Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

# Evolution of epigenetic chromatin states

Philip Yuk Kwong Yung and Simon J Elsässer

The central dogma of gene expression entails the flow of genetic information from DNA to RNA, then to protein. Decades of studies on epigenetics have characterized an additional layer of information, where epigenetic states help to shape differential utilization of genetic information. Orchestrating conditional gene expressions to elicit a defined phenotype and function, epigenetics states distinguish different cell types or maintain a long-lived memory of past signals. Packaging the genetic information in the nucleus of the eukaryotic cell, chromatin provides a large regulatory repertoire that capacitates the genome to give rise to many distinct epigenomes. We will discuss how reversible, heritable functional annotation mechanisms in chromatin may have evolved from basic chemical diversification of the underlying molecules.

**Address**
Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden

Corresponding author: Elsässer, Simon J (simon.elsasser@scilifelab.se)

**Current Opinion in Chemical Biology** 2017, **41**:36–42

This review comes from a themed issue on **Mechanistic biology**

Edited by **Gregory A Weiss**

**http://dx.doi.org/10.1016/j.cbpa.2017.10.001**

1367-5931/© 2017 Elsevier Ltd. All rights reserved.

The question of how one genome gives rise to many phenotypically different cell types has shaped the field of epigenetics since its conception over 50 years ago. More generally, any stable phenotypic variation on a constant genetic complement mandates the existence of information encoded on top of genetic information (thus 'epi-' genetic). Such information must be reversible in principle (thus not manifested in permanent genetic changes) but persistent beyond a transient period (e.g. a signaling response), for example persistence through mitosis and/ or meiosis.

Modern epigenetic research focuses on identifying and understanding the precise molecular variation that underlies phenotypic variation. Clues may be gained from an evolutionary perspective on the molecular diversification of the machinery that governs gene expression. At the core of epigenetic regulation are the factors and processes that regulate inheritance of and access to the information encoded in the DNA molecules itself, the so-called chromatin in eukaryotic cells.

## Chromatin at the root of eukaryotic evolution

The vastly increased genetic content and complexity of the eukaryotic, and in particular metazoan, cells requires a highly efficient mechanism to package the DNA molecules into the confined space of the nucleus. Histone proteins package the DNA molecules into the highly conserved structure of the nucleosome, forming the fundamental unit of chromatin. In Archaea, prototypic histones and, in addition, a number of other small basic proteins (such as Alba), are implicated in packaging DNA [1]. The exquisite structural conservation of nucleosomal histone-DNA contacts in Archaea, without the precise nucleosomal organization and regulatory tails of eukaryotic histones, suggest that histones evolved primarily as means to compact DNA [2••]. While the regulatory role of archeal chromatin is largely unknown, it shares basic features with eukaryotic chromatin, such as nucleosome-free regions at active transcription start sites [3].

Emil Heitz noted in 1928 that eukaryotic chromosomes consist of two fundamentally distinct compartments: "Euchromatin is genicly active, heterochromatin genicly passive. Heterochromatic chromosomes or pieces of chromosomes contain no genes or somehow passive genes [4]", shaping our understanding that such complex genome requires exquisite control over which parts of the genetic information are accessible to the gene expression machinery at any given point in time. Chromatin provides a biophysical barrier, in a way that renders large parts of the genome inaccessible to DNA-dependent biochemical processes such as transcription. Chromatin folding promotes physical interactions between genetic loci and restricts others. Such constraints mandate the existence of mechanisms that 'unmask', and coordinate in space and time, relevant genetic information.

## Chemical variation in the chromatin fiber

As exemplified by the macroscopic compartmentalization of euchromatin and heterochromatin, controlled variations in the chromatin fiber allow for the establishment and propagation of epigenetic states. They dynamically partition, in space and time, the genome in functionally distinct domains, be it the size of a single gene promoter region or a megabase stretch of the chromosome. Such variation arises from three main sources: the use of

histone variants, histone post-translational modifications (PTMs), and DNA modifications.

Histone variants are found in all studied eukaryotes with varying degree of amino acid sequence distance from their canonical counterpart [1]. Functional diversification early in the evolution of core histones through duplication and divergence of sequence is exemplified in the protozoan *Tetrahymena*: transcriptionally silent 'germline' micronuclei and euchromatic macronucleus are separate entities in the cell, facilitating the striking finding that two histone variants, hv1 and hv2, are exclusively associated with open chromatin of the transcriptionally active macronucleus [5,6,7]. Enriched at transcriptionally active regions of the genome, metazoan H2A.Z and H3.3 are functional orthologs of *Tetrahymena* hv1 and hv2, respectively. CenH3 is another universal H3 variant in eukaryotes. Centromeres are defined by the presence of CenH3 which provides the attachment point for microtubules during chromosome segregation. Epigenetic propagation of the cenH3 variant is necessary and in some organism sufficient for centromere maintenance [8].

