MolecularMedicine

Genetic Determinants of Epigenetic Patterns: Providing Insight into Disease

Emma Cazaly,[1] Jac Charlesworth,[1] Joanne L Dickinson,[1] and Adele F Holloway[2]

[1]Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia; and the [2]School of Medicine, University of Tasmania, Hobart, Tasmania, Australia

The field of epigenetics and our understanding of the mechanisms that regulate the establishment, maintenance and heritability of epigenetic patterns continue to grow at a remarkable rate. This information is providing increased understanding of the role of epigenetic changes in disease, insight into the underlying causes of these epigenetic changes and revealing new avenues for therapeutic intervention. Epigenetic modifiers are increasingly being pursued as therapeutic targets in a range of diseases, with a number of agents targeting epigenetic modifications already proving effective in diseases such as cancer. Although it is well established that DNA mutations and aberrant expression of epigenetic modifiers play a key role in disease, attention is now turning to the interplay between genetic and epigenetic factors in complex disease etiology. The role of genetic variability in determining epigenetic profiles, which can then be modified by environmental and stochastic factors, is becoming more apparent. Understanding the interplay between genetic and epigenetic factors is likely to aid in identifying individuals most likely to benefit from epigenetic therapies. This goal is coming closer to realization because of continual advances in laboratory and statistical tools enabling improvements in the integration of genomic, epigenomic and phenotypic data.

Online address: http://www.molmed.org
doi: 10.2119/molmed.2015.00001

WHY EPIGENETICS AND WHY NOW?

The genomics era brought with it dramatic advances in our understanding of the molecular basis of disease. High-density genome mapping strategies have proven particularly successful for the identification of genes underlying mendelian disorders, such as hemochromatosis, cystic fibrosis and muscular dystrophy (1). The advent of genome-wide association studies (GWAS) was heralded with the promise of providing a comprehensive map of genetic susceptibility to complex disease. While uncovering thousands of variants associated with disease risk, their promise is yet to be fully realized, with a persistent gap emerging between the fraction of disease accounted for by genetic variation and the heritability estimates for many traits (2). Several explanations have been proposed for this unexplained genetic component to disease susceptibility, including the impact of large deletions, inversions or copy number variants, complex gene–gene and gene–environment interactions, overestimated heritability, poor modeling and statistical application and common variants masking rare variants or driving synthetic association (2). Notably, deleterious variants occurring in coding regions account for the minority of disease-associated single nucleotide polymorphisms (SNPs), with estimates that over 90% of variants identified in GWAS are located in noncoding regions of the genome (3). At least some of these SNPs affect gene regulatory mechanisms, modifying gene expression by altering transcription factor binding and directing altered epigenetic profiles (4).

Studies involving genetically identical individuals such as monozygotic (MZ) twins have been invaluable in investigating the role of genetics and environment in complex disease. While providing insight into the genetic basis of disease, these studies have also strongly implicated a nongenetic contribution to many diseases. For example, MZ twins have much less than 100% concordance rates for common diseases such as Alzheimer's disease and certain cancers (5). The effect of the environment and random factors on the epigenome poses a possible explanation for this discordance, since while MZ twin epigenetic profiles show a high level of heritability early in development, they diverge with age and differing lifestyles and epigenetic marks differ according to disease state (5).

Address correspondence to *Adele F Holloway, School of Medicine, University of Tasmania, Private Bag 34, Hobart Tasmania 7000, Australia. Phone: +61-(0)3-6226-2670; Fax: +61-(0)3-6226-7704; E-mail: a.f.holloway@utas.edu.au.*
Submitted January 6, 2015; Accepted for publication March 26, 2015; Published Online (www.molmed.org) March 26, 2015.

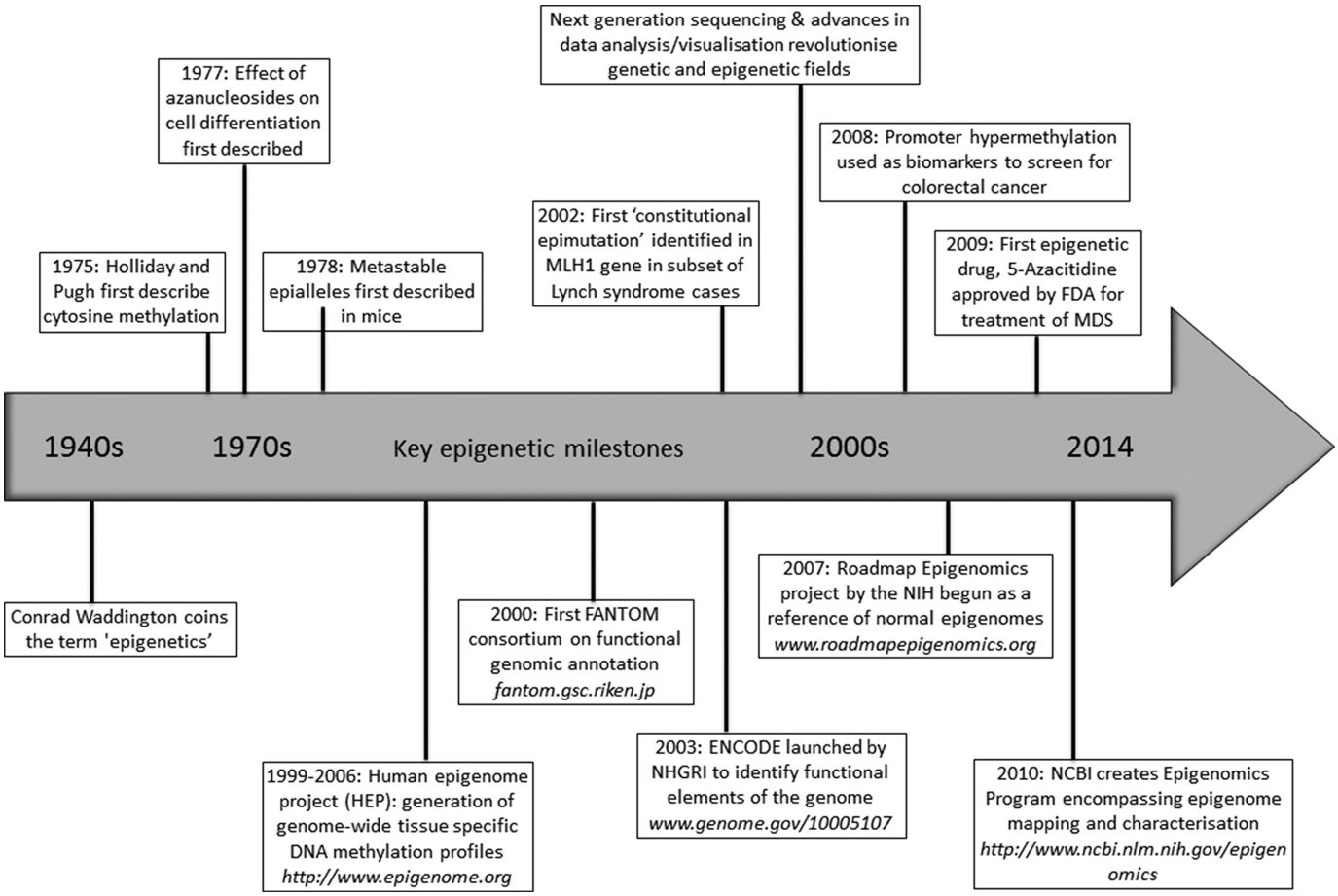

Figure 1. Timeline of key advances in the epigenetics field. Key milestones in the epigenetics field are shown.

