Published in final edited form as:

Acta Neuropathol. 2016 October ; 132(4): 503–514. doi:10.1007/s00401-016-1612-7.

The epigenome in Alzheimer's disease: Current state and approaches for a new path to gene discovery and understanding disease mechanism

Hans-Ulrich Klein^{1,2,3}, David A. Bennett⁴, and Philip L. De Jager^{1,2,3}

¹Program in Translational Neuropsychiatric Genomics and Partners Multiple Sclerosis Center, Ann Romney Center for Neurologic Diseases, Department of Neurology, Brigham and Women's Hospital, Boston MA

²Harvard Medical School, Boston MA

³Program in Medical and Population Genetics, Broad Institute, Cambridge MA

⁴Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago IL

Abstract

The advent of new technologies and analytic approaches is beginning to provide an unprecedented look at features of the human genome that affect RNA expression. These “epigenomic” features are found in a number of different forms: they include DNA methylation, covalent modifications of histone genes and non-coding RNAs. Some of these features have now been implicated in Alzheimer's Disease (AD). Here, we focus on recent studies that have identified robust observations relating to DNA methylation and chromatin in human brain tissue; these findings will ground the next generation of studies and provide a model for the design of such studies. Stemming from observations that compounds with histone deacetylase activity may be beneficial in AD, epigenome-wide studies in cortical samples from large numbers of human subjects have now shown that AD-associated epigenomic changes are reproducible, are not driven by genetic risk factors, and are widespread at specific locations in the genome. A fundamental question of whether such changes are causal remains to be demonstrated, but it is already clear that well-powered investigations of the human epigenome in the target organ of a neurodegenerative disease are feasible, are implicating new areas of the genome in the disease, and will be an important tool for future studies. We are now at an inflection point: as genome-wide association studies of genetic variants come to an end, a new generation of studies exploring the epigenome will provide an important new layer of information with which to enrich our understanding of AD pathogenesis and to possibly guide development of new therapeutic targets.

The National Institute of Health's Roadmap Epigenomics project defines “epigenetics” as the “emerging frontier of science that involves the study of changes in the regulation of gene activity and expression that are not dependent on gene sequence (<http://www.roadmapepigenomics.org/overview>).” Some of these changes are inherited from

parents, but many are acquired and can be stable for long periods of time. Epigenetics is usually used to describe the study of selected genes while epigenomics refers to the study of genome-wide patterns of changes in chromosome and chromatin that lead to changes in RNA expression. Over the last few years, the first generation of epigenome-wide association studies has been conducted in a number of different diseases, including Alzheimer's disease (AD). These studies were made possible because of emerging technologies that could interrogate the state of chromatin throughout the human genome in a moderate to high throughput manner: as with genetic studies a decade ago, this development enabled studies of sufficient size to produce robust results. Given the nascent nature of the field, we are just beginning to understand which study designs, sample sizes and analysis methods yield reproducible results. Epigenomic variation, unlike genetic variation, is highly plastic, responding to an individual's environment and life experiences; ultimately, every cell has its unique epigenome which makes tissue-level profiles challenging in their interpretation. Further, there are over 40 different epigenomic features, and no single feature is sufficient to establish a complete map of chromatin or to predict RNA expression from a given locus. These and other challenges are beginning to be appreciated and addressed in epigenome-wide association studies, and AD investigations have been at the forefront of such studies, producing the first independently replicated associations of an epigenomic mark with a disease. Here, we review the state of the field in AD epigenomics and outline a path forward to explore this key element of genomic variation that is already uncovering new biology and complements the known genetic architecture of the disease.

The brain's epigenome

Epigenetic mechanisms can alter gene expression without changing the DNA sequence and are essential for cell differentiation [9], aging [7], pathogenesis of various diseases [52,31], and in the brain for memory formation and learning [24]. Since epigenetic marks are reversible, epigenetic mechanisms have long been a target for clinical intervention [1]. The most intensively studied epigenetic mechanism is DNA methylation. Methylated cytosines usually occur in the CG context, but in contrast to most other mature cells, non-CG context methylation is frequently found in neurons [37]. Further, hydroxymethylation of CG (which occurs as an intermediate in demethylation) is a stable feature in certain chromosomal regions and is enriched in the neuronal DNA methylome [32]. We are just beginning to appreciate the complexity of the methylome, and its relation to gene expression, as active transcription is typically associated with segments enriched for unmethylated CG (CG islands) near the transcription start site but with CG methylation in the gene body and the 3' end of the gene [30]. Thus, one cannot simply average methylation measures over a gene: one must consider the context of each CG when interpreting its level of methylation.

A second class of epigenetic marks is summarized by the term "histone modifications" which refers to post-translational modifications such as acetylation, methylation or phosphorylation of amino acids in these proteins that are an important structural element of chromatin. Modifications occur at numerous different positions of the histone amino acid sequence, and they are found in complex combinations that are associated with different states of chromatin [2,55]. No one mark fully predicts RNA expression from a given locus; histone mechanisms have recently been well reviewed elsewhere [34]. Finally, non-coding

Acta Neuropathol. Author manuscript; available in PMC 2017 October 01.

One interesting observation in a study of frontal cortex samples from a large number of older individuals (n=708) [De Jager 2014] is that most of the epigenome sampled by the Illumina Humanmethylation450 platform is very similar (>98%) across individuals. This may be an overestimate because of the limited sampling of this beadset and of the fact that many CG dinucleotides are either fully methylated or unmethylated in these subjects. Nonetheless, it is clear that the epigenome of the frontal cortex is quite similar between two individuals. Further, while there is significant concern that loss of neurons with age or pathology could skew epigenomic profiles, we actually see very modest or non-significant correlation between the estimated proportion of neurons in cortical samples [23] and measures of AD pathology. Neuronal loss clearly does occur with AD pathology, and this question probably deserves to be re-examined as algorithms for predicting the proportion of neurons and of different subsets of neurons improve. Thus, for now, the possible role of this measure (neuronal proportion) may be best left to pre-specified secondary analyses in epigenome-wide studies.