## Origin of histone PTMs

A myriad of PTMs have been described on histone proteins, many of which are implied in a transient but potentially propagatable reversible modulation of DNA accessibility. Together with the sequence context of the PTM site, enzymes that establish ('write'), or remove ('erase') it, and modules that recognize ('read') the modified amino acid can often be traced back far in evolution. In particular, lysine residues are subject to chemically and structurally diverse modifications including acylation (foremost acetylation), methylation, and ubiquitination. Positively charged lysine residues provide the electrostatic basis for compacting the highly negatively charged DNA molecule. Modulation of the charge of lysine residues, i.e. by acetylation thus provides the most simple and direct means to control DNA accessibility. It is thus not surprising that phylogenetic analyses of histone proteins reveals a strong coevolution of the core histone fold providing biophysical packaging of DNA into the nucleosome and regulatory sites mainly clustered in the unstructured histone tails implied in regulating access to the underlying DNA [9,10]. The fundamental importance of histone PTMs in eukaryotic chromatin regulation is supported by the conservation of at least some well-studied histone PTMs (e.g. H3K9ac, H3K9/K27me3) down to protozoans at the root of the eukaryotic tree [9].

The evolution of major classes of PTMs, including phosphorylation, acetylation and methylation preceded the origin of eukaryotic cells. Lysine acetylation is present in all kingdoms of life albeit a nonenzymatic mechanism may have been predominant in early evolution [11]. $NAD^+$ dependent Sirtuin deacetylases are conserved in bacteria, Archaea and eukaryotes [12]. Archaeal chromatin protein Alba is known to be lysine-acetylated, which, akin to histone acetylation, lowers its affinity to DNA [13]. These data suggests that even the oldest known histone-like proteins were subject to PTMs akin to modern histones.

As another example, the SET domain, the catalytic core of most archeal and eukaryotic lysine methyltransferases originates in bacteria [14]. For example, the Gö1-SET catalyses lysine trimethylation of an Archaeal chromatin protein [15]. Thus it is reasonable to assume that an archetypical mode of regulation, e.g. via acetylation and methylation of conserved lysine residues was already in action at the origin of eukaryotic histone proteins.

Donor molecules for common PTMs (e.g. Acetyl-CoA for acetylation, NAD+ for deacetylation, ATP for phosphorylation, S-Adenosyl-Methionine for methylation, FAD or 2-ketoglutarate for demethylation) are primary metabolism intermediates. It is thus hypothesized that PTMs evolved as means of metabolic feedback regulation. While the archetypical substrates of such PTMs are likely proteins involved in metabolism, it appears that a large evolutionary expansion of enzyme families allowed building more complex regulatory patterns for epigenetic control of gene expression. Nevertheless, much evidence has accumulated in recent years that even the highly diversified 'writer' and 'eraser' systems of higher eukaryotes including humans are still tightly coupled to the metabolic state of the cell and entire organism through direct effects of substrate availability and product inhibition [16].

## Origin of DNA methylation

DNA bases are also subject to posttranslational modifications, providing the most direct means of 'annotating' the DNA sequence with functionally relevant information. Most well characterized in mammals, cytosine 5-methylation (5meC) varies greatly in global abundance across metazoans. Further chemical variation is achieved through oxidation of 5meC leading to hydroxymethyl, formylcytosine and carboxylcytosine (5hmC, 5fC, 5caC) [17]. The most ancient, reversible modifications of the DNA molecules are observed in all kingdoms of life. N6-adenine methylation (6meA) is widely utilized in bacteria as an essential component of the restriction-modification (RM) system against phage infection. Differential methylation status provides an efficient means to distinguish and target foreign vs. self genome. Bacteria also use DNA methylation as a means to regulate DNA replication, repair, gene expression and transposon repression [18,19]. These functions are performed by solitary DNA methyltransferases that evolved divergently from RM systems. This provides a basic paradigm for repurposing host-defence strategies for the regulation of other cellular processes. With a homologous set of writer and eraser,

Download English Version:

https://daneshyari.com/en/article/7693975

Download Persian Version:

https://daneshyari.com/article/7693975

Daneshyari.com