work by inducing epigenetic changes, as highlighted by de Andres and colleagues in 2015, who outlined substantial DNA methylation changes induced by treatment with one of these drugs, methotrexate, in RA patients (de Andres et al. 2015). Although this is a relatively new field, an article by Plant and colleagues in 2016 demonstrated the usefulness of peripheral blood epigenetic patterns to predict clinical responses in RA patients subsequently treated with the antitumor necrosis factor- $\alpha$  drug etanercept (Plant et al. 2016). In this study, they performed an epigenome-wide association analysis (EWAS). The discovery cohort included 36 pretreatment whole-blood samples from RA patients who later had a poor response to etanercept by a well-validated clinical disease activity index after three months of treatment and compared this with 36 pretreatment whole blood samples from patients who later had a very good response to etanercept (clinical disease activity index in remission after three months of therapy). They identified five CpG sites with significant differential methylation between responders and nonresponders; intriguingly, all five demonstrated reduced methylation among good responders compared to poor responders. Interestingly, two of the top five, and four of the top 15 most differentially methylated CpG sites between groups were located within exon 7 of the *LRAP1* gene; they went on to demonstrate that three *genetic* single nucleotide polymorphism mutations were highly correlated with the DNA methylation levels of two of these CpGs, indicating that epigenetic and genetic mechanisms in *LRAP1* interact to produce a good or bad TNF inhibitor response in RA patients. This sort of interaction, known as methylation quantitative trait loci or meth-QTL, has in recent years been demonstrated as quite important in a variety of clinical phenotypes and disease states, and will likely be even more intensely studied as the field moves toward whole-genome and—epigenome studies, collectively known as “big data”, which will be discussed in the following final section of this chapter.

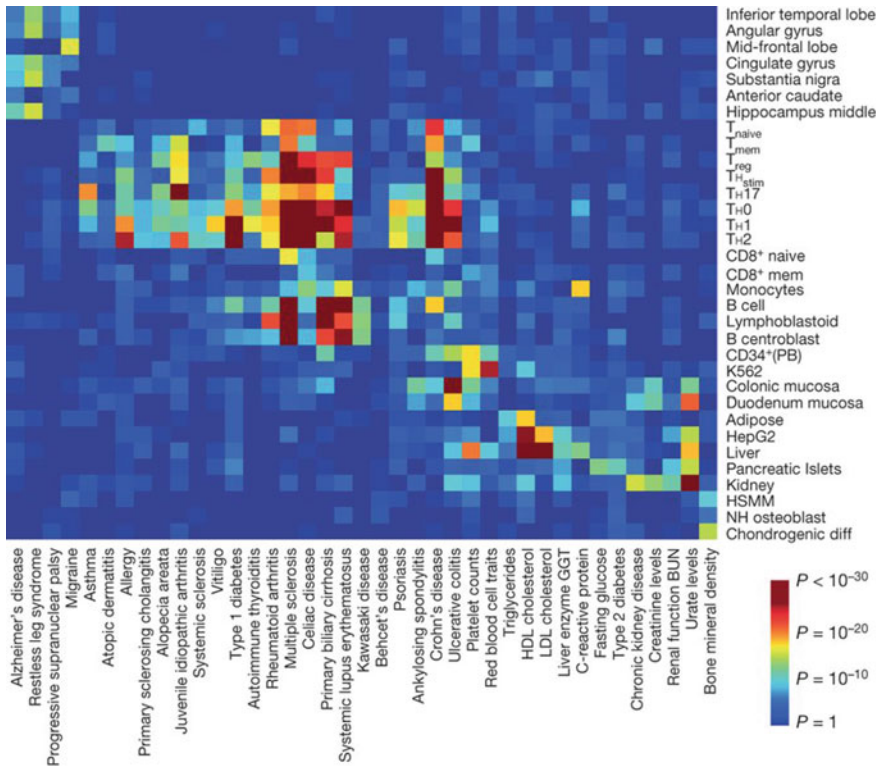
## **2.4 Epigenetics in the Modern World: Using Big Data to Understand the Pathogenesis of Complex Diseases in the Present Day**