The effect of aging on the epigenome

Some of the inter-individual variation in epigenomic profiles will come from the developmental stage at which an individual is sampled or the extent to which they have experienced advancing age: several studies of various tissues and species have reported age-associated changes in all layers of the regulatory landscape including transcription factor binding, histone marks, DNA methylation and ncRNA levels. The cellular aging process is characterized by a loss of heterochromatin, as demonstrated by DNA hypomethylation, a loss of the heterochromatin mark H3K9me3 and nuclear lamina dysfunction [12,54,46]. Age-associated epigenetic changes were also observed in euchromatic regions including changes of the active mark H3K4me3, the repressive mark H3K27me3 and the transcription elongation mark H3K36me3 [10,39,43]. These changes likely contribute to increased transcriptional activity with age, which ultimately leads to progressive functional decline [48].

For the human brain, most aging studies have focused on DNA methylation. In line with observations from other tissues or species, surveys of human prefrontal cortex samples throughout the lifespan have uncovered large differences over time both during the human brain's development in children and with aging [42]. Strong associations at certain loci with aging may be relevant to AD, given that its incidence is age-related; however, it also creates an important confounder that has to be accounted for in AD studies. Work performed across many different tissues has suggested the existence of an epigenetic clock that captures a "biological or epigenetic age" of the target tissue that may reflect the accumulated effects of different exposures and may be different from a subject's "chronological age" (Figure 3a) [28,26]. The two measures are strongly correlated, but many subjects deviate from the regression line, appearing to be either "older" or "younger" than an average person at their chronological age (Figure 3b). Application of the algorithm that uses DNA methylation data to estimate an "epigenetic age" of a brain tissue sample from an individual has yielded an observation that its difference from an individual's chronological age correlates with AD related traits such as aging-related cognitive decline [Levine 2015]. However, the original algorithm does not account for the effect of pathologies which accumulate with age and can

confound what loci are involved with aging, independent of pathophysiological processes. Further, the model is built across a broad age spectrum, with limited sampling of the older end of the spectrum where epigenomic changes may no longer be linear and may require a model that reflects alterations in the rate of change of epigenomic marks. As seen in Figure 3b, in a sample of subjects who died at an advanced age, the inferred epigenetic age using an algorithm derived across tissues is correlated to chronological age but systematically underestimates the age of the subjects. Another factor may be the presence of regional differences in the accumulation of methylation changes since one study suggests that different brain regions may have different “epigenetic ages” [29]. However, such changes may be attributable to differences in the accumulation of neuropathologies among brain regions.

The impact of various factors that can confound the identification of age-related effects was nicely illustrated in a recent paper [51]: it demonstrated that about one-half of loci in which DNA methylation is associated with age in a naïve model can be accounted for by the effects of demographic features such as sex, different AD-related pathologies, and estimates of cell-type proportion. Thus, while it is clear that age leads to large changes in brain methylation at many sites throughout the genome, the exact etiology of these changes remains unclear, and it is possible that they represent the summation of changes attributable to multiple different pathologies, including those that lead to AD. In Figure 3b, we illustrate the large effect of β -amyloid and tau pathologies on the aging associations, after sex is taken into account: nearly 40% of the aging associations are attributable to these neuropathologic measures.

Since much of AD pathophysiology increases with age during the asymptomatic phase of the disease and there is significant discrepancy between an antemortem, syndromic diagnosis of probable AD and a pathologic diagnosis of AD [3], epigenomic studies of AD are very vulnerable to the outcome measures and subject populations that are selected at the time of a study’s design. In this, studies of the epigenome are more similar to studies of the transcriptome than to genetic studies, although all AD studies are vulnerable to this issue.

First generation of epigenomic studies in AD

While a number of studies have either evaluated an epigenomic mark in a small targeted region of the genome surrounding a candidate gene or captured a global measure of the presence of a histone modification in a sample of brain tissue, their results are difficult to interpret because they have been performed in small numbers of subjects, do not account for possible confounders and often do not include replication of their results in an independent sample. Thus, as with targeted candidate-gene based studies in genetics, these modestly sized studies have often not yielded robust results. These studies have been comprehensively catalogued recently and are not discussed further [35,5]. One example that illustrates this challenge to interpretation of many current studies is the first microRNA study of a larger AD cohort which found 41 deregulated microRNA in the prefrontal cortex and 35 in the hippocampus. However, the authors stated that there is little overlap with results from previously published smaller studies [36]. While different brain regions and technical confounders could be a reason, sample size and lack of appropriate accounting of potential cofounders are the most likely causes for a lack of replication. Nonetheless, associations of a

handful of miRNAs, such as miR-132 [47], with AD have been replicated in several independent studies, implicating such ncRNA in the disease. These results suggest that more comprehensive evaluations and dissections of miRNA and probably other ncRNA should be conducted in large sample sizes since we have now reasonable evidence for at least some miRNA being involved in AD.

A different aspect of the epigenome has now been investigated by two scans of the brain's DNA methylome. Performed independently on cortical samples from different patient populations and using different but related neuropathologic outcome measures, they nonetheless replicate each other's results at loci such as *ANKK1* [14, 41]. These two papers are important to the field of AD epigenomic studies because (1) they establish the fact that robust, reproducible associations between DNA methylation and AD-related outcomes exist and (2) they document a study design and sample size that led to a successful gene discovery study. This information can now be used as a template for future studies. Such replication remains rare in large-scale epigenomic studies of any disease.