Our understanding of the factors affecting the epigenome is increasing at a rapid rate. The role of the underlying genetic sequence in determining epigenetic profiles and the mechanisms by which environmental and stochastic factors then modify these epigenetic patterns are becoming clearer. Our increased understanding of these mechanisms and their role in disease processes is being driven by rapid advances in laboratory and statistical tools and the creation of extensive public databases. This result makes it possible to integrate genomic, epigenomic and phenotypic data with a greater level of detail and scale than ever before (6) (see Figure 1 for a timeline of epigenetic milestones, including the advent of online annotation databases). This information is providing a valuable resource for the investigation of complex disease, revealing new opportunities for disease prevention in at-risk individuals, identifying new therapeutic targets and providing the prospect of increased sensitivity and specificity of disease diagnosis (7).

A BRIEF HISTORY OF EPIGENETICS

Conrad Waddington first coined the term "epigenetics" in the early 1940s to integrate the existence of two related phenomena: that genetically identical cells possess the capacity to differentiate into tissue-specific structures with correlated functions and that gene–environment interactions can affect phenotypes (reprinted in Waddington [8]). The term "epigenetics" has since come to refer to the environment surrounding the DNA, with a current working definition characterizing an epigenetic trait as "a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" (9). This term is most often used in reference to the inheritance of traits to a daughter cell during mitosis, but there is evidence, although still controversial, of germ line transmission of epigenetic traits between generations (transgenerational inheritance) (10,11).

Eukaryotic DNA is assembled into chromatin; repeating units of nucleosomes consisting of 147 base pairs of DNA wrapped around an octamer of histones. Mechanisms that affect the genomic environment include modifications to the DNA itself, absence/presence of histone modifications and histone variants, and also processes involving noncoding

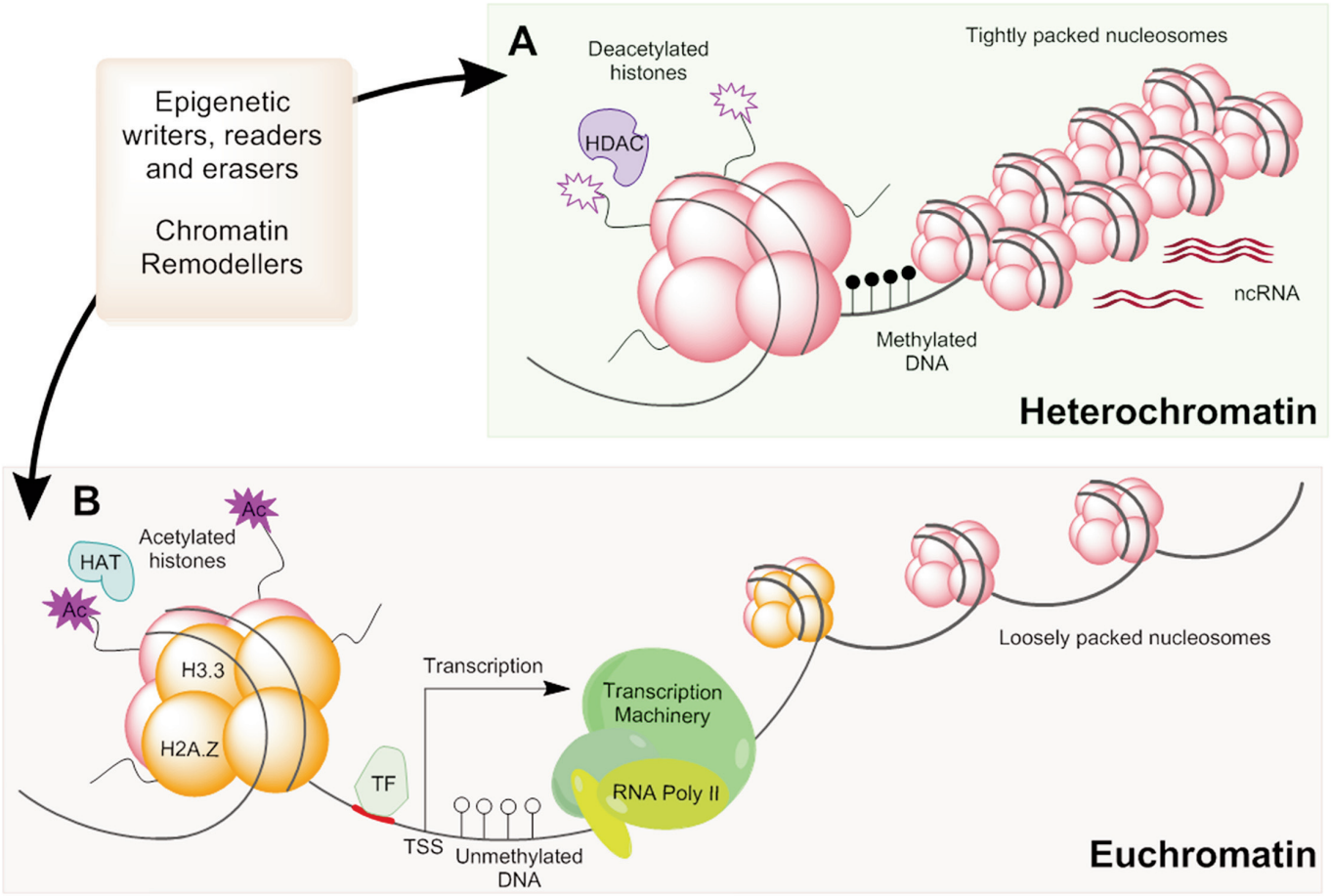

Figure 2. Epigenetic characteristics of heterochromatin and euchromatin. Nucleosome complexes consisting of eight core histone proteins and 147 bp of DNA are configured in higher order structures that regulate the accessibility of DNA to the transcriptional machinery. (A) Heterochromatin is characterized by DNA hypermethylation (filled black circles), de-acetylated histones (clear purple stars) and tightly packed nucleosomes. Noncoding RNA is involved in maintaining heterochromatin structure. (B) Euchromatin contains unmethylated CpGs (unfilled circles), acetylated histones (purple stars) and histone variants H2A.Z and H3.3 and is more accessible to the transcription machinery (green). Writers, readers and erasers of epigenetic modifications work in concert with chromatin remodeling complexes to move and modify nucleosomes, altering chromatin composition. TF, transcription factor; TSS, transcription start site.

