1506
UNIVERSITÀ
DEGLI STUDI
DI URBINO
CARLO BO

DIPARTIMENTO DI SCIENZE BIOMOLECOLARI

CORSO DI DOTTORATO DI RICERCA IN
SCIENZE DELLA VITA, SALUTE E BIOTECNOLOGIE

CICLO XXXII

Epigenomic profiling of archived FFPE tissues by enhanced PAT-ChIP technology and *in vivo* decoding of the aging-associated epigenetic drift and possible role of caloric restriction

SSD: MED/46

RELATORE
Chiar.mo Prof. **Fanelli Mirco**

DOTTORANDO
Persico Giuseppe

CO-RELATORE
Chiar.mo Prof. **Amatori Stefano**

ANNO ACCADEMICO 2018-2019

Table of Contents

Abstract

In the last 25 years, chromatin immunoprecipitation (ChIP) has become a powerful experimental approach to better understand the role of epigenetic modifications. Developed in the 1980s by Gilmour and Lis, chromatin immunoprecipitation makes possible the study of histone post-translational modifications (HPTMs) or other chromatin-associated proteins (e.g., transcription factors) to better understand their crucial role in reversible chromatin remodelling and regulation of gene expression.

Coupled with next-generation sequencing, chromatin-immunoprecipitation (ChIP-sep) allows the mapping of HPTMs over the entire genome unrevealing the so-called "epigenome".

Recent studies indicate that alterations of the epigenome is one of the most important hallmark of aging process, attracting much interest due the potential reversibility of epigenetic marks that makes them promising therapeutic targets to delay or minimize age-related diseases and potentially extend lifespan.

Although the majority of ChIP studies have been conducted on cultured cells with several limitations, the main being the alteration of the epigenetic profile in consequence of the adaptation of cells to tissue culture conditions. However, in 2010 Fanelli and colleagues introduced a modified version of ChIP, named pathology tissue-chromatin immunoprecipitation (PAT-ChIP), that allows chromatin extraction and immunoprecipitation from formalin-fixed and paraffin-embedded (FFPE) tissues. Formalin fixation followed by embedding in paraffin is the most cost-effective and simple method used to storage biopsy specimens in different therapeutics areas such as oncology. By extending the application of chromatin studies to clinical patient samples, PAT-ChIP makes possible epigenomic studies in a vast number of clinically

annotated tissues stored in pathology archives, providing an unprecedented opportunity to understand the epigenetic mechanisms underlying genome activity in progression of several human pathologies including aging-associated diseases.

However, due to the lack of standardization in the formalin fixation procedure, many FFPE tissue specimens stored in hospital archives or in tissue banks result heavily crosslinked and thus not suitable for genome-wide chromatin immunoprecipitation studies.

To overcome this problem, part of my Ph.D activity was dedicated to the improvement of PAT-ChIP to allow histone epigenomic studies also using "complex" biological samples. Thanks to these studies, a new procedure, called EPAT-ChIP, has been developed and recently published in "Clinical Epigenetics" journal (Amatori-Persico *et al.*, 2018). At the same time, in collaboration with the European Institute of Oncology of Milan, we exploited the already set PAT-ChIP-seq protocol to characterize four different epigenomic landscapes during aging process in mice livers and the possible effects of caloric restriction on them.

The first part of this work provides a detailed overview of all the procedures used to improve the original PAT-ChIP protocol, while the second part describes the global characterization of aging-associated epigenetic landscapes and the effects induced by CR.

1. Epigenetics

Conrad Waddington introduced the term *epigenetics* in the early 1940s defining it as "the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being". In the original sense of this definition, epigenetics referred to all molecular pathways that modulate the expression of a genotype into a particular phenotype. Today the term epigenetics is generally accepted as "the study of ***heritable*** and ***reversible*** changes in genomic function that do not entail a change in DNA sequence".

Unlike what happens with prokaryotic organisms, the organization and control of eukaryotic gene expression cells is very complex.

In eukaryotes, DNA is associated with protein complexes and packed into a highly organized and dynamic structure called *chromatin*, whose state of condensation regulates the accessibility of transcription factors to the DNA molecule. Chromatin is not a mere depository of the genomic content but rather a signal transduction platform for extracellular or intracellular signal that regulates all genomic function, including gene expression, DNA replication, cell division and genome stability. Upstream signals are translated into either transient or long-lasting changes of chromatin, thereby allowing chromatin to serve the double function of adapting cell to the environment changes while maintaining their lineage or identity. As mentioned before, these epigenetics modifications do not change the DNA sequence, but consist in the transfer of chemical groups to DNA (DNA methylation) or to a specific histone protein (histone modifications). These modifications (Figure 1) interfere with chromatin structure changing it in two distinct forms, that were originally defined by morphology as darkly stained constitutive *heterochromatin*, which remains condensed and repressed throughout the

cell cycle, and as lightly stained *euchromatin*, a de-condensed chromatin conformation much more accessible than heterochromatin and containing the majority of actively expressed genes (Felsenfeld and Groudine - 2003).

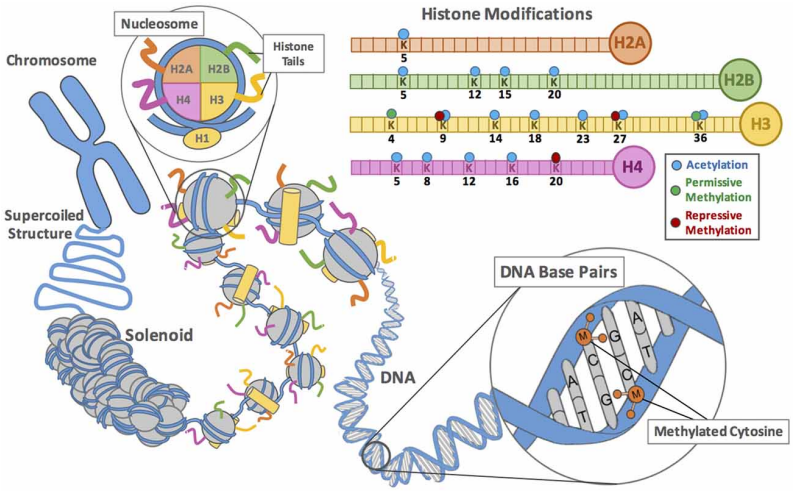


Figure 1. Epigenetic modification and chromatin structure. Each histone tail can undergo numerous post-translational modifications. In mammals the most common forms are acetylation and methylation of lysine. DNA can be chemically modified by methylation of cytosine of CpG dinucleotide. Epigenetic modifications through a remodelling of chromatin structure regulate the gene expression.