A few important points have emerged from these studies: first, the effect size for a given methylated CG dinucleotide's association with an AD-related outcome is modest but is greater than that of validated AD SNPs [14,50]. This means that, while future studies need to be substantial in terms of sample size, they do not need to be on the scale of genome-wide association studies [14]. As with genetic studies, effect sizes will vary by trait, and we see smaller effects in relation to syndromic diagnoses such as AD dementia than in quantitative measures of AD neuropathology, and this will impact power calculations that inform study designs. Second, unlike what has been reported by other studies such as in rheumatoid arthritis [38], it appears that the detected epigenomic associations to AD are not driven by genetic variation. Genetic variation that affect the level of methylation are extremely common throughout the genome [21] and have large effects on methylation levels; however, none of the validated AD epigenomic associations show evidence of being caused by genetic associations [14]. In fact, when genetic and epigenomic associations have been documented in the same locus, they appear to be independent of one another [14]: for example, when the genetic and methylation data from the *BIN1* locus are modeled jointly in a single analytic framework, the association of the CG dinucleotide is not explained by the susceptibility SNP. The SNP and CG are found near each other within the gene (Figure 1), but each of them are influencing disease risk separately. Thus, a picture of converging influences – some genetic and some epigenomic – on certain key loci for AD susceptibility is emerging. Interestingly, a more complex picture emerges in the *CR1* locus, where it appears that an epigenomic feature may have an opposite effect on AD, depending on whether the AD susceptibility allele at this locus is present [11]. In *CR1*, the risk allele is associated with a duplication of a subset of exons that leads the CR1 protein to being structurally different and could explain why the same CG dinucleotide could have opposite effects on disease risk in the presence or absence of the risk allele. Finally, the quantum of DNA methylation variation remains to be better delineated. Unlike SNPs that identify a specific point in the genome as being involved in a disease, DNA methylation and chromatin marks identify a chromosomal segment as being associated. Further, as with SNPs that form the haplotypes that are the effective unit of genetic variation, there is an extensive correlation structure among methylation levels (or chromatin marks) both locally (in *cis*) and distally (in *trans*), such that

we do not yet fully understand how best to evaluate the number of independent tests being performed in an epigenome-wide study. In this, epigenomic data are more similar to transcriptomic data.

One can also take a less stringent threshold of significance and attempt to infer broader patterns in the disease-associated loci: two studies have done this and come up with different interpretations. In one case where β -amyloid plaques were the primary outcome measure [14], we see skew toward greater methylation among the associated sites and a strong enrichment for associated CG dinucleotides to be located in weak enhancers and weakly transcribed regions. There is also a strong depletion of associations in actively transcribed transcription start sites. However, in another smaller study, comparing AD subjects to non-AD subjects, associated sites were more likely to increase in methylation with AD but poised promoter regions appeared to be enriched for associated CG dinucleotides measured using the same Illumina platform [50]. These examples of possible broader patterns in epigenomic changes, evaluated in different cortical regions, therefore need further characterization before a clearer picture of broader patterns can emerge. Nonetheless, it is clear that AD-related epigenomic change is not random and is not a simple catastrophic decay of the epigenome in which we would not expect reproducible associations in specific loci.

In the DNA methylome-wide study for association with the burden of neuritic amyloid plaques [14], associated loci were found throughout the genome, and some of these associations were attenuated when estimates of cell type proportion were considered in secondary analyses. For example, the *RHBDF2* association is attenuated when an estimate of microglial cell proportion is added to the model, suggesting that the association may be driven, in part, by microglia. Because these models of cell type proportion are imperfect and can correlate with cell functions such as activation, results of these secondary analyses have to be interpreted cautiously: at the very least, this analysis tells us that the proportion of microglia is changing in the target tissue in relation to amyloid pathology, consistent with many earlier studies [15], and it provides a quantitative estimate of that presence that is not available from the analysis of tissue sections. In addition, it prioritizes a particular gene, *RHBDF2*, for future investigation within microglia, while understanding that it is possible that its methylation level is simply a proxy for the proportion of microglia. In that case, *RHBDF2* becomes a very useful tool for future studies. In the study of neuritic plaque pathology, the different validated associations appear to be distributed across multiple different cell types, consistent with our understanding of neuritic plaque pathology accumulating extracellularly and affecting different CNS cells. The same is not true of the other neuropathologic feature of AD, tau pathology: here, it appears that there are large-scale alterations in the state of DNA methylation in the presence of tau pathology. Unlike the changes noted with neuritic plaques [14] that are more distributed, tau associations derived from profiles of a histone modification that marks transcriptionally active chromatin (H3K9 acetylation) in the same collection of brains are clustered in certain segments of the genome (unpublished data). These results are consistent with an observation of widespread relaxation of heterochromatin in a *Drosophila* model of Tau pathology that was confirmed to be present in a small set of human AD brain samples [20].

Epigenomic profiling of blood in AD

Given the availability of blood-derived DNA from large numbers of subjects characterized for AD, there is a lot of interest in profiling the epigenome of peripheral blood to identify susceptibility loci. The challenge is that there is, as yet, little data on the utility of this approach. While it is reasonable to expect that some life experiences or exposures could affect the epigenome of the brain and of the peripheral blood in the same manner, most probably will not. A recent study in a modest number of samples reported that paired brain and blood profiles from the same subjects are largely not correlated [25]. Another small study is consistent with this observation: it reports little correlation between cortical methylation profiles and CD4+ T cell methylation profiles from the same individuals [53]. Further, while the small brain data set demonstrated known AD associations, the CD4+ T cell data did not. There are many caveats to these experiments, including the selection of cells to profile in peripheral blood, as monocytes that infiltrate the brain in AD could be more pertinent than lymphocytes. However, it is unlikely that there is a strong, direct correlation between the epigenomes of the brain and blood such that measuring a given epigenomic feature in blood offers a reasonable proxy for the same feature in the brain for most genomic loci. Sampling certain immune cell populations that infiltrate the brain from peripheral blood may be meaningful, but, given the small proportion of such cells in cortical tissue, it will not be useful as a proxy for tissue profiles. Peripheral blood data may be more useful in identifying biomarkers that reflect the state of the subject but do not provide a direct evaluation of the brain.