RNA and chromatin remodeling complexes (12) (Figure 2). Although there is contention over the extent of heritability of the outcomes of many of these mechanisms and whether they should therefore strictly be regarded as epigenetic mechanisms (9), it is clear that they all influence genome function.

EPIGENETIC MECHANISMS

DNA Methylation

DNA methylation, the addition of a methyl group to a cytosine residue immediately preceding a guanine (CpG dinucleotides), is the most widely studied epigenetic modification. CpG dinucleotides are enriched in clusters called CpG islands, associated with the promoter regions of up to 60% of genes (13). DNA methylation is generally associated with gene silencing, inhibiting gene expression by recruiting proteins that facilitate chromatin condensation and to a lesser extent by physically blocking transcription factor binding (13).

DNA methylation patterns are influenced by the underlying genetic sequence, stochastic changes, environmental factors and other epigenetic mechanisms, resulting in differing methylation patterns across populations, age, tissue and loci (14). DNA methyltransferases (DNMTs) ensure tissue-specific DNA methylation patterns, once established, are maintained through mitosis with high precision and fidelity (13). In contrast, at a global level, methylation is substantially wiped clean during gametogenesis to provide the developing embryo the capacity for totipotency and prevent accumulation of epigenetic changes from previous generations (15). This reprogramming occurs in two waves: first, during pre-implantation

and, second, after primordial germ cell migration, including the removal of imprinted marks (16).

Genomic reprogramming involves both active and passive demethylation, which has long perplexed the scientific community because of their inability to identify the enzyme responsible for this demethylation. Insight into this process has been gained only recently after the discovery of the additional DNA modification, hydroxymethylation, enriched in Purkinje cells in the brain (17) and also embryonic stem cells (18,19). The ten-eleven translocation (TET) family of proteins are responsible for this modification, and there is evidence that it is an intermediate in the process of active demethylation (20). However, the high levels of the modification, particularly in neural cells and during embryonic development, suggests that it may have a yet-to-be-elucidated role in regulating the genome also.

Chromatin Conformation and Structure

Packaging of the vast quantity of DNA into the eukaryotic nucleus is facilitated by the assembly of DNA into nucleosomes followed by their compaction into higher-order structures. This structural organization also plays a fundamental role in regulating accessibility of the DNA to the cellular transcription machinery (12) (Figure 2).

Up to a dozen posttranslational modifications of histone proteins have been reported to date, including methylation, acetylation, phosphorylation and ubiquitination (12), and technological advances have made it possible to map these modifications genome-wide (21). These modifications have together been proposed to form a histone code, which can be interpreted by cellular proteins to specify downstream functions (22). In addition, chromatin structure is altered by the actions of ATP-dependent chromatin remodeling enzymes and the exchange of canonical histones with histone variants. This step creates a highly dynamic, adaptable epigenetic landscape that plays a key role in regulating genome function and provides an interface between the environment and the genome. Although histone modifications can be subject to rapid turnover, there is also evidence that they can be stably inherited during cell division and thus contribute to the maintenance of cellular identity. However, this apparent epigenetic memory may be initiated by other factors such as DNA methylation, noncoding RNA (ncRNA), transcription factors or the underlying DNA sequence (12). Three recent studies point to an important role of genetic variants in determining histone modification patterns (23–25). In these studies, hundreds of variants were associated with changes to histones and gene expression, with the underlying mechanism thought to be altered transcription factor binding.

Noncoding RNA

In recent years, the existence of a complex network of ncRNAs transcribed from the human genome has become apparent. These ncRNAs have regulatory functions and play a key role in the establishment and maintenance of other epigenetic marks (26), with evidence that they constitute a mechanism for transgenerational epigenetic inheritance (27). There is also mounting evidence for the involvement of ncRNAs in disease development, particularly in cancer (28).

GENETIC DETERMINANTS OF EPIGENETIC PATTERNS

Twin Studies

Diminished genetic noise and the innate advantage of perfect age and often sex matching, frequently coupled with similar environmental and socioeconomic upbringing, ensures twin studies play an invaluable role in understanding epigenetic mechanisms (29). The first twin studies examining DNA methylation focused on X-chromosome inactivation, finding that the selection of which X-chromosome is inactivated is not as random as previously thought, but is influenced to a degree by underlying heritable patterns (30). Vickers *et al.* (30) also showed that, with increasing age, there was a greater skew in inactivation patterns, suggesting twins become epigenetically dissimilar with age. This result was later reinforced by several larger studies that found a high degree of epigenetic heritability among MZ twins that decreased with age and that found larger discrepancies between twins that lead different lifestyles (5).

More recently, the advent of methylation array technology has enabled more extensive studies that have found MZ twins to be more epigenetically similar than dizygotic twins. These studies also found that the most heritable CpG sites correlated with functional regions and promoters, indicating these regions are under tighter genetic control (31). Additionally, MZ DNA methylation patterns may be influenced by the chorionicity of the prenatal environment. Perhaps, surprisingly, monochorionicity (a single shared placenta) has been linked to more divergent methylation patterns, yet chorionicity is not always accounted for in MZ twin studies (32). These studies provide evidence for epigenetic differences in genetically identical individuals, suggesting epigenetic metastability independent of genotype can occur in humans, as has been previously shown in mouse models (33). However, they also point to certain regions where genetic influences exert greater control on the epigenome.

Further evidence from studies examining DNA methylation patterns among family members and unrelated individuals found SNPs induce subtle epigenetic variation. One of the first studies to examine the relationship between genetic variants and DNA methylation patterns found evidence for allele-specific methylation (ASM) outside imprinted regions, some linked to allele-specific gene expression (34). Examination of the effect of genetic variants on DNA methylation in a three-generation family and among unrelated individuals found that heterozygous SNPs associated with different methylation patterns (35). This differen-

tial methylation also correlated with gene expression. Globally, genetic variants were found to be more influential than imprinting, and most of the ASM (75%) observed in the family was also present in unrelated individuals, suggesting genotype influences heritable regions of differential methylation.