With the advancement in scientific research, the key players underlying these changes have been identified as ***epigenetic modifiers*** of chromatin. These epigenetic players are categorized as ***writers***, ***readers*** *and* ***erasers***. *Writers* are defined as those activities that introduce various chemical modifications on DNA and histones, *readers* as specialized domain containing proteins that identify and interpret those modifications, while *erasers* are the enzymes proficient in removing these chemical tags (LaSalle *et al*., 2013).

Increasing evidence shows that environmental and lifestyle factors (understood as "typical way of life or manner of living characteristic of

an individual or group") may influence epigenetic mechanisms with the potential to change the health status.

1.1 DNA methylation

The most widely studied epigenetic modification in mammals, including humans, is DNA methylation.

It is worthy to note that, unlike histone modifications, DNA methylation is present in both prokaryotic and eukaryotic organisms. In prokaryotes DNA methylation involves both cytosine and adenine bases, whereas in eukaryotes it occurs manly on cytosine localized in CpG reach regions, known as CpG islands.

In general, DNA methylation acts as a stable and heritable epigenetic mark generally associated with repressed chromatin states and inhibition of transcriptional initiation, playing a key role in genomic imprinting, where hyper-methylation at one of the two parental alleles leads to monoallelic expression. A similar gene-dosage reduction is observed in X-chromosome inactivation in females.

DNA methylation can inhibit gene expression by various mechanisms: *i)* it can promote the recruitment of methyl-CpG-binding domain (MBD) proteins. MBD family members in turn recruit histone- modifying and chromatin-remodelling complexes (Ghosh *et al.,* 2001), *ii)* can also directly inhibit transcription by precluding the recruitment of DNA binding proteins such as transcription factors (TF) from their target sites.

DNA methylation is mediated by DNA methyltrasferases (DNMTs) that catalyze the transfer of a methyl group from *S*-adenosyl methionine (SAM) to the cytosine (Lyko - 2018). In mammals, five members of the DNMT family have been reported: DNMT1, DNMT2, DNMT3a, DNMT3b and DNMT3L, but only DNMT1, DNMT3a and DNMT3b

possess methyltransferase activity.

DNMTs families are classified into *de novo* DNMTs (DNMT3A and DNMT3B) and maintenance DNMTs (DNMT1). DNMT3A and DNMT3B are thought to be responsible for establishing the pattern of methylation during embryonic development while *de novo* DNMTs are highly expressed in embryonic stem (ES) cells and downregulated in differentiated cells.

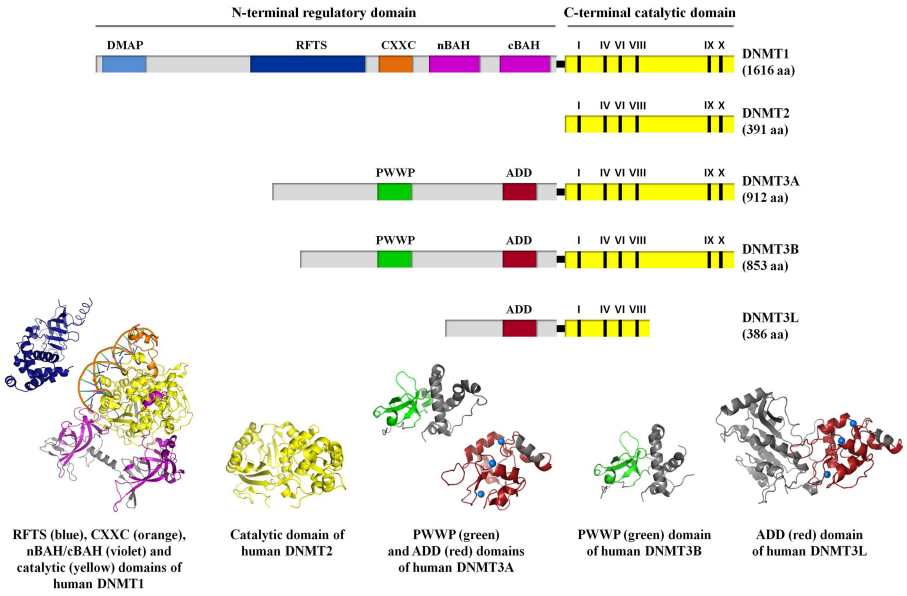


Figure 2. DNA methyltrasferase enzyme family. DNMTs families are classified into *de novo* DNMTs (DNMT3A and DNMT3B) and maintenance DNMTs (DNMT1). These enzymes catalysed the transfer of a methyl group from *S*-adenosyl methionine (SAM) to the cytosine localized in CpG reach regions.

1.2 Histone modification

Histones and their post-translational modifications (PTMs) play a crucial role in the organization of chromatin structure and regulation of gene transcription. Generally, histone classifications comprise the main histones or their variants H1, H2A, H2B, H3, and H4.

The fundamental building block of chromatin is know as *nucleosome* and consists of DNA wrapped around an octamer of histones (Strah and Allis - 2000). Each octamer contains two units of each variant histone H2A, H2B, H3, and H4 forming the histone core, while DNA linker connecting nucleosomes associates with the main form or variants of the linker histone H1. The four histones (H2A, H2B, H3, H4) are relatively similar in structure and are highly conserved through evolution; their N-terminal tail, which protrude from the center of the nucleosome core, are the main – and most characterized, protein regions in which post-translation modifications occur. HPTMs include acetylation, methylation, phosphorylation, ubiquitination, and SUMOylation, modifications that occur on several amino acid residues of histones tails. Each combination of HPTM/a.a. residues, determines a specific epigenetic effect (Figure 2).

