CellPress

Review

Epigenetics: It's Getting Old. Past Meets Future in Paleoepigenetics

David Gokhman,[1] Eran Meshorer,[1,2,*] and Liran Carmel[1,*]

Recent years have witnessed the rise of ancient DNA (aDNA) technology, allowing comparative genomics to be carried out at unprecedented time resolution. While it is relatively straightforward to use aDNA to identify recent genomic changes, it is much less clear how to utilize it to study changes in epigenetic regulation. Here we review recent works demonstrating that highly degraded aDNA still contains sufficient information to allow reconstruction of epigenetic signals, including DNA methylation and nucleosome positioning maps. We discuss challenges arising from the tissue specificity of epigenetics, and show how some of them might in fact turn into advantages. Finally, we introduce a method to infer methylation states in tissues that do not tend to be preserved over time.

Trends

How are ancient epigenomes reconstructed?

Which epigenetic layers could be reconstructed?

The pros and cons of tissue specificity in paleoepigenetics.

A novel method to infer methylation in unobtainable tissues.

Unearthing Epigenetic Layers

The epigenome is viewed today as a collection of regulatory layers that control when, where, and how genes are turned on and off. These layers are passed through cellular or organismal generations and include modifications to the DNA (i.e., DNA methylation) and to the proteins that package it (e.g., histone modifications), as well as regulation by noncoding RNAs (e.g., miRNAs) and changes in the 3D conformation of the genome. While it is still debated which layers are epigenetic and to what extent they are heritable (Box 1) [1–3], it is nevertheless accepted that alterations in regulatory layers can propel substantial phenotypic changes [4]. Such alterations can stem from sequence mutations, but also from environmental factors, or simply be a result of stochastic processes [5]. This combination of plasticity and heredity led to the growing recognition that epigenetic evolution occurs in short timescales, precedes sequence adaptation [6], and could underlie phenotypic differences between closely related species [6–11]. In light of this, studying recent adaptations of a species requires a comparison of epigenomes of close evolutionary relatives [12–14].

Unfortunately, very often, the extant sister group of a species is deeply diverged from it, allowing only crude resolution in determining the timing of evolutionary events. For example, the closest extant relatives of humans are the chimpanzee and the bonobo, from which we diverged ~5–8 million years ago [15–17]. As a result, it is usually impossible to determine whether an evolutionary change along our lineage happened recently and is unique to modern humans, or whether it occurred in our deep past, at times when our ancestors displayed many ancestral properties such as a brain the size of a chimpanzee's. However, exciting developments in the rising field of ancient DNA (aDNA) provide access to genomes of extinct species, and thus pave the way for much finer temporal analyses.

The Rise of Ancient Genomics

Recent years have witnessed the successful high-quality sequencing of two individuals from archaic human groups – a Denisovan at 30× coverage [18], and a Neanderthal at 52× coverage

[1]Department of Genetics, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel
[2]The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

*Correspondence: meshorer@huji.ac.il (E. Meshorer) and liran.carmel@huji.ac.il (L. Carmel).

 http://dx.doi.org/10.1016/j.tree.2016.01.010

Box 1. The Different Perceptions of Epigenetics

The term 'epigenetics' was coined by Conrad Waddington [84] to describe the interactions between genes and their products to produce phenotypes. This incredible insight was published in 1942, in the pre-gene era. Nowadays, the term 'epigenetics' is used in different ways. The narrower and more traditional definition is a 'stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence' [85]. The broader everyday definition refers to the complete set of regulatory layers that have the potential of being inherited, whether genetically driven or non-genetically driven. According to this definition, any change to DNA methylation, histone modifications, chromatin packaging, small RNA, etc. is considered part of the epigenome. In this review, we refer to epigenetics in its broader definition.

[19]. These works provided the full DNA sequence of these archaic humans, allowing in some cases to associate phenotypes with genetic differences [19,20]. However, as noted earlier, genetics alone cannot explain the full spectrum of phenotypic adaptations. These ancient genomes revealed that the number of fixed amino acid substitutions that separate present-day humans from archaic ones stands at only 96 (in 87 proteins) [19]. This relatively small number stems from the fact that we share a very recent common ancestor with these archaic humans, around 550 000–765 000 years ago [19]. Such recent divergence was too young for the accumulation of many amino acid changes, but sufficient to give rise to numerous noncoding sequence changes of potentially regulatory roles. However, our ability to predict the regulatory effect of a sequence change is very limited, and it is therefore necessary to develop ways to map epigenetic layers in aDNA. On first glimpse, it might look like an insurmountable endeavor; aDNA is broken and degraded, and thus inherently inactive and includes little to no remnants of cellular context. Moreover, while some extracellular proteins survive for periods of time exceeding those of DNA, most proteins rapidly break down, leaving little to no trace of their activity patterns in the premortem cell [21–23]. Nevertheless, recent works demonstrated that at least some of the epigenetic signals might be accurately reconstructed [24–26], thus providing information that could not have been gained from genetics alone [27,28]. In this review, we describe these recent developments and discuss future possibilities in this novel field, which we refer to as 'paleoepigenetics'.

Reconstructing Archaic DNA Methylation Maps

DNA methylation, the conversion of cytosine to 5′-methyl-cytosine by DNA methyltransferases, is a fundamental epigenetic mark, involved in the regulation of gene activity. In mammals, methylation usually occurs in the context of a cytosine followed by a guanine (CpG). We are still far from predicting the effect of a local change in methylation on the expression level of a gene. The strongest predictive power is in promoter regions, where hypermethylation is associated with gene silencing [29]. Several recent works demonstrated that premortem DNA methylation patterns can be reconstructed from aDNA sequences, either by direct measurement or by computational algorithms (Table 1). Direct measurement employs protocols that are regularly used in modern samples (e.g., bisulfite conversion or methyl-enrichment methods, followed by sequencing [30,31]). These methods can produce single nucleotide resolution maps, but their applicability to aDNA depends on several factors. Bisulfite sequencing (BS-seq) involves the conversion of unmethylated cytosines into uracils, followed by sequencing, and therefore requires special allocation of DNA, as this conversion does not allow the reuse of the same sample in the future (Table 1). This makes rare samples, or samples with minute quantities of DNA, not suitable for BS-seq. Enrichment-based methods, by contrast, are based on the precipitation of methylated cytosines, and thus generally do not modify the DNA. However, these methods are inherently biased towards CpG-rich regions and long fragments of DNA [30,32]. Finally, both methods will measure systematically skewed levels of methylation in samples that have gone through elevated levels of degradation, that is, the spontaneous hydrolytic deamination of methylated cytosines into thymines [30,32]. Altogether, these are the methods of choice for typically younger, better preserved and relatively abundant aDNA samples (Table 1).