In the final section of this chapter, we will move from a historical perspective of epigenetics to discuss the state-of-the-art in epigenetics research, as well as offer some glimpses into the future of epigenetics research as it relates to chronic disease research generally, and disorders of the immune system specifically. As technology has advanced and allowed researchers to generate increasingly intricate maps of both

genetic risk and epigenetic associations in chronic disease states, new techniques have been developed to make sense of this massive amount of data. Several studies have been recently published, including in the field of rheumatology, which seeks to computationally synthesize large genetic, epigenetic, proteomic, and transcriptomic datasets in key cell types and draw conclusions regarding the ways in which these—omics domains interact regionally to contribute to the development of the disease. In a way counterintuitively, large, detailed data sets and complex computational analysis have also allowed researchers to take a step back and draw more general conclusions about the genes, organ systems, and pathways that are altered in chronic diseases generally. As the cost of complete genomic sequencing and base-specific whole-genome epigenetic analyses continue to fall, we will no doubt see more of these integrated analyses in the future; hopefully offering a more complete understanding of the fundamental processes underlying disease pathogenesis driven by machine-learning-driven insights, and perhaps offer novel treatment strategies.

A nice example of this computationally driven view of the integration of underlying genetic sequence and epigenetic modification in the development of autoimmune diseases generally can be found in an article by Farh et al. (2015). In this study, they collected genome-wide association study (GWAS) data from 39 independent, well-powered GWAS studies representing a variety of complex systemic diseases from published datasets. They then clustered diseases based on shared genetic susceptibility loci in order to produce a map relating the diseases to each other on a genetic level; remarkably, many disparate disorders shared a number of similar features. For example, 69% of single nucleotide polymorphisms (SNPs) collected in their meta-analysis were shared among at least two autoimmune diseases. Next, they developed a novel computational approach to estimate the probability that SNPs associated with multiple autoimmune diseases represented a causal SNP as opposed to an “innocent bystander”, a method they termed Probabilistic Identification of Causal SNPs (PICS). Next, they generated immune cell subtype-specific “maps” (including, for example, CD4<sup>+</sup> and CD8<sup>+</sup> T cell subsets, B cells, and monocytes) based on both data generated within their laboratory and data previously published through the NIH Epigenomics Project and the Encyclopedia of DNA Elements (ENCODE, a publicly accessible research project aiming to identify and catalogue all functional elements within the human genome), including a 56 cell types. They computed a genome-wide map of histone posttranslational modification regulatory elements and then clustered individual cell types based on these patterns.

Finally, they inferred the cell types most likely driving particular autoimmune diseases by overlaying these two datasets; that is, they looked for cell types in which disease-specific genetic mutations were located in regions known to be epigenetically “active” (Fig. 2.5). This allowed them to guess which cell types are most likely involved in specific autoimmune disease. Several of the associations were predictable; for example, both T-cell stimulation and B cells shared enrichment of epigenetically active enhancers. Interestingly, these SNPs were also common within generalized *stimulus-dependent* enhancers; epigenetically active enhancers from *unstimulated* T cells did not overlap with these SNPs. Further, and a good example of the complexity of epigenetic regulation, they found that many disease-



**Fig. 2.5** Cell-type specificity of human diseases, inferred from epigenetic mapping. Adopted from Farh et al. (2015)

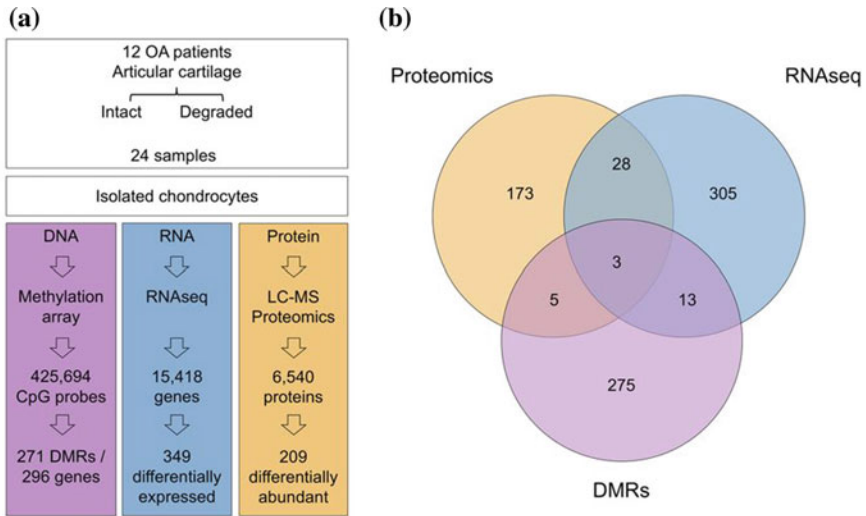
and cell-type-associated SNPs were located within regions which regulated the transcription of noncoding RNAs, another epigenetic control mechanism. In the end, they estimate that genetic changes in enhancer regions producing mixed genetic-epigenetic effects overall account for around ~60% of all disease-associated genetic variants (Houseman et al. 2012; Martino et al. 2012).