Designing a statistically powerful epigenomic study in human subjects

In the sections above, we reviewed the state of the field of AD epigenomics in human subjects; the discovery of independently validated AD loci such as *ANKK1* and the study of the properties of DNA methylation profiles have given us important information with which to design a second generation of epigenomic studies. What do such studies look like given our current findings? The first consideration is the selection of the target tissue or cell type. Given the limited sampling of tissues profiled to date, human cortex is preferred since we have validated observations that give us information with which to conduct power calculations; blood may be useful but remains questionable as a substrate for epigenomic studies of AD today. Given the use of post-mortem cortical tissue, accessing samples available from archives of frozen tissue is preferred as fixed tissue is more difficult to profile using current techniques [4].

Many different epigenomic marks could be profiled and will offer complementary perspectives of the state of chromatin in the tissue sample: no one epigenomic mark captures complete information on chromatin conformation. Currently, DNA methylation profiling is available from experimental protocols that are most easily deployed in moderate to large scale, although moderate scale chromatin immunoprecipitation with sequencing (ChIP-seq) is now feasible as well. Each platform produces a different form of data, and thus the analytic approach must take into account fundamental differences in the nature of the data, such as whether the epigenomic mark is measured at a single point (as with CG methylation using an Illumina beadset) or over a region, as is the case with ChIP-seq methods. The

nature of the data will therefore yield important differences in how the targeted epigenomic feature is linked to a physical position on the chromosome. It is important to emphasize that the quantum of variation in the epigenome is not the single nucleotide as with SNPs but rather a chromosomal region of variable size, depending on the chromatin mark being measured. Another issue is that many platforms do not sample the entire human genome: for example, the Illumina HumanMethylation450 beadset interrogates a large number of CG dinucleotides throughout the genome but only samples ~5% of all potentially methylated CG dinucleotides [13]. This is well illustrated by Figure 1 which presents the results of this array for the *BIN1* locus: the subset of targeted CGs is unevenly distributed and enriched in certain features of known genes, making generalizations from the available data difficult. Finally, these methods also do not distinguish another layer of complexity such as modification of methylation into 5-hydroxymethylation, which appears to be particularly prominent in the nervous system [32].

For the validated, CG-specific association that exists to date, we have reported that the methylation level of a given CG dinucleotide explains, on average, 5.0% of the variance in neuritic amyloid pathology burden [14]. This is much larger than the less than 1% explained by known AD susceptibility variants for the same trait. This means that studies with ~500 subjects will be well powered to discover novel associations with quantitative measures of AD pathology. The effect size for a syndromic diagnosis of AD is more modest and is therefore less attractive in the primary analysis since more subjects will be needed to achieve an equivalent level of statistical power, as demonstrated in Figure 3.

Addressing causality – meaning of epigenomic associations

As with putative risk factors evaluated in epidemiologic studies, an association between an epigenomic feature and a disease-related outcome in a cross-sectional study cannot address the issue of causality [6]. That is, the feature could be associated with the outcome because either (1) it contributes to the process that leads to the outcome or (2) it is one of the events that is a consequence of the outcome's occurrence. This ambiguity is important to realize, and, once an epigenomic association is robustly validated by independent studies, it clearly needs to be explored further to resolve its place in the sequence of disease-related events. Studies of human brains are generally limited to a cross-sectional design because samples are primarily accessed at the time of autopsy. This leaves investigators with the challenging task of being creative in designing *in vitro* or *in vivo* studies in model systems to explore the issue of causality. New emerging epigenetic editing tools might be extremely helpful for this task [27]. On the other hand, studies of other tissues such as blood offer the opportunity to investigate human subjects longitudinally, enabling researchers to directly explore the issue of causality in the target human tissue. In this case, change in epigenomic profiles over time can identify loci where chromatin changes contribute to the onset of disease.

It is important to remember that genome-wide association studies (GWAS) have the same cross-sectional design and the same issue that causality is not formally demonstrated by the association study. One makes the reasonable assumption that the vast majority of sampled genetic variation were present in the organism at the time of conception, and one therefore assumes that the genetic variant has a causal role in a disease such as AD. However, this

assumption breaks down as we begin to investigate (1) rare variations that could represent an early somatic mutation event in the tissue being sampled (i.e. blood) that is not representative of the organism or (2) regions of the genome where rearrangements and somatic mutation occur (i.e. immunoglobulin and T cell receptor loci). It is also likely that other, as yet undescribed, regions of somatic alteration occur in the human genome. Thus, we have to keep in mind that not all genetic variation measured in blood or other tissues is completely representative of an individual's genetic complement, although this assumption will hold in the vast majority of common and other variants sampled in GWAS.

Regardless of the relative location of the epigenomic event in relation to the outcome of interest, the association is meaningful as it provides important information with which to either prevent an outcome or to mitigate its occurrence. So, an associated epigenomic event is interesting either way, but its interpretation and the design of follow-on studies will depend on its temporal relationship to the outcome. Requiring a definitive interpretation of causality in the initial report of an epigenomic discovery study of human autopsy tissue is an unrealistic threshold: as with genetic studies, an appropriately sized study with replication in independent subjects is the goal for an initial epigenome-wide association study. Well-designed follow-up studies represent a substantial investment of time and resources, and their interpretation will only be as good as the relevance of the model system that is being used.