Other studies estimate around 20% of heterozygous SNPs are linked to ASM (36). Alternative terms have been proposed for these variants, including methylation-associated SNPs (mSNPs) (37), CpG-SNPs (36) and methylation quantitative trait loci (meQTL) (38). Hundreds of these genetic variants have been linked to DNA methylation patterns and, similar to genetic quantitative trait loci, these variants can affect gene expression and phenotype (34,36–38). While SNPs in the vicinity of CpG sites influence methylation levels, the most obvious variant affecting methylation is a mutation at the CpG site itself. Indeed, most of the SNPs linked to ASM are located at the CpG site and are designated meSNPs (36,39). These meSNPs also influence the methylation levels of neighboring CpGs, particularly those close by (within 45 base pairs [bp]) but have also been show to affect CpGs up to 10 kb away (39).

GENETIC VARIANTS CAUSING EPIGENETIC CHANGE IN DISEASE

Although there is no shortage of studies demonstrating a role for epigenetic changes in driving disease, there are now a number of examples in which *cis*-acting variants have been clearly demonstrated to drive the disease-associated epimutations. The term "epimutation" refers to an altered epigenetic state resulting in altered transcriptional activity of a gene. Such *cis*-acting variants have been shown to alter DNA methylation patterns and gene expression in a variety of human tissues (34,36). These changes are likely an indirect result of altered binding of transcription factors, which can either lead to altered recruitment of chromatin modifiers and remodelers and subsequent epigenetic changes or can cause changes in gene expression that predispose the gene to epigenetic silencing (40). Comprehensive genetic analysis facilitated by improved technology has revealed that several diseases involving epigenetic dysfunction have genetic origins.

One such example is the X-linked neurodevelopmental disorder, fragile X syndrome. The genetic defect involves expansion of a trinucleotide repeat sequence (CGG) at the promoter of the fragile X mental retardation gene (*FMR1*), with up to 45 repeats in unaffected individuals and up to 200 in affected individuals (41). The repeat expansion results in methylation of the region and subsequent epigenetic silencing of the gene.

A more distally acting example of sequence variation contributing to disease through altered gene regulation involves the Myc transcription factor. Activation of the Myc transcription factor is suggested to occur in up to 70% of cancers, arising through a range of mechanisms including translocations, gene amplification, enhanced protein translation and stability, or indirectly through signaling pathways that regulate Myc (42). In addition, a number of GWAS have found multiple SNPs on chromosome 8q24 associated with different types of cancer (43). These SNPs occur in a gene desert but have since been found to influence regulation of the Myc oncogene located hundreds of kilobases away. These regions have been shown to contain distal enhancers of the *Myc* gene and highlight that changes to long-range chromatin structures can result in altered gene expression (44).

Imprinting Disorders

Imprinted genes escape the initial phase of epigenetic reprogramming after fertilization, retaining their parental methylation marks and are expressed in a parent of origin manner (45). Around 100 such imprinted genes are currently known, with this number continuing to increase (45). Imprinting provides clear evidence that epigenetic modifications can be inherited through meiosis, and specific diseases ensue when there is an abnormality in either the erasure of existing marks or reestablishment and maintenance of new marks (45). These diseases can result from underlying genetic defects or be due to epimutations (46).

Prader-Willi syndrome and Angelman syndrome are neurological disorders with distinct phenotypes that occur when the same imprinted region on chromosome 15 is nonfunctional. While the vast majority of cases are caused by a single gene mutation or chromosomal deletion, between 1% (Prader-Willi syndrome) and up to 4% in Angelman syndrome are due to an imprinting defect, with the majority of these being primary epimutations, occurring in the absence of DNA sequence mutations (47). Maternally inherited defects lead to Angelman syndrome, whereas paternal imprinting errors lead to Prader-Willi syndrome (47).

Several different disorders result from disruption of epigenetic regulation at the *IGF2/H19* locus, a region in which heritable factors have been shown to have a greater impact on DNA methylation than the accumulation of stochastic and environmental-induced changes (48). DNA methylation at the imprinting control region upstream of the paternal *H19* allele normally silences *H19* expression and activates *IGF2*, whereas the maternal *IGF2* allele is silenced (49). In the Silver-Russell syndrome, a rare developmental disorder, 45% of cases are attributed to an epimutation in the imprinting control region of the paternal *H19* allele (50).

Beckwith-Weidemann syndrome is a congenital overgrowth syndrome with 83% of cases occurring sporadically, of which around 60% are thought to involve epimutations of two imprinting control regions regulating *H19*, *IGF2*, *KCNQ1* and *CDKN1C* (51). Female MZ twins represent a high proportion of Beckwith-Weidemann syndrome cases, and a study of five discordant female MZ twins found that all affected twins had a defect at the imprinted locus *KCNQ1OT1*, which encodes a noncoding RNA that regulates the expression of other imprinted genes (52). Loss of imprinting at the *IGF2/H19* locus is also an

epigenetic cause of around one-third of Wilms tumor, the most common renal cancer in children (49). The defect is also associated with heightened colorectal cancer risk (53) and esophageal squamous cell carcinoma (54).

Lynch Syndrome

Whereas epigenetic changes are well described in disease, particularly cancer, there are now several clear examples of cancer-associated epigenetic changes being genetically driven. These types of interactions can help to explain features of these diseases such as late onset, environmental effects, tissue specificity and also familial associations that do not follow mendelian inheritance patterns. Combining genomic and epigenomic data is also proving to be of value in the search for prognostic signatures in cancer (55), as seen in the Lynch syndrome, an autosomal dominant cancer susceptibility condition. Approximately two-thirds of Lynch syndrome cases result from heterozygous loss-of-function mutations in DNA mismatch repair genes, most commonly mutL homolog 1 (*MLH1*) and mutS protein homolog 2 (*MSH2*) (56).

However, such mutations are not apparent in around one-third of Lynch syndrome cases, some of which (~4% for *MLH1* [56]) can be explained by epimutations in *MLH1* and *MSH2*. These epimutations lead to transcriptional inactivation of the gene, essentially having the same effect as a genomic sequence mutation seen in other Lynch syndrome cases. One possible mechanism underlying these epimutations involves primary DNA methylation changes independent of any sequence change, resulting in labile epimutations, which can be reversed in the germline and are therefore inherited in an unpredictable, nonmendelian manner or not passed on at all.

Alternatively, secondary epimutations may result from underlying sequence changes, including promoter deletions and SNPs (57); for example, the c.-27C>A germline variant in the 5′UTR of the *MLH1* gene has been linked to cancer susceptibility through transcriptional silencing (58). In these cases, the disease follows a more predictable inheritance pattern, since the epimutation is driven by a genetic variant. As yet undiscovered sequence mutations may also be the underlying carcinogenic mechanism in subsets of cancers such as Cowden syndrome, where some individuals have hypermethylation epimutations in the absence of known sequence mutations (59).