The next problem, of course, is identifying how these genetic/epigenetic interactions within enhancer regions in specific immune cell subsets are actually contributing to disease risk or pathogenesis. For example, the most obvious potential explanation is that these changes conspire to alter gene regulation, inducing overexpression or underexpression or, intriguingly, interrupting the appropriate *regulation* of the gene in question, allowing it to be turned on or off at inappropriate times. They investigated this possibility by modeling how these variations (be they genetic or epigenetic) alter the binding of transcription factors. Although they did find enrichment of PICS SNPs within or proximate to many transcription factor binding sites, these represented only a minority (around 7%), and was more or less the percentage one would expect to find at random from non-disease-related SNPs. Unfortunately, the authors were unable to give a definitive pathophysiologic mechanism, although they speculate that future

research to better define the structure and function of gene enhancers, which the majority of these sites fell within, may allow a better understanding.

As outlined above, one would expect that this sort of analysis would also give us another important bit of information: the cell types which are most likely to be involved in a particular disease based on this genetic/epigenetic inference. The authors compared their list of likely “causal” SNPs for a given disease with the histone “fingerprint” of a wide variety of cell types included in both the NIH Epigenomics Project and ENCODE datasets mentioned above. Several of these associations were expected; for example, migraine disease and Alzheimer’s were both predicted to involve brain tissue, just as inflammatory bowel diseases like ulcerative colitis mapped to gastrointestinal tissue. Some were more surprising, however. For example, systemic lupus erythematosus, Kawasaki disease (a form of large vessel autoimmune vasculitis), and primary biliary cirrhosis all mapped mostly to B cells. One can imagine large genetic/epigenetic screens such as this one being used in the future to help researchers narrow down the increasingly large field of potential players in a particular condition for more in-depth large-scale-omics analyses; future similar studies may also allow researchers to predict the most likely involved cell types to study for potential therapeutics.

Multi-omics, integrated computational analyses can also be quite informative when tasked with closely examining a single pathogenic tissue. As an example of this, in 2018 Steinberg and colleagues published an integrative and multi-omics analysis of cartilage from osteoarthritis (OA) patients, and offers a good example of how such state-of-the-art studies can be used to distill enormous amounts of data to actionable targets. In their study, they collected osteochondral samples from patients undergoing total knee joint replacements, both eroded and intact sections from within the same joint, as well as control cartilage samples from non-OA individuals, in both a discovery and two replication cohorts. They then set out to examine proteins that were differentially present in the two disease states using liquid chromatography-mass spectrophotometry (LC-MS) technique, mRNA gene expression using RNA-Seq, and epigenetic patterns (specifically, DNA methylation) using Illumina’s genome-wide DNA methylation microarray system. As expected, they found a large number of differences in each domain; 209 proteins were differentially abundant, 349 genes were differentially expressed from an mRNA perspective, and 271 differentially methylated regions were identified from an epigenetic perspective (Fig. 2.6). The important next step, which has only been possible with technological and computational advances over the past few years, was the integration of these data across multiple-omics domains. They found 49 genes which differed in at least 2 domains, and three genes that exhibited significant evidence for OA involvement across all three domains: aquaporin 1 (*AQP1*), the collagen 1 gene (*COL1A1*), and *CLEC3B* gene, which encodes tetranectin, a regulator of fibrinolysis. All three of these genes were upregulated at both a protein and mRNA level, and also exhibited reduced DNA methylation at all CpG probes associated with them included in the Illumina arrays. Interestingly, of the 49 genes with differential regulation across at least 2 domains, fully one third had not been previously implicated in OA pathogenesis.