Nonetheless, certain secondary analyses can be used to provide an initial assessment of causality that can guide further experiments: for example, the effect of the reported epigenomically-associated AD loci such as *ANKK1* is present in subjects prior to the clinical manifestation of AD, suggesting that these changes are not part of the terminal, dementing process [14]. Further, the fact that two AD susceptibility loci – *ABCA7* and *BIN1* – harbor both genetic and epigenomic risk factors supports the possibility that at least some of the epigenomically-associated loci may be involved in AD susceptibility, as is also suggested by a protein:protein interaction network analysis that links up genetically and epigenomically-defined susceptibility genes [14].

Future direction – should we study the epigenome in AD?

Despite their limitations, the current set of published and emerging studies document a number of loci where changes in epigenomic features are reproducibly occurring in the cortex of older individuals who have accumulated AD-related pathology. Thus, we now know that measuring the human brain's epigenome in AD is meaningful, and we have a model with which to design larger studies that will identify additional loci where chromatin changes are associated with AD-related traits. The first generation of studies has identified a small set of such AD associated loci, but we do not yet understand either the mechanism by which they occur or their exact nature as we have only catalogued discrete changes in methylation at each site. Fine epigenetic mapping of each of these sites with multiple different epigenomic marks and perhaps certain transcription factors that are implicated by the associated regions are needed to explore the nature of these epigenomic associations. In particular, profiling the same tissue samples for 5-hydroxymethylation should prove quite revealing in assessing whether this DNA methylation mark that is enriched in neurons may

have a greater role in AD than might be suggested by its presence on only a minority of potentially methylated sites. Further, new techniques such as profiles that use the Assay for Transposase Accessible Chromatin (ATAC) will yield higher-resolution profiles of chromatin that are not specific to a given mark that is correlated with a chromatin state but rather interrogate the structure of chromatin directly [8]. These and other methods need to be deployed in large scale on the same samples so that we can understand (1) the extent of interindividual variation in the epigenome and (2) the most informative layer(s) of “omic” data that should be obtained to capture the majority of the information relating to the state of chromatin. No single mark will be sufficient: we need to accumulate profiles using multiple non-redundant marks and integrate the results into a single framework. Then, measures of ncRNAs need to be layered on to DNA and chromatin measures to further enhance the overall perspective on patterns of gene regulation in AD since reproducible associations of miRNA with AD have emerged.

The issue of causality should be pursued, although both *in vitro* systems such as human induced pluripotent stem cells differentiated to neurons [49] and *in vivo* murine or Drosophila models have limitations. In one example, an epigenomic profile of the CK-p25 mouse model, seems to have primarily indicated the presence of a large number of infiltrating myeloid cells, which may not be representative of other AD models [22] but may capture aspects of the human disease. Overall, a better approach is needed to select models in which to mechanistically dissect epigenomic associations: one needs to empirically identify mouse and other model systems which capture elements of the biology being investigated. A comparison of mouse and human epigenomic changes could be one way to select a mouse model that is optimal to follow up a human observation, and, in fact, it is likely to define a population of such models which will have to be considered simultaneously to capture the heterogeneity of human disease.

For practical reasons, current studies have focused on tissue-level profiles that are confounded by issues such as variation in the frequency of different cell populations in the tissue. Models with which to estimate the proportion of neurons in a cortical sample [23]), while useful for an initial evaluation, remain relatively crude as they may over-correct the data if used in the generation of the epigenomic measure being analyzed. Miniaturization of epigenomic assays is now allowing the design of cell-type specific profiles of tissues, and single cell approaches are beginning to allow us to characterize the complexity of cell states among a single cell type. These studies that are more time-consuming are reasonable given the fact that robust epigenomic associations exist at the tissue level, and they will lead to the next level of insight in the role of the epigenome in AD.

The plasticity of the epigenome creates challenges in deconstructing the observed epigenomic associations at the tissue level. However, they also highlight one of the main attractions of studying the epigenome: it may reflect the longstanding effect of certain environmental exposures or life experiences on the function of an organism. These exposures may be difficult to measure or may have occurred many years prior to the onset of any study, making it very difficult to understand their role without a concrete measure of exposure in the tissue of an individual with which to anchor a study. GWAS have explained <50% of the variance in most studied human traits: we therefore need another approach to

identify complementary genomic changes that alter gene expression to cause diseases such as AD. Transcriptomic studies will address the same issue but may be more vulnerable to agonal events and reflect just what the target tissue was expressing at the time of death. The epigenome captures information not just about actively transcribed genes but also about genes that may be in a conformation where they have the potential to be expressed in response to a given stimulation. The epigenome thus offers information about both the actively transcribed and potentially transcribed genes. Nonetheless, because no single mark fully captures the state of the epigenome, the best studies will measure both the epigenome and the transcriptome of a tissue or set of purified cells from genetically characterized individuals, allowing for an integrated analysis of multiple sources of variation in the function of the brain in AD.

Will therapies that modulate the epigenome be effective in AD? The current set of studies have been useful to identify specific loci in which DNA methylation is related to AD and could become the target of therapeutic intervention. However, it is too early to pinpoint a single overall mechanism of epigenomic dysfunction that yields the robust changes that are being reported. Rather, given the heterogeneity of the disease, it is likely that a number of different mechanisms will be implicated and that no single compound will reverse all AD-related epigenomic changes. Nonetheless, compounds that modulate the epigenome could well be part of future treatment paradigms; for example, histone deacetylase inhibitors have been studied and shown empirically to have effects on certain mouse models [19]. However, these and other compounds need to be re-examined carefully since we now have clear genomic regions whose epigenetic alterations need to be corrected instead of looking for broad changes in acetylation levels that affect a mouse model. Optimism for such approaches is warranted as the epigenomic changes that have been identified are accumulating over the lifespan: they are not present at birth. Thus, given the plasticity of the epigenome, they may be reversed, or, at least, we may be able to develop strategies to prevent the changes from occurring by blocking the epigenomic machinery when the risk-associated exposure or life experience occurs during a patient's life.