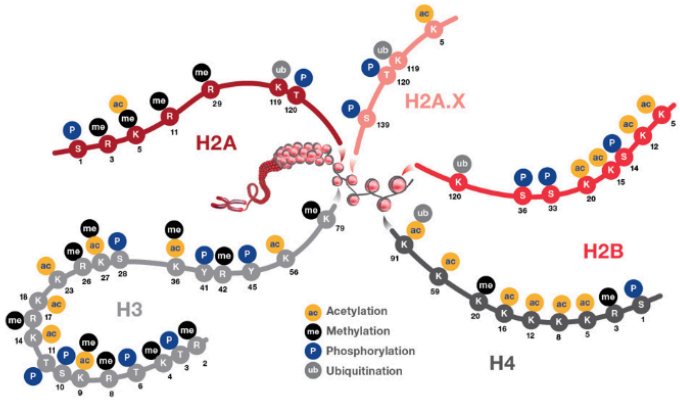


Figure 3. The main histone post-translational modifications (HPTMs). Different HPTMs occur on several amino acid residues determining specific epigenetic effects.

As mentioned before, these modifications, through the switch from euchromatin to heterochromatin state and vice versa, regulate several biological processes such as gene regulation, DNA repair, chromosome condensation, and spermatogenesis.

In mammals, high levels of methylation of lysine 9 and lysine 27 of histone H3 (H3K9, H3K27) (Sharakhov and Sharakhova - 2015) and concomitant low levels of acetylation of histone tails are typically associated with heterochromatin while trimethylation of lysine 4 or acetylation of lysine 27 of H3 (H3K4me3 or H3K27ac) are associated with euchromatin.

Histone acetylation

Allfrey and colleagues described for first histone acetylation in 1964. Now we know that the acetylation of lysines is highly dynamic and regulated by the opposing action of two families of enzymes, histone acetyltransferases (*HATs*) and histone deacetylases (*HDACs*).

HATs utilize acetyl-CoA as cofactor and catalyse the transfer of an acetyl group to the ε -amino group of lysine side chains. This covalent modification neutralizes the lysine's positive charge and this action has the potential to weaken the interactions between histones and DNA.

There are two major classes of HATs: *type-A* and *type-B*.

The type-A HATs are located in the nucleus and can be classified into at least three separate groups depending on amino acid sequence homology and conformational structure: GNAT, MYST and CBP/p300 families (Hodawadekar and Marmorstein - 2007).

These proteins are often found associated in large multiprotein complexes (Yang *et al*., 2007) that play important roles in controlling enzyme recruitment, activity and substrate specificity

The type-B HATs are predominantly cytoplasmic, and acetylate free histones but not those already deposited into chromatin. Type-B HATs acetylate newly synthesized histone H4 at K5 and K12 (as well as certain sites within H3), and this pattern of acetylation is important for deposition of the histones, after which the marks are removed (Parthun *et al*., 2007).

HDAC enzymes oppose the effects of HATs by removing lysine acetylation, an action that restores the positive charge of lysine. This potentially stabilizes the local chromatin architecture and is consistent with HDACs being predominantly transcriptional repressors. There are four classes of HDAC: ***Classes I*** and ***II*** contain enzymes that are most closely related to yeast *scRpd3* and *scHda1*, respectively; ***class IV*** has only a single member, *HDAC11*, while ***class III*** (referred to as sirtuins) are homologous to yeast *scSir2*. This latter class, in contrast to the other three classes, requires NAD^+ as specific cofactor for its activity (Seto and Yoshida - 2014).

In general, HDACs have relatively low substrate specificity; their specificity is indeed modulated by the incorporation of HATs in large multi-subunit protein complex. For instance, *HDAC1* is found together with *HDAC2* within the *NuRD*, *Sin3a* and *Co-REST* complexes (Yang *et al*., 2007). Thus, it is difficult to determine which activity (specific HDAC and/or complex) is responsible for a specific effect.

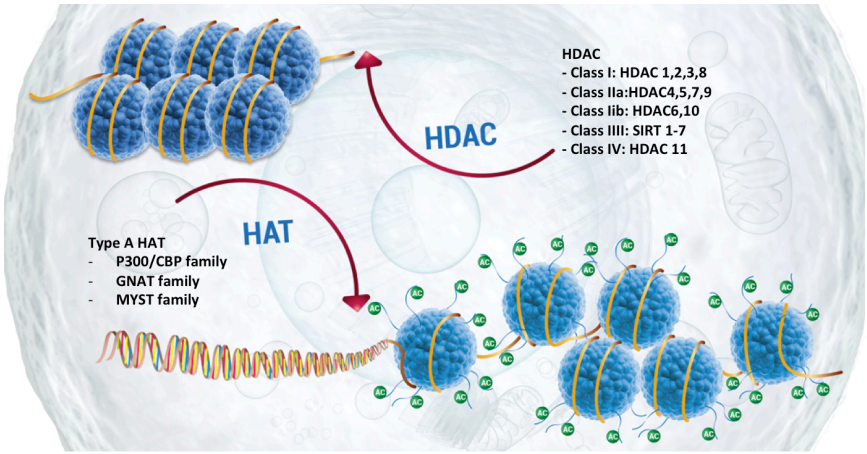


Figure 4: Histone acetylation. Histone acetylation is regulated by two families of enzymes: *histone acetyltransferases* (*HATs*) and *histone deacetylases* (*HDACs*). The HATs utilize acetyl-CoA as cofactor and catalyse the transfer of an acetyl group to the ε -amino group of lysine side chains while HDAC enzymes by removing lysine acetylation restore the positive charge of lysine.

Histone methylation

Unlike acetylation, histone methylation exerts its effect without altering the charge of the histone proteins. Moreover, in this case, the lysine ε-amino group of proteins can accept up to three methyl groups, resulting in either mono-, di-, or trimethyl lysine, (me1, me2, or me3) exerting distinct functions. Histone methylation is catalysed by enzymes named histone methyltransferases (HMT) that can be divided in two main categories: ***lysine-specific (KMT)*** and ***arginine-specific (PRMT)*** (Bannister and Kouzarides - 2005).