**Fig. 2.6** Overview of genes identified as associated with osteoarthritis in human cartilage from a multi-omics perspective. Adopted from Steinberg et al. (2017)

An important consequence of this sort of multi-omics domain is the ability of researchers to hone in on potentially important, but previously unrecognized, avenues for treatment. In this study, they applied their list of 49 genes with cross-domain association with OA to Drugbank (Law et al. 2014), a comprehensive databank which lists information on drug targets, and identified ten drugs which had actions on nine of the 49 dysregulated proteins, all of which already had Food and Drug Administration marketing authorization for use in human patients. Some were expected, including non-steroidal anti-inflammatory drugs, but a few were novel, including vitamin K1 (phyloquinone), an antipsychotic/antiemetic (trifluoperazine), and a drug used to treat elevated cholesterol (ezetimibe), among others. As technology continues to develop, and the costs of performing this sort of global analysis drops, future epigenetic studies will no doubt more and more frequently include just this sort of large, multi-omics approach to data analysis, and will likely substantially benefit patients by identifying previously unrecognized druggable targets in various epigenetically driven diseases.

## 2.5 Summary

In this chapter, we have highlighted the history of the discovery of epigenetic control mechanisms, which in many ways paralleled the discovery of the DNA code itself. We then highlighted a few ways in which epigenetics has changed the way we think about both basic biological processes and the pathogenesis of the complex human



disease. We began with a discussion of the epigenetics of “normal” aging, and went on to discuss the recent development of an epigenetic “clock”. Alterations in the rate of epigenetic aging have been demonstrated to be associated with a variety of chronic human diseases; furthermore, interventions that slow aging also appear to slow the rate of *epigenetic* aging, suggesting that antiaging interventions may have widespread epigenetic effects. We then discussed a few examples of a nonpathogenic role for epigenetics in clinical medicine: the development of epigenetic biomarkers both for the presence of disease, but also as predictors of response to particular therapies. Finally, we exemplified two examples of complex, modern epigenetic studies, which leverage massive amounts of data from multiple biological domains to infer the ways in which epigenetic modifications affect gene expression at a transcriptomic, proteomic, and whole-organism level. The ways in which epigenetic mechanisms have contributed to our understanding of disease pathogenesis is nothing short of remarkable; given its meteoric rise in importance over the past 100+ years, there is no doubt that the study of epigenetics in human disease has a bright future indeed.

## References

- Ahlenstiel CL, Lim HGW, Cooper DA et al (2012) Direct evidence of nuclear Argonaute distribution during transcriptional silencing links the actin cytoskeleton to nuclear RNAi machinery in human cells. *Nucleic Acids Res* 40:1579–1595
- Allfrey VG, Faulkner R, Mirsky AE (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci USA* 51:786–794
- Ameyar-Zazoua M, Rachez C, Souidi M et al (2012) Argonaute proteins couple chromatin silencing to alternative splicing. *Nat Struct Mol Biol* 19:998–1004
- Avery OT, Macleod CM, McCarty M (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J Exp Med* 79:137–158
- Barr ML, Bertram EG (1949) A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature* 163:676
- Basyuk E, Suavet F, Doglio A et al (2003) Human let-7 stem-loop precursors harbor features of RNase III cleavage products. *Nucleic Acids Res* 31:6593–6597
- Benayoun BA, Pollina EA, Brunet A (2015) Epigenetic regulation of ageing: linking environmental inputs to genomic stability. *Nat Rev Mol Cell Biol* 16:593–610
- Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A (2009) An operational definition of epigenetics. *Genes Dev* 23:781–783
- Berget SM, Moore C, Sharp PA (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci USA* 74:3171–3175
- Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409:363–366
- Bestor TH, Ingram VM (1983) Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc Natl Acad Sci USA* 80:5559–5563
- Brait M, Banerjee M, Maldonado L et al (2017) Promoter methylation of MCAM, ER $\alpha$  and ER $\beta$  in serum of early stage prostate cancer patients. *Oncotarget* 8:15431–15440
- Briggs R, King TJ (1952) Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. *Proc Natl Acad Sci USA* 38:455–463