Underlying genetic drivers have also been linked to epimutations in sporadic cases of renal cell cancer, where SNPs were associated with promoter hypermethylation of the von Hippel-Lindau (*VHL*) gene in tumor tissue, a gene previously shown to be genetically altered in individuals with the familial form of the cancer (60). Similarly, in colorectal cancer, a C>T point mutation at an enhancer element of the mismatch repair gene O(6)-methylguanine-DNA methyltransferase (*MGMT*) has been linked to aberrant promoter methylation and gene silencing (61). Given the recent technological advances that are enabling integration of genetic, epigenetic and phenotypic data, it is likely that more of examples of diseases resulting from genetic drivers of epigenetic change will be described in the future.

Genetic Mutations in Epigenetic Modifiers

Mutations in genes encoding epigenetic modifiers also contribute to complex diseases, again with cancer being the best described example. Translocations, mutations or overexpression of modifiers such as DNMTs, histone modifying enzymes or chromatin remodeling proteins are well documented in many cancers (62). Aberrant epigenetic modifiers can directly affect regulation of target genes as well as interacting with specific genetic variants of common disease-causing SNPs (63). Whereas the study of other complex diseases are at an earlier stage, there is accumulating evidence to suggest that disruption to epigenetic modifiers plays a role in a range of other diseases, including diabetes, immune diseases and intellectual disabilities such as autism (63).

The simplest example of a genetic mutation driving an epigenetic change and contributing to disease is when the change occurs in a gene encoding an epigenetic modifying enzyme. Some of the more recently described examples include Kabuki syndrome with mutations in the histone methyltransferase gene *MLL2* (64) and Coffin-Siris syndrome involving mutations in SWI/SNF subunit genes (65). Such disorders have been recently reviewed (66). Another straightforward example of this is in the ICF syndrome (immunodeficiency, centromeric instability and facial anomalies), which usually arises because of biallelic mutations in the gene encoding the DNMT3B methylating enzyme (67). The syndrome is a consequence of loss of DNMT activity resulting in genomic hypomethylation. Genomic hypomethylation is a rare disease, which is invariably fatal in early childhood. However, the disease displays phenotypic variability, which is likely due to the differing effects of individual mutations on DNMT3B activity (67).

Rett syndrome, an X-linked neurodevelopmental disorder usually affecting girls, is also due to genetic defects in an epigenetic modifier (in this case, the methyl CpG binding protein MeCP2) (68). The disease displays delayed onset, with children developing normally until 1–2 years of age, when they present with progressive neurological dysfunction. The largely neurological phenotype of this disease is likely a result of the requirement for tight regulation of a number of important neural targets of MeCP2 (69). Variability in disease phenotype is likely a function of the range of mutations that give rise to the disease as well as an effect of X-inactivation skewing.

THE EPIGENOME: AN INTERFACE BETWEEN THE GENOME AND THE ENVIRONMENT

While the underlying genome plays a role in determining epigenetic profiles, stochastic factors and environmental cues including diet, exercise and toxins bring

about subsequent changes in the epigenome, as previously reviewed comprehensively (70). There is evidence that at least some of these changes can then be passed down through meiosis as transgenerational epigenetic inheritance (26,71,72). Whereas the evidence for this and mechanisms involved are beyond the scope of this article, they have recently been comprehensively reviewed (11).

Stochastic Factors

It is now thought that randomly induced epigenetic patterns may also contribute to variation in development and aging as well as providing a possible mechanism for the rapid selection of epigenotypes in response to environmental pressures. X-chromosomal inactivation is a classic example of how epigenetic profiles can be regulated by stochastic factors. These factors may also explain discordance between MZ twins (73).

While the effect of environment on epigenetic profiles, particularly DNA methylation, is widely acknowledged, stochastic changes may in fact be more common than environmentally induced changes, with a study examining 4,000 human genes, finding 300 to have random monoallelic expression (74). Epigenetic stochasticity can be defined as a combination of epigenetic variation in the germline and somatic instability. Similar to Richards' "facilitated epigenetic variation" model (75), Feinberg and Irizarry's "inherited stochastic variation model" proposes genetic sequence variation underlies the propensity for epigenetic variation, since certain DNA sequences are not only directly responsible for particular traits but also increase natural methylation variation for that trait (76). Various stochastic and environmental factors then influence DNA methylation at these variably methylated regions, resulting in increased phenotypic differences, which are then acted on by Darwinian selection in a similar manner to selection pressures affecting purely genetic traits. Subsequent studies found the sites of greatest DNA methylation variability in colon cancer corresponded to the sites of greatest variability in other cancers, including lung, breast and ovarian cancers, with these sites normally having distinct tissue-specific DNA methylation patterns (77). Thus, heritable DNA methylation variation could provide some contribution to the unexplained heritable genetic component of common complex diseases.

THE PROMISE OF EPIGENETIC THERAPY

Whereas genetic mutations and chromosomal defects permanently alter the genome, epigenetic alterations, whether driven by changes to the underlying genome or by environmental or stochastic influences, can potentially be pharmacologically reversed or modified, providing the promise of restoring gene function, altered as a result of epigenetic changes in disease. There is currently considerable interest in the development and clinical translation of pharmacological agents that target either the writers or the readers of the epigenetic code. Because these are mainly enzymes, they provide an easier target than other gene regulators such as transcription factors. For obvious reasons, the use of such agents is at the most progressed stage in the treatment of cancers, with two DNMT and two HDAC inhibiting compounds approved for cancer treatment in the United States and numerous others in clinical trials (78) (Figure 3).

In 2004, 5-azacytidine (aza-C) became the first U.S. Food and Drug Administration (FDA)-approved epigenetic drug when it was approved for treatment of myelodysplastic syndrome (79). An analog of cytidine, the nucleoside is incorporated into nucleic acid of dividing cells with a preference for RNA over DNA. The presence of nitrogen at carbon-5 blocks the addition of a methyl group by DNMTs, preventing the methylation of the DNA after cell division. The bound DNMTs, unable to detach due to the nitrogen, form permanent adducts with the nucleic acid resulting in functional depletion of DNMT from the cell (Figure 3C). In addition, DNA replication is blocked (80), and the DNA is functionally compromised, activating the p53 damage pathway, leading to degradation of the DNA (81) (Figure 3D).