The histone methylation reaction, which is catalysed by carrying a methyl group (-CH_3) derived from S-adenosylmethionine (SAM) on a lysine or arginine residue, can trigger the formation of either transcriptionally active euchromatin or transcriptionally inactive heterochromatin,

depending on the specific methylated residue, the number of methyl groups added and the availability of factors that remodel chromatin.

For example, lysine 4 could be subjected to three different methylation events, whose signals are progressively localized closer to TSSs as the modification moves from *mono-* to *di-* to trimethylation (Barski *et al.*, 2007).

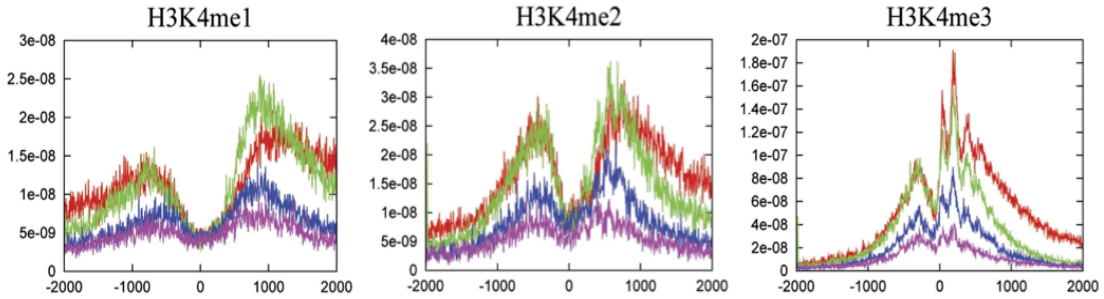


Figure 5. H3K4 methylation status. Different H3K4 methylation profiles near the transcription start sites of highly active (red), two stages of intermediately active (green and blue) and silent genes (purple) - Barski et al., 2007.

Two major regions of enrichment are normally detected for each modification: –900 and +1000 from TSS for H3K4me1, –500 and +700 from TSS for H3K4me2, and –300 and +100 from TSS for H3K4me3.
On the other hand, trimethylation of lysine 27 of histone H3 (H3K27me3) trigger heterochromatin assembly being related to gene silencing and transcriptional repression (Boyer *et al.*, 2006). The H3K27me3 signals are modestly elevated on silent genes and reduced when genes are expressed. H3K27me3 can be bind by Polycomb Group proteins (PcG proteins), known to mediate chromatin repression, and to be implicated in inactivation of the X chromosome and cell proliferation control. Like trimethylation of lysine 27, also trimenthylation of lysine 9 on histone H3 (H3K9me3) is associated with costitutive heterochromatin and therefore with gene silencing. This modification is involved, for example, in the

inactivation of the X chromosome and is associated with highly condensed centromeric and pericentromeric regions (Sullivan e*t al*., 2004).

In general, methyl groups are believed to turnover more slowly than many other PTMs and histone methylation was originally thought to be irreversible, at least until the discovery of histone H3K4 demethylase *LSD1* (Lysine Specific Demethylase 1, also known as KDM1A) by Byvoet *et al.,* in 1972. Frequently, LSD1 is frequently found to be associated with a transcriptional co-repressor protein (CoREST) and histone deacetylase (HDAC) 1/2 to form a complex able to demethylate H3K4me1/2, but when LSD1 is complexed with the androgen receptor, it demethylates H3K9. This has the effect of switching the activity of LSD1 from a repressor function to that of a coactivator (Klose and Zhang - 2007).
In 2006 JMJD2, the first trimethyl lysine demethylase that demethylates H3K9me3 and H3K36me3, was also described (Whetstine *et al*., 2006).

1.3 Histone code and chromatin states

Form the discovery of histone modifications scientists have tried to attribute a specific function at each modification. As mentioned before, modifications such as H3K4me3 or H3K27ac are known to be associated with active genes, while H3K27me3 is associated with repressive genes. However, is know that the histones can be modified at different sites and that core histones forming the nucleosome can carry several modifications at same time, giving rise to cross-talk among the different

marks. Communication among histone modifications (Fig. 6) can occur within the same site, in the same histone tail and among different histone tails. The relationships between DNA methylation and histone modifications, has been called "*epigenetic code*" (Jenuwein and Allis - 2001) and seems to be much more complex than believed because act through different mechanisms:

- First, communication at the level of a single histone tail (*cis effect*); for example, methylation on H3K9, can inhibit acetylation of the H3 tail and methylation of H3K4 (*Fischle et al., 2003*);
- Second, interactions at the level of nucleosomes mean that the modifications on different histones can affect each other (*trans effect*). For example, trimethylation on H3K9 is required for the induction of H4K20 trimethylation (An - 2007).
- Third, DNA methylation and histone modification pathways can influence each other and establish the epigenetic landscape important for development, somatic cell reprogramming, and tumorigenesis. Relationships between DNA methylation and histone H3 methylation, particularly H3K4, H3K9, and H3K27, have been observed. In addition, there is also a strong anti-correlation between different histone methylations; for example, it is known that the presence of the H3K4me mark prevents de novo methylation of CpG islands in the embryo (Cedar and Bergman 2009).

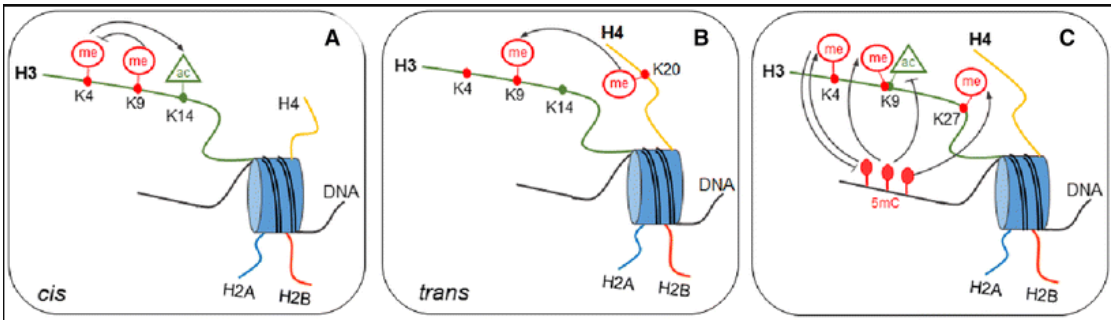


Figure 6. Epigenetic code: The Cross-talk between chromatin marks and DNA methylation. (A) Communication at the level of a single histone tail is defined *cis effect*; (B) different histones can affect each other (the *trans* effect); (C) DNA methylation and histone modification cooperate to enstablish the epigenetic landscape.