- Britten RJ, Davidson EH (1969) Gene regulation for higher cells: a theory. *Science* 165:349–357
- Brownell JE, Zhou J, Ranalli T et al (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84:843–851
- Burgess RC, Misteli T, Oberdoerffer P (2012) DNA damage, chromatin, and transcription: the trinity of aging. *Curr Opin Cell Biol* 24:724–730
- Castle PE (2015) PSA testing for prostate cancer screening. *Lancet Oncol* 16:e2–e3
- Chen BH, Marioni RE, Colicino E et al (2016) DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging* 8:1844–1865
- Chong S, Youngson NA, Whitelaw E (2007) Heritable germline epimutation is not the same as transgenerational epigenetic inheritance. *Nat Genet* 39:574–575; author reply 575–576
- Chow LT, Gelinis RE, Broker TR, Roberts RJ (1977) An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 12:1–8
- Clark RJ, Felsenfeld G (1971) Structure of chromatin. *Nat. New Biol* 229:101–106
- Coit P, Renauer P, Jeffries MA et al (2015) Renal involvement in lupus is characterized by unique DNA methylation changes in naïve CD4<sup>+</sup> T cells. *J Autoimmun* 61:29–35
- Cruikshanks HA, McBryan T, Nelson DM et al (2013) Senescent cells harbour features of the cancer epigenome. *Nat Cell Biol* 15:1495–1506
- Dang W, Steffen KK, Perry R et al (2009) Histone H4 lysine 16 acetylation regulates cellular lifespan. *Nature* 459:802–807
- Davis RL, Weintraub H, Lassar AB (1987) Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51:987–1000
- Day K, Waite LL, Thalacker-Mercer A et al (2013) Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. *Genome Biol* 14:R102
- de Andres MC, Perez-Pampin E, Calaza M et al (2015) Assessment of global DNA methylation in peripheral blood cell subpopulations of early rheumatoid arthritis before and after methotrexate. *Arthritis Res Ther* 17:233
- De Cecco M, Criscione SW, Peckham EJ et al (2013) Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. *Aging Cell* 12:247–256
- Doerfler W (1981) DNA methylation—a regulatory signal in eukaryotic gene expression. *J Gen Virol* 57:1–20
- Doi N, Zenno S, Ueda R et al (2003) Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. *Curr Biol* 13:41–46
- Eb P (1895) From the greeks to darwin: an outline of the development of the evolution idea. *Nature* 52:361
- Farh KK-H, Marson A, Zhu J et al (2015) Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 518:337–343
- Ferlitsch M, Reinhart K, Pramhas S et al (2011) Sex-specific prevalence of adenomas, advanced adenomas, and colorectal cancer in individuals undergoing screening colonoscopy. *JAMA* 306:1352–1358
- Fernández-Ruiz I (2017) Metabolism: Calorie restriction for healthy ageing. *Nat Rev Cardiol* 14:190
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391:806–811
- Fortin PR, Abrahamowicz M, Neville C et al (1998) Impact of disease activity and cumulative damage on the health of lupus patients. *Lupus* 7:101–107
- Fraga MF, Ballestar E, Paz MF et al (2005) Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci USA* 102:10604–10609
- Freser I (2010) Przewodnik dotyczący stosowania dyrektywy 2006/42/WE w sprawie maszyn. Komisja Europejska, Przedsiębiorstwa i Przemysł
- Grant CD, Jafari N, Hou L et al (2017) A longitudinal study of DNA methylation as a potential mediator of age-related diabetes risk. *Geroscience* 39:475–489