Aza-2-deoxycytidine (decitabine) is the deoxy form of aza-C and is solely incorporated into DNA, avoiding the indirect effects on RNA and protein synthesis of 5-aza-C. Approved in 2006 for treatment of myelodysplastic syndrome, it is the only other demethylating agent currently approved by the FDA. Other cytidine analogs such as zebularine and 5-fluoro-deoxycytidine are in clinical trials, as are direct DNMT inhibitors including procaine, procanamide and hydralazine (78). The two other FDA-approved epigenetic drugs fall under the umbrella of histone deacetylase inhibitors. Both vorinostat in 2006 and romidepsin in 2009 were approved for treatment of cutaneous T-cell lymphoma (78) (Figure 3A). Interest in targeting epigenetic modifiers in cancer and other diseases continues to grow, with a wide range of targets now being explored in both preclinical and clinical trials. For example, inhibitors of several histone methyltransferases, including EZH2 and DOT1L are also in preclinical trials for certain lymphomas and leukemias, respectively (78) (Figure 3B).

Combining various epigenetic therapies may prove to be the most effective strategy because of the high degree of biological interaction between DNA methylation, histone modifications and chromatin remodeling complexes. Indeed, clinical trials examining the synergistic action of these therapies are promising (82). These therapies are also most likely to be effective when combined with conventional cancer treatments (78). Epigenetic pharmaceuticals are still in their infancy, and clearly more is to be learned about the underlying mechanisms determining epigenetic states, since only some cancers can be reprogrammed to a normal state and demethylating agents and HDAC inhibitors are unable to bring about permanent expression changes. This result is particularly true if the abnormal epigenetic state is driven by underlying ge-

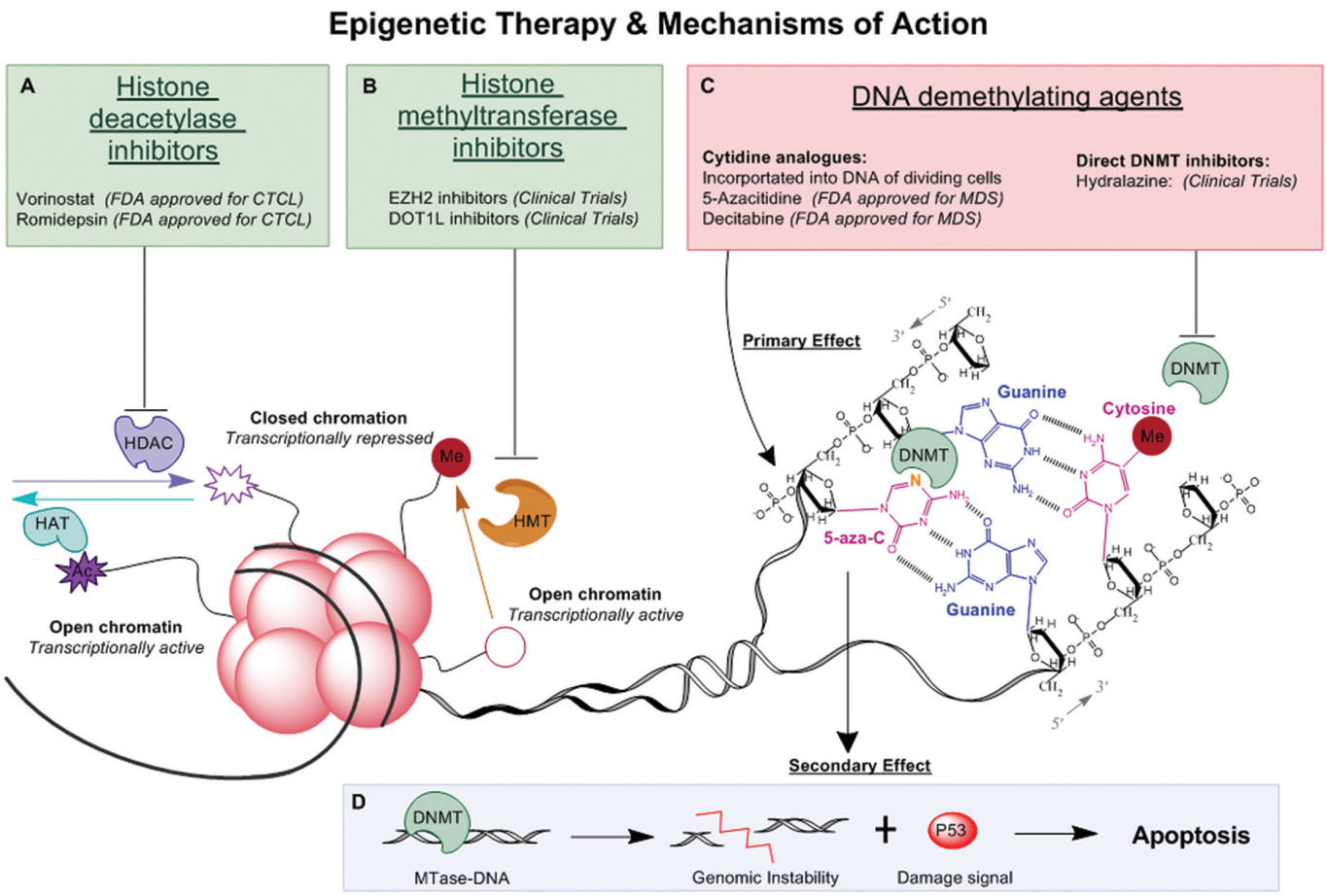

Figure 3. Advances in epigenetic therapy. Epigenetic modifiers and modifications provide targets for therapeutic intervention in disease. (A) Two histone deacetylase inhibitors are FDA approved for treatment of subtypes of leukemia. (B) Inhibitors of a range of epigenetic modifiers, including histone methyltransferase enzymes, are in preclinical trials. (C) DNA demethylating agents decrease genomic methylation, restoring aberrantly silenced gene expression by acting directly to inhibit methylating enzymes or as cytidine analogs are incorporated into nucleic acid of dividing cells, preventing methylation. (D) Cytidine analogs have the secondary effect of activating the p53 damage pathway and inducing apoptosis.

netic factors and would therefore require ongoing pharmacological intervention to prevent reversion of the epigenetic state.

Disease screening and diagnosis may also be vastly improved with the inclusion of epigenetic information. In 2012, Teschendorff *et al.* (83,84) were able to predict risk of cervical neoplasia 3 years before morphological changes by examining DNA methylation variability. Feinberg and Irizarry (85) suggest that if such tests are used to identify subgroups for further traditional follow-up screening (for example, mammogram, colonoscopy), the positive predictive value of these more invasive and expensive tests could rise to over 90% and greatly reduce cancer deaths.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Many questions remain regarding the underlying mechanisms that determine epigenetic patterns. Rapid improvements in technology and our increasing ability to effectively analyze and interpret the immense quantities of data produced are allowing the generation of comprehensive epigenomic maps and characterization of the differences in epigenetic state between individuals and the changes that occur during development, aging and disease processes. This information is also aiding in our understanding of the underlying mechanisms involved, including the relative contribution of environmental influences, stochastic factors and genetic variants. Numerous studies across a range of tissue types and populations are providing strong evidence for a key role for genetic variants in establishing inherited methylation patterns.