Understanding this code is one of the most compelling challenges in epigenomic studies field. Several bioinformatical tools have been developed in the intent of analyse the epigenomic landscape of several HPTMs at same time, to identify the so-called "*chromatin states*" (Baker – 2011). Identification of these states, characterized by different HPTMs levels at same locus (Fig. 7), allows to distinguish regions with different functions such as promoters, enhancers, transcribed, repressed, and repetitive regions, providing a systematic annotation of DNA elements and regulatory control regions.

For example, H3K4me1 alone marks primed enhancers, while H3K4me1 combined with H3K27ac mark active enhancers (Creyghton *et al.*, - 2010). H3K4me3 is predominant feature of active promoter, whereas poised (bivalent) promoter state is characterized by the occupancy of H3K4me3 and H3K27me3 modifications (Harikumar and Meshorer – 2015).

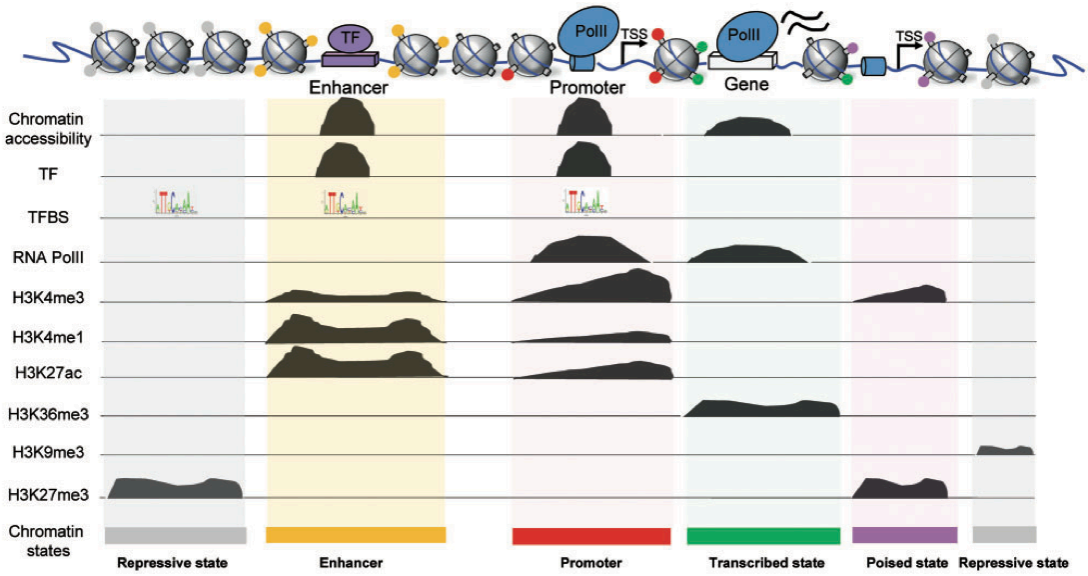


Figura 7. Chromatin states. The integrative analysis of histone modifications, TF and RNA pol II binding site allows to distinguish combinations of these features (chromatin states) associated with different functional regions such as promoters, enhancers, transcribed, repressed, and repetitive regions, providing a systematic annotation of DNA elements and regulatory control regions.

2. Chromatin immunoprecipitation

The biological significance of interactions of nuclear proteins with DNA in the context of gene expression, cell differentiation, or disease has immensely been enhanced by the advent of chromatin immunoprecipitation (ChIP). The technique involves the extraction and fragmentation of chromatin from a biological matrix, its selection through antibodies that specifically recognize a chromatin-associated protein (e.g., transcription factor) or histone post-translational modification, and the study of the DNA sequences present in the isolated chromatin.
Coupled with Next-Generation Sequencing approach (ChIP-seq), the technique make possible to know virtually all the binding sites of different transcription factors, RNA polymerase II or to study the histone post-translational modification at genome-wide level.

The original chromatin immunooprecipitation protocol involved the use of UV light to cross-linking proteins and DNA; later Solomon *et al*., (1988) introduced formaldehyde as a cross-linking agent able to form covalent bonds between DNA and proteins distant from each within 2Å.

Today there are two main methodologies for carrying out ChIP experiments. The first one, called X-ChIP, is based on chromatin fixation with formaldehyde and subsequent extraction and fragmentation by sonication. This variant minimizes the chances of chromatin rearrangements during preparation and precipitation and allows to study also non-histone proteins (e.g. transcription factors) which interact with DNA in a less stable way than histones.
Another variant called native chromatin immunoprecipitation (N-ChIP) is performed using native (not fixed) chromatin, sheared by micrococcal

nuclease (Mnase) digestion, which digests linker DNA, leaving nucleosomes intact (Turner - 2001).

The standard ChIP protocol methods (Fig. 8) involves the following steps:

- Crosslinking of DNA and associated protein performed in living cells or tissues (this step is omitted in N-ChIP); When ChIP is performed on tissue cultured cells, this step is generally carried out using formaldheyde at final concentration of 1% incubated for 10–15 min at 37 °C. If tissue specimens are used, the crosslinking condition step must be adjusted in terms of time of incubation in formaldehyde, considering both size and complexity of tissue.
- After cell lysis step through incubation in a solution containing detergents, chromatin fragmentation and subsequent extraction take place through physical processes (sonication) and/or through an enzymatic digestion with micrococcal nuclease.
- Using a specific antibody, DNA fragments associated with the protein of interest (or associated with the HPTMs of interest) are immune-selected. This is a critical step, that may determines the feasibility of the entire procedure.
- Immunoprecipitated complexes are washed to remove non-specifically bound chromatin and then, if fixation was performed, the cross-link is reverted through incubation at high temperatures in the presence of high salt concentration.
- The DNA from the isolated chromatin is purified and analysed by quantitative PCR (qPCR) for single locus analysis, or by Next-Generation Sequencing (NGS) approach – for for genome-wide studies.