- Hadorn E (1965) Problems of determination and transdetermination. *Brookhaven Symp Biol* 148:161
- Hannah A (1951) Localization and function of heterochromatin in *Drosophila melanogaster*. *Adv Genet* 4:87–125
- Hannum G, Guinney J, Zhao L et al (2013) Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell* 49:359–367
- Harvey W (1651) *Exercitationes de generatione animalium. Quibus accedunt quaedam de partu; de membranis ac humoribus uteri; & de conceptione.* <https://doi.org/10.5962/bhl.title.44364>
- Harvey W (1653) Anatomical exertitions, concerning the generation of living creatures: to which are added particular discourses, of births, and of conceptions, &c. By William Harvey, Doctor of Physick, and Professor of Anatomy, and Chirurgery, in the Colledge of Physitians of London. James Young
- Hecht A, Laroche T, Strahl-Bolsinger S et al (1995) Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell* 80:583–592
- Hershey AD, Chase M (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* 36:39–56
- Holliday R (1979) A new theory of carcinogenesis. *Br J Cancer* 40:513–522
- Holliday R (1987) The inheritance of epigenetic defects. *Science* 238:163–170
- Holliday R, Pugh JE (1975) DNA modification mechanisms and gene activity during development. *Science* 187:226–232
- Horvath S (2013) DNA methylation age of human tissues and cell types. *Genome Biol* 14:R115
- Horvath S, Ritz BR (2015) Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. *Aging* 7:1130–1142
- Horvath S, Gurven M, Levine ME et al (2016) An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol* 17:171
- Horvath S, Phillips N, Heany SJ et al (2018) Perinately acquired HIV infection accelerates epigenetic aging in South African adolescents. *AIDS*. <https://doi.org/10.1097/QAD.0000000000001854>
- Horwitz DA (2008) Regulatory T cells in systemic lupus erythematosus: past, present and future. *Arthritis Res Ther* 10:227
- Houseman EA, Accomando WP, Koestler DC et al (2012) DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 13:86
- Hu Z, Chen K, Xia Z et al (2014) Nucleosome loss leads to global transcriptional up-regulation and genomic instability during yeast aging. *Genes Dev* 28:396–408
- Irvin MR, Aslibekyan S, Do A et al (2018) Metabolic and inflammatory biomarkers are associated with epigenetic aging acceleration estimates in the GOLDN study. *Clin Epigenetics* 10:56
- Ito S, D'Alessio AC, Taranova OV et al (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466:1129–1133
- Jähner D, Stuhlmann H, Stewart CL et al (1982) De novo methylation and expression of retroviral genomes during mouse embryogenesis. *Nature* 298:623–628
- Jatkoe TA, Karnes RJ, Freedland SJ et al (2015) A urine-based methylation signature for risk stratification within low-risk prostate cancer. *Br J Cancer* 112:802–808
- Jones PA (1985) Altering DNA methylation with 5-azacytidine. *Cell* 40:485–486
- Jylhävä J, Pedersen NL, Hägg S (2017) Biological age predictors. *EBioMedicine* 21:29–36
- Kim DH, Villeneuve LM, Morris KV, Rossi JJ (2006) Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol* 13:793–797
- Ko M, Huang Y, Jankowska AM et al (2010) Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* 468:839–843
- Koelsch KA, Webb R, Jeffries M et al (2013) Functional characterization of the MECP2/IRAK1 lupus risk haplotype in human T cells and a human MECP2 transgenic mouse. *J Autoimmun* 41:168–174
- Kossel A (1884) Über einen peptonartigen bestandteil des zellkerns. *Z Physiol Chem* 511–515