Overall, the field of epigenomic studies in AD is just emerging, but it is clear that it will play an important role in characterizing the set of events that lead from genetic susceptibility variants and environmental risk factors to intermediate traits such as the accumulation of amyloid and tau pathologies and eventually to a syndromic diagnosis of AD. Plastic by nature, the epigenome presents a compelling target for drug development; however, it is unlikely that a single mechanism – whether related to methylation, histones, or ncRNAs – will provide a comprehensive therapeutic solution. Nonetheless, epigenomic modulators could well become important tools in the portfolio of AD specialists, once the right population and disease stage for intervention have been identified.

References

1. Adwan L, Zawia NH. Epigenetics: a novel therapeutic approach for the treatment of Alzheimer's disease. *Pharmacol Ther.* 2013; 139:41–50. [PubMed: 23562602]
2. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res.* 2011; 21:381–395. [PubMed: 21321607]

3. Bennett DA, Wilson RS, Boyle PA, Buchman AS, Schneider JA. Relation of neuropathology to cognition in persons without cognitive impairment. *Ann Neurol.* 2012; 72:599–609. [PubMed: 23109154]
4. Bennett DA, Wilson RS, Arvanitakis Z, Boyle PA, de Toledo-Morrell L, Schneider JA. Selected findings from the Religious Orders Study and Rush Memory and Aging Project. *J Alzheimers Dis.* 2013; 33(Suppl 1):S397–403. [PubMed: 22647261]
5. Bennett DA, Yu L2, Yang J2, Srivastava GP3, Aubin C3, De Jager PL3. Epigenomics of Alzheimer's disease. *Transl Res.* 2015; 165:200–20. [PubMed: 24905038]
6. Birney E, Smith GD, Greally JM. Epigenome-wide association studies and the interpretation of disease -omics. *PLoS Genet.* 2016; 12:e1006105. [PubMed: 27336614]
7. Booth LN, Brunet A. The aging epigenome. *Mol Cell.* 2016; 62:728–744. [PubMed: 27259204]
8. Buenrostro, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods.* 2013; 10:1212–1218.
9. Cantone I, Fisher AG. Epigenetic programming and reprogramming during development. *Nat Struct Mol Biol.* 2013; 20:282–289. [PubMed: 23463313]
10. Cheung I, Shulha HP, Jiang Y, Matevossian A, Wang J, Weng Z, Akbarian S. Developmental regulation and individual differences of neuronal H3K4me3 epigenomes in the prefrontal cortex. *Proc Natl Acad Sci U S A.* 2010; 107:8824–8829. [PubMed: 20421462]
11. Chibnik LB, Yu L, Eaton ML, Srivastava G, Schneider JA, Kellis M, Bennett DA, De Jager PL. Alzheimer's loci: epigenetic associations and interaction with genetic factors. *Ann Clin Transl Neurol.* 2015; 2:636–647. [PubMed: 26125039]
12. Cruickshanks HA, McBryan T, Nelson DM, Vanderkraats ND, Shah PP, van Tuyn J, Singh Rai T, Brock C, Donahue G, Dunican DS, Drotar ME, Meehan RR, Edwards JR, Berger SL, Adams PD. Senescent cells harbour features of the cancer epigenome. *Nat Cell Biol.* 2013; 15:1495–1506. [PubMed: 24270890]
13. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. *Epigenomics.* 2011; 3:771–784. [PubMed: 22126295]
14. De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, Eaton ML, Keenan BT, Ernst J, McCabe C, Tang A, Raj T, Replogle J, Brodeur W, Gabriel S, Chai HS, Younkin C, Younkin SG, Zou F, Szyf M, Epstein CB, Schneider JA, Bernstein BE, Meissner A, Ertekin-Taner N, Chibnik LB, Kellis M, Mill J, Bennett DA. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat Neurosci.* 2014; 17:1156–1163. [PubMed: 25129075]
15. De Strooper B, Karran E. The Cellular Phase of Alzheimer's Disease. *Cell.* 2016; 164:603–15. [PubMed: 26871627]
16. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012; 489:57–74. [PubMed: 22955616]
17. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods.* 2012; 9:215–216. [PubMed: 22373907]
18. Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shores N, Whitton H, Ryan RJ, Shishkin AA, Hatan M, Carrasco-Alfonso MJ, Mayer D, Luckey CJ, Patsopoulos NA, De Jager PL, Kuchroo VK, Epstein CB, Daly MJ, Hafler DA, Bernstein BE. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature.* 2015; 518:337–343. [PubMed: 25363779]
19. Foscher A. Targeting histone-modifications in Alzheimer's disease. What is the evidence that this is a promising therapeutic avenue? *Neuropharmacology.* 2014; 80:95–102.
20. Frost B, Hemberg M, Lewis J, Feany MB. Tau promotes neurodegeneration through global chromatin relaxation. *Nat Neurosci.* 2014; 17:357–66. [PubMed: 24464041]
21. Gibbs JR, an der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL, Arepalli S, Dillman A, Rafferty IP, Troncoso J, Johnson R, Zielke HR, Ferrucci L, Longo DL, Cookson MR, Singleton AB. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* 2010; 6:e1000952. [PubMed: 20485568]