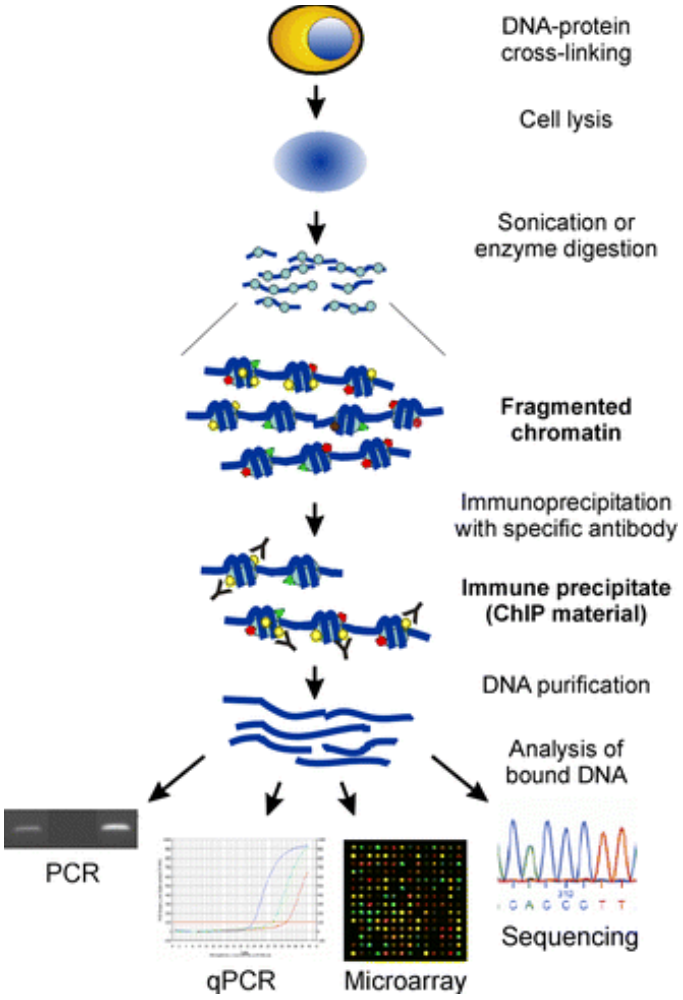


Figure 8. Chromatin immunoprecipitation steps and various methods of analysis. Fixed cells are lysated and chromatin is fragmented and extracted. Then, chromatin is immunoselected using specific antibodies directed against the protein of interest and the DNA associated with that protein is analysed through different methods.

3. Pathology-tissue chromatin immunoprecipitation (PAT-ChIP)

As discussed above, since its introduction, chromatin immunoprecipitation coupled with next-generation sequencing has become the most powerful approach to investigate specific chromatin-associated proteins (e.g., transcription factors, HPTMs).

Chromatin immunoprecipitation is conducted on chromatin extracted from cultured cells or from fresh/frozen tissues. However, in the first case it has been demonstrated that cultured cells change thair epigenome in function of the culture conditions, while fresh samples are always not available, especially in clinical practice. This is because clinical samples immediately after excision are normally formalin-fixed by immersion in a solution of neutral-buffered formalin for about 16-24 hours and embedded in paraffin (the so-called formaldehyde fixed and paraffin-embedded - FFPE - tissues). Fixation with formalin and subsequent paraffin-embedding are routinely used to preserve biopsies and maintain intact their cellular structure, including cross-linked DNA and proteins as well. However, due to the method of preservation, obtaining biomolecules from these samples has been an hard challenge for many years and the use of this biological resource has been limited mainly to immunohistochemistry and *in situ* hybridization-based techniques, allowing to study only few targets at the same time. In addition, these approaches require a prior knowledge of the target, allowing hypothesis-driven rather than discovery-based studies.

In 2010 Fanelli and colleagues have developed a new technique named pathology tissue chromatin immunoprecipitation (PAT-ChIP), a technique allowing extraction, immunoselection and high-throughput

analysis of chromatin derived from FFPE samples making possible epigenomic studies using innumerable FFPE tissues stored in hospital archives and tissue banks around the world, providing an unprecedented opportunity to understand the epigenetic mechanisms underlying genome activity in disease etiology and progression.

The original ChIP protocol was modified through introduction of initial steps designed to eliminate the presence of paraffin and to progressively rehydrate the samples. Being FFPE tissues usually heavily crosslinked samples, due long incubation times (usually 16–24 h and more at room temperature) and the high concentrations of formaldehyde (around 4%), the chromatin isolation required a combination of enzymatic digestion and physical approach (sonication).

3.1 Limitation of the original PAT-ChIP protocol

During years PAT-ChIP was applied by several research groups, (Sharma *et al*., 2013 – Fang *et al*., 2015 – Serra *et al*., 2013) giving new impetus to chromatin studies to the identification of new potential epimarkers in function of the clinical information of patients. However, some limitations of protocol have emerged. One of these was given by the heterogeneity of FFPE tissue slices, this problem was successfully circumvented by Amatori and colleagues in 2014 using laser capture microdissection (LCM) to increase the purity and homogeneity of the celllular populations under investigation. Nevertheless, the main problem and limitation of PAT-ChIP, is related to fixation status of tissues. Indeed, we have experience that the performances of the technique can be hindered when highly-fixed samples are used. Due to the lack of

standardization in the formalin fixation procedure in clinical practice, some users leave tissue biopsies in formalin solutions for a longer period (up to 72h) with a relatively low impact on standard analysis (e.g. immunohistochemistry) but compromising any further molecular analysis, in particular chromatin immunoprecipitation.