- Lang L (2011) New test measures DNA methylation levels to predict colon cancer. *Gastroenterology* 140:5
- Lao VV, Grady WM (2011) Epigenetics and colorectal cancer. *Nat Rev Gastroenterol Hepatol* 8:686–700
- Laskey RA, Gurdon JB (1970) Genetic content of adult somatic cells tested by nuclear transplantation from cultured cells. *Nature* 228:1332–1334
- Law V, Knox C, Djombou Y et al (2014) DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res* 42:D1091–D1097
- Lee JT, Bartolomei MS (2013) X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell* 152:1308–1323
- Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75:843–854
- Lee Y, Ahn C, Han J et al (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425:415–419
- Lee TI, Jenner RG, Boyer LA et al (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125:301–313
- Levine ME, Hosgood HD, Chen B et al (2015a) DNA methylation age of blood predicts future onset of lung cancer in the women’s health initiative. *Aging* 7:690–700
- Levine ME, Lu AT, Bennett DA, Horvath S (2015b) Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer’s disease related cognitive functioning. *Aging* 7:1198–1211
- Levine ME, Lu AT, Chen BH et al (2016) Menopause accelerates biological aging. *Proc Natl Acad Sci USA* 113:9327–9332
- Levine ME, Lu AT, Quach A et al (2018) An epigenetic biomarker of aging for lifespan and healthspan. *BioRxiv* 276162. <https://www.biorxiv.org/search/276162>
- Lin S-Y, Hsieh S-C, Lin Y-C et al (2012) A whole genome methylation analysis of systemic lupus erythematosus: hypomethylation of the IL10 and IL1R2 promoters is associated with disease activity. *Genes Immun* 13:214–220
- Luo X, Huang R, Sun H et al (2016) Methylation of a panel of genes in peripheral blood leukocytes is associated with colorectal cancer. *Sci Rep* 6:29922
- Lyon MF (1961) Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* 190:372
- Maegawa S, Hinkal G, Kim HS et al (2010) Widespread and tissue specific age-related DNA methylation changes in mice. *Genome Res* 20:332–340
- Maegawa S, Lu Y, Tahara T et al (2017) Caloric restriction delays age-related methylation drift. *Nat Commun* 8:539
- Maierhofer A, Flunkert J, Oshima J et al (2017) Accelerated epigenetic aging in Werner syndrome. *Aging* 9:1143–1152
- Martino D, Maksimovic J, Joo J-H et al (2012) Genome-scale profiling reveals a subset of genes regulated by DNA methylation that program somatic T-cell phenotypes in humans. *Genes Immun* 13:388–398
- Menigatti M, Truninger K, Gebbers J-O et al (2009) Normal colorectal mucosa exhibits sex- and segment-specific susceptibility to DNA methylation at the hMLH1 and MGMT promoters. *Oncogene* 28:899–909
- Mi X-B, Zeng F-Q (2008) Hypomethylation of interleukin-4 and -6 promoters in T cells from systemic lupus erythematosus patients. *Acta Pharmacol Sin* 29:105–112
- Mizuguchi G, Tsukiyama T, Wisniewski J, Wu C (1997) Role of nucleosome remodeling factor NURF in transcriptional activation of chromatin. *Mol Cell* 1:141–150
- Morgan TH (1911) An attempt to analyze the constitution of the chromosomes on the basis of sex-limited inheritance in *Drosophila*. *J Exp Zool A Ecol Genet Physiol* 11:365–413
- Muller HJ (1930) Types of visible variations induced by X-rays in *Drosophila*. *J Genet* 22:299–334
- O’Sullivan RJ, Kubicek S, Schreiber SL, Karlseder J (2010) Reduced histone biosynthesis and chromatin changes arising from a damage signal at telomeres. *Nat Struct Mol Biol* 17:1218–1225