Over recent years, we have gained considerable insight into the role of genetic variants and epigenetic change in diseases. Attention is now turning to understanding the interactions between genetic and epigenetic factors and their concerted roles in disease processes. This

approach is providing advances in disease prevention, diagnosis and surveillance. It is also offering hope that a heightened understanding of how inherited factors regulate gene expression through epigenetic mechanisms will provide more personalized diagnostic tools and effective treatments for complex disease.

ACKNOWLEDGMENTS

The authors were supported by grants from the David Collins Leukaemia Foundation and The Cancer Council Tasmania. JL Dickinson is an Australian Research Council Future Fellow. E Cazaly was supported by a scholarship from the Royal Hobart Hospital Cancer Auxiliary.

The authors acknowledge the many researchers whose work has contributed to this field, but they are not specifically cited here because of space and referencing constraints.

DISCLOSURE

The authors declare that they have no competing interests as defined by *Molecular Medicine*, or other interests that might be perceived to influence the results and discussion reported in this paper.

REFERENCES

1. Botstein D, Risch N. (2003) Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat. Genet.* 33:228–37.
2. Eichler EE, *et al.* (2010) Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.* 11:446–50.
3. Hindorff LA, *et al.* (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U. S. A.* 106:9362–7.
4. Furey TS, Sethupathy P. (2013) Genetics: genetics driving epigenetics. *Science.* 342:705–6
5. Steves CJ, *et al.* (2012) Ageing, genes, environment and epigenetics: what twin studies tell us now, and in the future. *Age Ageing.* 41:581–6.
6. Laird PW (2010) Principles and challenges of genome-wide DNA methylation analysis. *Nat. Rev. Genet.* 11:191–203.
7. Montavon C, *et al.* (2012) Prognostic and diagnostic significance of DNA methylation patterns in high grade serous ovarian cancer. *Gynecol. Oncol.* 124:582–8.
8. Waddington CH. (2012) The epigenotype. *Int. J. Epidemiol.* 41:10–3.
9. Berger SL, *et al.* (2009) An operational definition of epigenetics. *Genes Dev.* 23:781–3.
10. Skinner MK. (2011) Environmental epigenetic transgenerational inheritance and somatic epigenetic mitotic stability. *Epigenetics.* 6:838–42.
11. Pembrey M, *et al.* (2014) Human transgenerational responses to early-life experience: potential impact on development, health and biomedical research. *J. Med. Genet.* 51:563–72.
12. Dawson MA, Kouzarides T. (2012) Cancer epigenetics: from mechanism to therapy. *Cell.* 150:12–27.
13. Cedar H, Bergman Y. (2012) Programming of DNA methylation patterns. *Annu. Rev. Biochem.* 81:97–117.
14. ENCODE Project Consortium. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature.* 489:57–74.
15. Seisenberger S, *et al.* (2012) The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell.* 48:849–62.
16. Sasaki H, Matsui Y. (2008) Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat. Rev. Genet.* 9:129–40.
17. Kriaucionis S, Heintz N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science.* 324:929–30.
18. Ito S, *et al.* (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature.* 466:1129–33.
19. Cimmino L, *et al.* (2011) TET family proteins and their role in stem cell differentiation and transformation. *Stem Cell.* 9:193–204.
20. Hackett JA, *et al.* (2013) Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science.* 339:448–52.
21. Barski A, *et al.* (2007) High-resolution profiling of histone methylations in the human genome. *Cell.* 129:823–37.
22. Wang Y, *et al.* (2004) Beyond the double helix: reading and writing the histone code. *Novartis Found. Symp.* 259:3–17.
23. McVicker G, *et al.* (2013) Identification of genetic variants that affect histone modifications in human cells. *Science.* 342:747–9.
24. Kilpinen H, *et al.* (2013) Coordinated effects of sequence variation on DNA binding, chromatin structure, and transcription. *Science.* 342:744–7.
25. Kasowski M, *et al.* (2013) Extensive variation in chromatin states across humans. *Science.* 342:750–2.
26. Daxinger L, Whitelaw E. (2012) Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat. Rev. Genet.* 13:153–62.
27. Liebers R, *et al.* (2014) Epigenetic regulation by heritable RNA. *PLoS. Genet.* 10:e1004296.
28. Kasinski AL, Slack FJ. (2011) Epigenetics and genetics. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy. *Nat. Rev. Cancer.* 11:849–64.
29. Bell JT, Spector TD. (2011) A twin approach to unraveling epigenetics. *Trends Genet.* 27:116–25.
30. Vickers MA, *et al.* (2001) Assessment of mechanism of acquired skewed X inactivation by analysis of twins. *Blood.* 97:1274–81.
31. Kaminsky ZA, *et al.* (2009) DNA methylation profiles in monozygotic and dizygotic twins. *Nat. Genet.* 41:240–5.
32. Race JP, *et al.* (2006) Chorion type, birthweight discordance and tooth-size variability in Australian monozygotic twins. *Twin Res. Hum. Genet.* 9:285–91.
33. Morgan HD, *et al.* (1999) Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* 23:314–8.
34. Kerkel K, *et al.* (2008) Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. *Nat. Genet.* 40:904–8.
35. Gertz J, *et al.* (2011) Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. *PLoS. Genet.* 7:e1002228.
36. Shoemaker R, *et al.* (2010) Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Res.* 20:883–9.
37. Zhang D, *et al.* (2010) Genetic control of individual differences in gene-specific methylation in human brain. *Am. J. Human Genet.* 86:411–9.
38. Bell JT, *et al.* (2011) DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* 12:R10.
39. Zhi D, *et al.* (2013) SNPs located at CpG sites modulate genome-epigenome interaction. *Epigenetics.* 8:802–6.
40. Hesson LB, *et al.* (2010) Epimutations and cancer predisposition: importance and mechanisms. *Curr. Opin. Genet. Dev.* 20:290–8.
41. Fu Y-H, *et al.* (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell.* 67:1047–58.
42. Meyer N, Penn LZ. (2008) Reflecting on 25 years with MYC. *Nat. Rev. Cancer.* 8:976–90.
43. Haiman CA, *et al.* (2007) Multiple regions within 8q24 independently affect risk for prostate cancer. *Nat. Rev. Cancer.* 39:638–44.
44. Sotelo J, *et al.* (2010) Long-range enhancers on 8q24 regulate c-Myc. *Proc. Natl. Acad. Sci. U. S. A.* 107:3001–5.
45. Ferguson-Smith AC. (2011) Genomic imprinting: the emergence of an epigenetic paradigm. *Nat. Rev. Genet.* 12:565.
46. Banno K, *et al.* (2012) Epimutation and cancer: a new carcinogenic mechanism of Lynch syndrome (Review). *Int. J. Oncol.* 41:793–7.
47. Buiting K, *et al.* (2003) Epimutations in Prader-Willi and Angelman syndromes: a molecular study of 136 patients with an imprinting defect. *Am. J. Hum. Genet.* 72:571–7.
48. Heijmans BT, *et al.* (2007) Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. *Hum. Mol. Genet.* 16:547–54.
49. Ludgate JL, *et al.* (2013) Global demethylation in loss of imprinting subtype of Wilms tumor. *Genes Chromosomes Cancer.* 52:174–84.
50. Fuke T, *et al.* (2013) Molecular and clinical studies in 138 Japanese patients with Silver-Russell syndrome. *PLoS One.* 8:e60105.