FA is a tight (2 Å) crosslinking agent that efficiently produces both protein–nucleic acid and protein–protein crosslinks. Amino and imino groups of amino acids (lysines, arginines, and histidines) and of DNA (primarily adenines and cytosines) are involved in formalin fixation forming a Schiff's base that can participate in a second linkage with an additional amino group and condense to give the final DNA–protein complex.
Extensive crosslinking to which FFPE archival samples can be exposed produces a dense network of crosslinked cellular biomolecules that can render chromatin extraction extremely challenging.

Additionaly, less efficiency in chromatin extraction is not the only problem caused by formalin fixation, since fixation has a strong impact on lysine residues, one of the most studied target where occurs several PTMs occur, determining the masking of the epitope and hampering subsequent immunoselection.

4. Next Generation Sequencing (NGS) and bioinformatics analysis

Genome-wide mapping of binding sites for transcription factors, or other DNA-binding proteins, cofactors and histone post-translational modifications, is essential for deciphering the gene regulatory networks that underlie various biological processes. The main tool for investigating these mechanisms is chromatin immunoprecipitation coupled with next-generation sequencing, also known as ChIP-seq. The ability to sequence tens or hundreds of millions of short DNA fragments in a single run is enabling studies that were considered impossible only ten years ago. There are diverse NGS techniques developed over the last few years although the majority share the same operating principle based on the sequencing by synthesis (SBS) technology (e.g Illumina platform – Fig. 9).

First of all, the DNA from ChIP is ligated to specific and know oligonucleotide called *adapters* and amplified to prepare a library (see library preparation in material and method section).

Then sequencing takes place by polymerization, adding one by one the four nucleotides (labelled or not, and chemically blocked at 3'-OH) whose progressive incorporation is detected and registered.

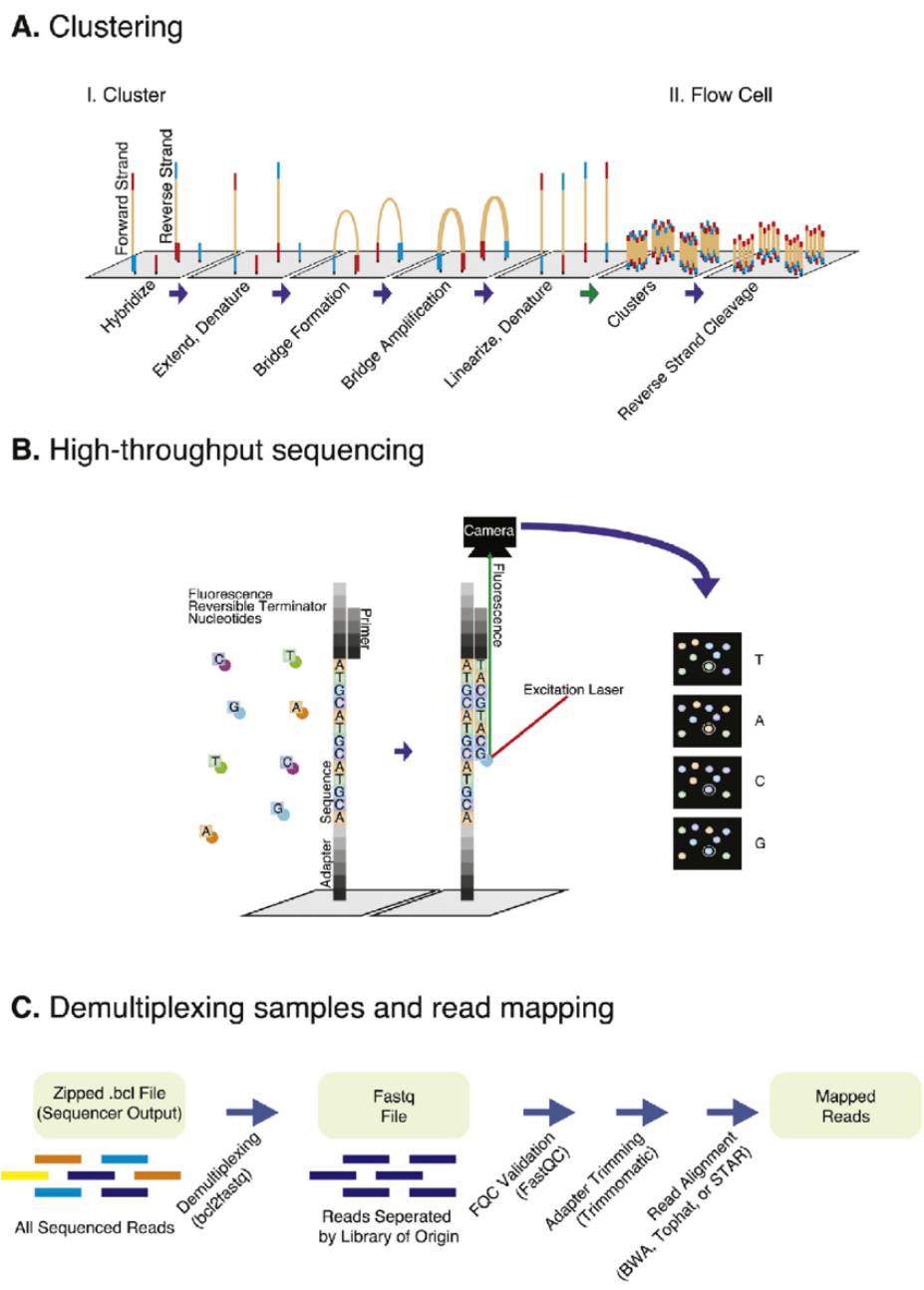


Figure 9. Illumina sequencing and data processing workflow.

A) Denaturated NGS library fragments are hybridize on a chip bearing millions of oligonucleotides covalently linked and complementary to the adapters. Bridge amplification is performed to amplify DNA and to generate clusters of the same fragment. Reverse filament is then cut and removed by washing to obtain clusters that consist of only forward filaments. **B)** Fragments are primed and sequenced. **C)** Raw data is demultiplexed into individual libraries and assessed for quality. Removing adapter reads reduces technical noise. Finally reads are aligned onto the assembly of interest.