- Ohno S, Kaplan WD, Kinoshita R (1959) Formation of the sex chromatin by a single X-chromosome in liver cells of *Rattus norvegicus*. *Exp Cell Res* 18:415–418
- Okada M, Ogasawara H, Kaneko H et al (2002) Role of DNA methylation in transcription of human endogenous retrovirus in the pathogenesis of systemic lupus erythematosus. *J Rheumatol* 29:1678–1682
- Pazin MJ, Kamakaka RT, Kadonaga JT (1994) ATP-dependent nucleosome reconfiguration and transcriptional activation from preassembled chromatin templates. *Science* 266:2007–2011
- Peterson CL, Herskowitz I (1992) Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. *Cell* 68:573–583
- Piotrowski PC, Duriagin S, Jagodzinski PP (2005) Expression of human endogenous retrovirus clone 4-1 may correlate with blood plasma concentration of anti-U1 RNP and anti-Sm nuclear antibodies. *Clin Rheumatol* 24:620–624
- Plant D, Webster A, Nair N et al (2016) Differential methylation as a biomarker of response to etanercept in patients with rheumatoid arthritis. *Arthritis Rheumatol* 68:1353–1360
- Rai TS, Cole JJ, Nelson DM et al (2014) HIRA orchestrates a dynamic chromatin landscape in senescence and is required for suppression of neoplasia. *Genes Dev* 28:2712–2725
- Reilly MT, Faulkner GJ, Dubnau J et al (2013) The role of transposable elements in health and diseases of the central nervous system. *J Neurosci* 33:17577–17586
- Riggs AD (1975) X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet* 14:9–25
- Ruggiero K, Farran-Matas S, Martinez-Tebar A, Aytes A (2018) Epigenetic regulation in prostate cancer progression. *Curr Mol Biol Rep* 4:101–115
- Sawalha AH, Webb R, Han S et al (2008) Common variants within MECP2 confer risk of systemic lupus erythematosus. *PLoS One* 3:e1727
- Schoen RE, Pinsky PF, Weissfeld JL et al (2012) Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. *N Engl J Med* 366:2345–2357
- Siegel R, Ward E, Brawley O, Jemal A (2011) Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin* 61:212–236
- Slutels F, Zwart R, Barlow DP (2002) The non-coding air RNA is required for silencing autosomal imprinted genes. *Nature* 415:810–813
- Stedman E, Stedman E (1950) Cell specificity of histones. *Nature* 166:780–781
- Steinberg J, Ritchie GRS, Roumeliotis TI et al (2017) Integrative epigenomics, transcriptomics and proteomics of patient chondrocytes reveal genes and pathways involved in osteoarthritis. *Sci Rep* 7:8935
- Sturtevant AH (1913) The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *J Exp Zool A Ecol Genet Physiol* 14:43–59
- Sun D, Luo M, Jeong M et al (2014) Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell* 14:673–688
- Sun Q, Zhu T, Wang C-Y, Ma D (2015) Binding of human SWI1 ARID domain to DNA without sequence specificity: a molecular dynamics study. *J Huazhong Univ Sci Technol Med Sci* 35:469–476
- Tahiliani M, Koh KP, Shen Y et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324:930–935
- Tang C, Li Y, Lin X et al (2014) Hypomethylation of interleukin 6 correlates with renal involvement in systemic lupus erythematosus. *Cent Eur J Immunol* 39:203–208
- Taunton J, Hassig CA, Schreiber SL (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* 272:408–411
- Teschendorff AE, Menon U, Gentry-Maharaj A et al (2010) Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res* 20:440–446
- Thakur N, Tiwari VK, Thomassin H et al (2004) An antisense RNA regulates the bidirectional silencing property of the *Kcnq1* imprinting control region. *Mol Cell Biol* 24:7855–7862

- van der Krol AR, Mur LA, de Lange P et al (1990) Inhibition of flower pigmentation by antisense CHS genes: promoter and minimal sequence requirements for the antisense effect. *Plant Mol Biol* 14:457–466
- Vidal-Bralo L, Lopez-Golan Y, Mera-Varela A et al (2016) Specific premature epigenetic aging of cartilage in osteoarthritis. *Aging*. <https://doi.org/10.18632/aging.101053>
- Waddington CH (1939) *An introduction to modern genetics*, by C.H. Waddington
- Waddington CH (1952) *The epigenetics of birds*. University Press, Cambridge. Google Scholar
- Wang L, Lin Y-L, Li B et al (2014) Aberrant promoter methylation of the cadherin 13 gene in serum and its relationship with clinicopathological features of prostate cancer. *J Int Med Res* 42:1085–1092
- Warner JR, Soeiro R, Birnboim HC et al (1966) Rapidly labeled HeLa cell nuclear RNA. I. Identification by zone sedimentation of a heterogeneous fraction separate from ribosomal precursor RNA. *J Mol Biol* 19:349–361
- Wassenegger M, Heimes S, Riedel L, Sanger HL (1994) RNA-directed de novo methylation of genomic sequences in plants. *Cell* 76:567–576
- Waterhouse PM, Graham MW, Wang MB (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc Natl Acad Sci USA* 95:13959–13964
- Watson JD, Crick FH (1953) The structure of DNA. *Cold Spring Harb Symp Quant Biol* 18:123–131
- Weintraub H, Tapscott SJ, Davis RL et al (1989) Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc Natl Acad Sci USA* 86:5434–5438
- Wilson VL, Jones PA (1983) DNA methylation decreases in aging but not in immortal cells. *Science* 220:1055–1057
- Wood JG, Helfand SL (2013) Chromatin structure and transposable elements in organismal aging. *Front Genet* 4:274
- Wyatt GR (1951) Recognition and estimation of 5-methylcytosine in nucleic acids. *Biochem J* 48:581–584
- Zhang W, Li J, Suzuki K et al (2015) Aging stem cells. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. *Science* 348:1160–1163
- Zhao M, Zhou Y, Zhu B et al (2016) IFI44L promoter methylation as a blood biomarker for systemic lupus erythematosus. *Ann Rheum Dis* 75:1998–2006