22. Gjonjeska E, Pfenning AR, Mathys H, Quon G, Kundaje A, Tsai LH, Kellis M. Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature*. 2015; 518:365–9. [PubMed: 25693568]
23. Guintivano J, Aryee MJ, Kaminsky ZA. A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics*. 2013; 8:290–302. [PubMed: 23426267]
24. Guzman-Karlsson MC, Meadows JP, Gavin CF, Hablitz JJ, Sweatt JD. Transcriptional and epigenetic regulation of Hebbian and non-Hebbian plasticity. *Neuropharmacology*. 2014; 80:3–17. [PubMed: 24418102]
25. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*. 2015; 10:1024–32. [PubMed: 26457534]
26. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sada S, Klotzle B, Bibikova M, Fan JB, Gao Y, Deconde R, Chen M, Rajapakse I, Friend S, Ideker T, Zhang K. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*. 2013; 49:359–367. [PubMed: 23177740]
27. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, Gersbach CA. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol*. 2015; 33:510–517. [PubMed: 25849900]
28. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013; 14:R115. [PubMed: 24138928]
29. Horvath S. The cerebellum ages slowly according to the epigenetic clock. *Aging*. 2015; 7:294–306. [PubMed: 26000617]
30. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012; 13:484–492. [PubMed: 22641018]
31. Klein HU, De Jager PL. Uncovering the role of the methylome in dementia and neurodegeneration. *Trends Mol Med*. 2016 ePub ahead of print.
32. Kriaucionis S, Heintz N. The nuclear DNA base, 5-hydroxymethylcytosine is present in brain and enriched in Purkinje neurons. *Science*. 2009; 324:929–930. [PubMed: 19372393]
33. Lambert JC, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 2013; 45:1452–1458. [PubMed: 24162737]
34. Lawrence M, Daujat S, Schneider R. Lateral thinking: How histone modifications regulate gene expression. *Trends Genet*. 2016; 32:42–56. [PubMed: 26704082]
35. Lardenoije R, Iatrou A, Kenis G, Kompotis K, Steinbusch HW, Mastroeni D, Coleman P, Lemere CA, Hof PR, van den Hove DL, Rutten BP. The epigenetics of aging and neurodegeneration. *Prog Neurobiol*. 2015; 131:21–64. [PubMed: 26072273]
36. Lau P, Bossers K, Janky R, Salta E, Frigerio CS, Barbash S, Rothman R, Sierksma AS, Thathiah A, Greenberg D, Papadopoulou AS, Achsel T, Ayoubi T, Soreq H, Verhaagen J, Swaab DF, Aerts S, De Strooper B. Alteration of the microRNA network during the progression of Alzheimer's disease. *EMBO Mol Med*. 2013; 5:1613–1634. [PubMed: 24014289]
37. Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, Yu M, Tonti-Filippini J, Heyn H, Hu S, Wu JC, Rao A, Esteller M, He C, Haghighi FG, Sejnowski TJ, Behrens MM, Ecker JR. Global epigenomic reconfiguration during mammalian brain development. *Science*. 2013; 341:1237905. [PubMed: 23828890]
38. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, Reinius L, Acevedo N, Taub M, Ronninger M, Shchetynsky K, Scheynius A, Kere J, Alfredsson L, Klareskog L, Ekström TJ, Feinberg AP. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat Biotechnol*. 2013; 31(2):142–147. [PubMed: 23334450]
39. Liu L, Cheung TH, Charville GW, Hurgu BM, Leavitt T, Shih J, Brunet A, Rando TA. Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. *Cell Rep*. 2013; 4:189–204. [PubMed: 23810552]
40. Lorio MV, Piovan C, Croce CM. Interplay between microRNAs and the epigenetic machinery: an intricate network. *Biochim Biophys Acta*. 2010; 1799:694–701. [PubMed: 20493980]

41. Lunnon K, Smith R, Hannon E, De Jager PL, Srivastava G, Volta M, Troakes C, Al-Sarraj S, Burrage J, Macdonald R, Condliffe D, Harries LW, Katsel P, Haroutunian V, Kaminsky Z, Joachim C, Powell J, Lovestone S, Bennett DA, Schalkwyk LC, Mill J. Methyloomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat Neurosci.* 2014; 17:1164–1170. [PubMed: 25129077]
42. Numata S, Ye T, Hyde TM, Guitart-Navarro X, Tao R, Wininger M, Colantuoni C, Weinberger DR, Kleinman JE, Lipska BK. DNA methylation signatures in development and aging of the human prefrontal cortex. *Am J Hum Genet.* 2012; 90:260–272. [PubMed: 22305529]
43. Pu M, Ni Z, Wang M, Wang X, Wood JG, Helfand SL, Yu H, Lee SS. Trimethylation of Lys36 on H3 restricts gene expression change during aging and impacts life span. *Genes Dev.* 2015; 29:718–731. [PubMed: 25838541]
44. Roadmap Epigenomics Consortium et al. Integrative analysis of 111 reference human epigenomes. *Nature.* 2015; 518:317–330. [PubMed: 25693563]
45. Satterlee JS, Beckel-Mitchener A, McAllister K, Procaccini DC, Rutter JL, Tyson FL, Chadwick LH. Community resources and technologies developed through the NIH Roadmap Epigenomics Program. *Methods Mol Biol.* 2015; 1238:27–49. [PubMed: 25421653]
46. Salta E, Sierksma A1, Vanden Eynden E1, De Strooper B2. miR-132 loss de-represses ITPKB and aggravates amyloid and Tau pathology in Alzheimer's brain. *EMBO Molec Med.* 2016:e201606520.
47. Scaffidi P, Misteli T. Lamin A-dependent nuclear defects in human aging. *Science.* 2006; 312:1059–1063. [PubMed: 16645051]
48. Southworth LK, Owen AB, Kim SK. Aging mice show a decreasing correlation of gene expression within genetic modules. *PLoS Genet.* 2009; 5:e1000776. [PubMed: 20019809]
49. Sullivan S, Young-Pearse TL. Induced pluripotent stem cells as a discovery tool for Alzheimer's disease. *Brain Res.* 2015 (online).
50. Watson CT, Roussos P1,2,3, Garg P1, Ho DJ1, Azam N1, Katsel PL2,4, Haroutunian V2,3,4, Sharp AJ5. Genome-wide DNA methylation profiling in the superior temporal gyrus reveals epigenetic signatures associated with Alzheimer's disease. *Genome Med.* 2016; 8:5. [PubMed: 26803900]
51. Yang J, Yu L2, Gaiteri C2, Srivastava GP3, Chibnik LB3, Leurgans SE2, Schneider JA4, Meissner A5, De Jager PL3, Bennett DA2. Association of DNA methylation in the brain with age in older persons is confounded by common neuropathologies. *Int J Biochem Cell Biol.* 2015; 67:58–64. [PubMed: 26003740]
52. You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell.* 2012; 22:9–20. [PubMed: 22789535]
53. Yu L, Chibnik LB, Yang J, McCabe C, Xu J, Schneider JA, De Jager PL, Bennett DA. Methylation profiles in peripheral blood CD4+ lymphocytes versus brain: The relation to Alzheimer's disease pathology. *Alzheimers Dement.* 2016 (online).
54. Zhang W, Li J, Suzuki K, Qu J, Wang P, Zhou J, Liu X, Ren R, Xu X, Ocampo A, Yuan T, Yang J, Li Y, Shi L, Guan D, Pan H, Duan S, Ding Z, Li M, Yi F, Bai R, Wang Y, Chen C, Yang F, Li X, Wang Z, Aizawa E, Goebel A, Soligalla RD, Reddy P, Esteban CR, Tang F, Liu GH, Belmonte JC. Aging stem cells. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. *Science.* 2015; 348:1160–1163. [PubMed: 25931448]
55. Zhou VW, et al. Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet.* 2011; 12(1):7–18. [PubMed: 21116306]