51. Murrell A, *et al.* (2004) An association between variants in the IGF2 gene and Beckwith-Wiedemann syndrome: interaction between genotype and epigenotype. *Hum. Mol. Genet.* 13:247–55.
52. Weksberg R, *et al.* (2002) Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith-Wiedemann syndrome. *Hum. Mol. Genet.* 11:1317–25.
53. Cui H, *et al.* (2003) Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science.* 299:1753–5.
54. Murata A, *et al.* (2014) IGF2 DMR0 methylation, loss of imprinting, and patient prognosis in esophageal squamous cell carcinoma. *Ann. Surg. Oncol.* 21:1166–74.
55. Yi JM, *et al.* (2011) Genomic and epigenomic integration identifies a prognostic signature in colon cancer. *Clin. Cancer Res.* 17:1535–45.
56. Ward RL, *et al.* (2013) Identification of constitutional MLH1 epimutations and promoter variants in colorectal cancer patients from the Colon Cancer Family Registry. *Genet. Med.* 15:25–35.
57. Hitchins MP, Lynch HT. (2014) Dawning of the epigenetic era in hereditary cancer. *Clin. Genet.* 85:413–6.
58. Hitchins MP, *et al.* (2011) Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the 5′UTR. *Cancer Cell.* 20:200–13.
59. Bennett KL, *et al.* (2010) Germline epigenetic regulation of KILLIN in Cowden and Cowden-like syndrome. *JAMA.* 304:2724–31.
60. Moore LE, *et al.* (2011) Von Hippel-Lindau (VHL) inactivation in sporadic clear cell renal cancer: associations with germline VHL polymorphisms and etiologic risk factors. *PLoS Genet.* 7:e1002312.
61. Ogino S, *et al.* (2007) MGMT germline polymorphism is associated with somatic MGMT promoter methylation and gene silencing in colorectal cancer. *Carcinogenesis.* 28:1985–90.
62. You JS, Jones PA. (2012) Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell.* 22:9–20.
63. Bjornsson H. (2004) An integrated epigenetic and genetic approach to common human disease. *Trends Genet.* 20:350–8.
64. Ng SB, *et al.* (2010) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* 42:790–3.
65. Tsurusaki Y, *et al.* (2012) Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet.* 44:376–8.
66. Berdasco M, Esteller M. (2013) Genetic syndromes caused by mutations in epigenetic genes. *Hum. Genet.* 132:359–83.
67. Hansen RS, *et al.* (1999) The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc. Natl. Acad. Sci. U.S.A.* 96:14412–7.
68. Liyanage VRB, Rastegar M. (2014) Rett syndrome and MeCP2. *Neuromol. Med.* 16:231–64.
69. Chen WG, *et al.* (2003) Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science.* 302:885–9.
70. Feil R, Fraga MF. (2012) Epigenetics and the environment: emerging patterns and implications. *Nat. Rev. Genet.* DOI: 10.1038/nrg3142.
71. Veenendaal MVE, *et al.* (2013) Transgenerational effects of prenatal exposure to the 1944–45 Dutch famine. *BJOG.* 120:548–53.
72. Guerrero-Bosagna C, Skinner MK. (2011) Environmentally induced epigenetic transgenerational inheritance of phenotype and disease. *Mol. Cell. Endocrinol.* 354:1–6.
73. Petronis A. (2010) Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature.* 465:721–7.
74. Gimelbrant A, *et al.* (2007) Widespread monoallelic expression on human autosomes. *Science.* 318:1136–40.
75. Richards EJ. (2006) Inherited epigenetic variation: revisiting soft inheritance. *Nat. Rev. Genet.* 7:395–401.
76. Feinberg AP, Irizarry RA. (2010) Evolution in health and medicine Sackler colloquium: stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc. Natl. Acad. Sci. U. S. A.* 107:1757–64.
77. Hansen KD, *et al.* (2011) Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* 43:768–75.
78. Bojang P Jr, Ramos KS. (2014) The promise and failures of epigenetic therapies for cancer treatment. *Cancer Treat. Rev.* 40:153–69.
79. Kaminskas E, *et al.* (2005) FDA drug approval summary: azacitidine (5-azacytidine, Vidaza™) for injectable suspension. *Oncologist.* 10:176–82.
80. Kuo HK, *et al.* (2007) 5-Azacytidine-induced methyltransferase-DNA adducts block DNA replication in vivo. *Cancer Res.* 67:8248–54.
81. Karpf AR, *et al.* (2001) Activation of the p53 DNA damage response pathway after inhibition of DNA methyltransferase by 5-aza-2′-deoxycytidine. *Mol. Pharmacol.* 59:751–7.
82. Garcia-Manero G, *et al.* (2006) Phase 1/2 study of the combination of 5-aza-2′-deoxycytidine with valproic acid in patients with leukemia. *Blood.* 108:3271–9.
83. Teschendorff AE, Widschwendter M. (2012) Differential variability improves the identification of cancer risk markers in DNA methylation studies profiling precursor cancer lesions. *Bioinformatics.* 28:1487–94.
84. Teschendorff AE, *et al.* (2012) Epigenetic variability in cells of normal cytology is associated with the risk of future morphological transformation. *Genome Med.* 4:24.
85. Feinberg AP. (2014) Epigenetic stochasticity, nuclear structure and cancer: the implications for medicine. *J. Intern. Med.* 276:5–11.

Cite this article as: Cazaly E, Charlesworth J, Dickinson JL, Holloway AF. (2015) Genetic determinants of epigenetic patterns: providing insight into disease. *Mol. Med.* 21:400–9.