Figure 1. Genetic and epigenomic variation have independent effects on disease susceptibility in the *BIN1* locus

At the top of the figure, we include a diagram of chromosome 2, with a vertical red line denoting the location of the *BIN1* locus. Below the chromosome, a double line presents the physical position of the segment of chromosome 2 that is presented in this figure (X axis for all subsequent components of the figure). The first ribbon contains the SNPs (blue dots) that have been interrogated in relation to AD susceptibility in this locus. The Y axis reports the $-\log_{10}(\text{p value})$ for the association from AD, taken from the IGAP study for each SNP [33]. The second ribbon presents CG dinucleotides (pink dots) interrogated in relation to AD susceptibility as part of a DNA methylation study of the frontal cortex in AD [DeJager 2014]. The peak of the genetic and epigenomic associations are close to one another but do not overlap exactly. When modeled together, the SNP does not explain the association of the CG methylation level with AD, as reported in [14]. Below the association results, we present the chromatin state of this segment of chromosome 2, based on the Roadmap Epigenomics data and the ChromHMM algorithm [44] for three different brain regions: the prefrontal cortex (PFC), the hippocampus (HC), and the substantia nigra (SN). Each color denotes a different chromatin state, as outlined in the color key. The next ribbon presents the level of sequence conservation across mammalian species. The one below presents the location of CG islands, and the bottom ribbon shows the location of the exons of the gene found in this chromosomal segment, *BIN1*.

Figure 2. “Epigenetic age” based on DNA methylation and its relation to neuropathology

(a) The age of a subject can be predicted based on the subject’s epigenetic profile. The predicted age is termed “epigenetic age” and, for most tissues, is strongly correlated with the chronological age as indicated in the scatter plot. If the epigenetic age of a subject is larger than the chronological age (red dot), the subject’s epigenome appears to be “older” than the average epigenome of an individual with that chronological age in the reference set that was used to train the prediction model. Hence, the epigenetic age must be interpreted with respect to the reference set, which can comprise various different tissues and cell types or be a homogeneous set of brain tissues, for example. (b) A published method [28] was used to estimate the epigenetic age of each aging subject from whom we have a DNA methylation profile of the dorsolateral prefrontal cortex ($n=740$) [DeJager 2014]. The scatter plot shows each subject as a dot relative to their estimated epigenetic age based on DNA methylation data and chronological age. While the two ages are strongly correlated ($r=0.67$) (the regression line is presented in red), the estimated epigenetic age systematically underestimates the chronological age. This may be due to the age range and the various different tissue types from which the algorithm was developed. (c) The diagram illustrates the point made in a recent manuscript [51] that neuropathologies must be considered when studying aging-associated alterations of brain DNA methylation levels. The red circles indicates that 7,336 out of 420,132 CG dinucleotides were significantly associated with age in a statistical model adjusted for gender and technical covariates. When the model was

further adjusted for six common neuropathologies (green circle), 39% of those CG dinucleotides were no longer significantly associated with age.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. Required sample sizes for AD methylome-wide association studies

The plot shows the number of significantly differentially methylated CG dinucleotides detected for various sample sizes between 50 and 740 subjects. The required sample size depends remarkably on the studied variable as demonstrated by the different curves for neuritic plaque (NP, continuous pathologic variable), pathologic Alzheimer diagnosis (pathologic AD, dichotomous pathologic variable) and clinical Alzheimer diagnosis (clinical AD, dichotomous clinical variable). The dashed lines show results after a stringent multiple testing adjustment using the family-wise error rate (FWER); solid lines show results when using the more liberal false discovery rate (FDR) approach. Methylation levels were logit-transformed and all 420,132 CG dinucleotides for tested for an association with the variable of interest in a linear model adjusted for gender, age and technical covariates.