The raw data coming from the sequencer are in *FASTQ* files, a text-based format for storing both a biological sequence and its corresponding quality score

After quality check and filtering of low quality reads, the remaining reads are *aligned* against an organism-specific reference genome using mapping programs such as Bowtie2 or BWA generating a Binary Alignment Map (BAM) file, a file type easier to read and process for later bioinformatics analysis. Once obtained a BAM file, read densities can be visualized on UCSC genome browser (or other software auch as Integrative Genomics Viewer – IGV) where it is possible to visualize the distribution of the sequenced reads over the entire genome. If the antibody selection was specific, a considerable proportion of the mapped reads will be dispersed by chance throughout the genome (the so-called background), while the others will cluster together constituting reads-enriched regions termed "peaks". Peaks can differ substantially depending on the protein or histone modification investigated. For example, the majority of transcription factors and many histone modifications like H3K4me3 and H3K27ac tend to have narrow peaks, with a size ranging from several hundred to a few thousand base pairs while other histone modifications such as H3K27me3 or H3K9me3 tend to form broad genomic domains with diffusive ChIP-seq signals, which can span up to thousands kilo base pairs.

Peaks of enrichment are identified using different softwares, the most used being model-based analysis of ChIP-seq (MACS and MACS2) which assign they a p-value reflecting their significance. Moreover, when an input control data is present, is possible to filter the peaks that are called and to assign each peak a *false discovery rate* (FDR) score, which is the likelihood that the peak is not valid.

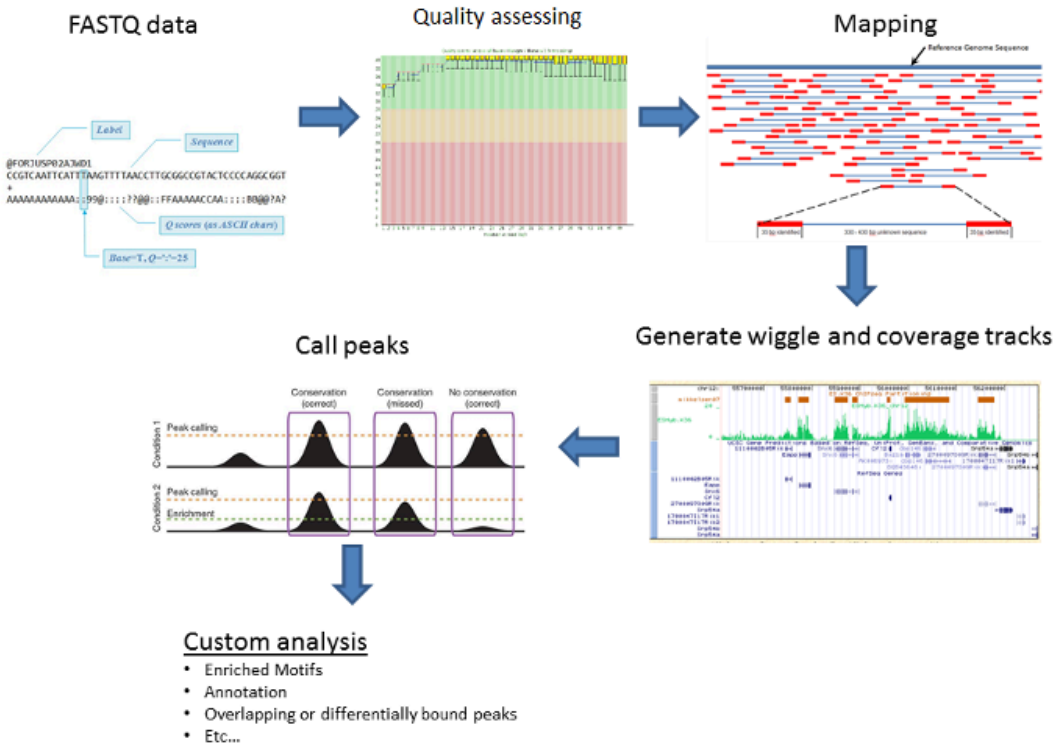


Figure 10. ChIP-seq data analysis workflow. Reads are filtered basing on a quality score and high-quality reads are mapped on a reference genome. Reads density can be visualized along the genome and enriched regions (peaks) can be identified and used for comparative analysis.

Other tools are then used downstream to annotate these enriched regions. Analysis can include (but is not limited to) peak comparison among samples to observe presence/absence of specific peaks, global peaks distribution around transcription start site (TSS), or to annotate the genomic features of peak, gene ontology (GO) or pathway analysis, and recurrent motif search, checking for quantitative significant changes in binding levels.

However, most experiments nowadays are designed to study several histone modifications at once and to this end various computational algorithms are developed.

ChromHMM and Segway, for example, allows to identify the specific combination patterns of histone modifications and classify the genome into a preselected number of *chromatin states*.

Using a multivariate Hidden Markov Models (HMM), ChromHMM split the genome in 200-nucleotide (or more) intervals, called bins. For each bins the tools determines the presence or absence (1 or 0) of each mark and use it to learn a chromatin-state model, and create an annotation of state occurrences across the genome.

5. Theories of aging

The epigenetics field has rapidly developed into one of the most influential areas of scientific research due to involvement in the regulation of several and essential biological processes such as embryo development, cellular differentiation and tumorigenesis. The number of pathologies linked to the dysregulation of epigenetic systems continues to grow and with it, the list of potential targets for epigenetic-based therapeutics.

By the last century, one of the most intriguing challenges for scientific community is to understand the aging process. Recent studies indicate that epigenetic alterations represent one of the most important hallmark of it, attracting much interest due potential reversibility of epigenetic marks makes them promise therapeutic targets to delay or minimize age-related diseases and potentially extend lifespan.

Denham Harman (1950s) describes aging as the result of the progressive accumulation of changes in the body which occurs over the time and which that increase in the chance of getting sick or dying.

In the past couple of centuries, scientists proposed several aging theories but none of them is sufficiently able to explain the aging process. Some of these theories indeed match with others, while others are completely different.

In a frequently cited paper published in 1990, Zhores Medvedev had attempted to make a rational classification of theories of aging, but over the years, gerontologists have resigned to the futility of formulating a unified theory of aging because the large body of descriptive data underlines the multifaceted, different and complex nature of aging.

The rates of aging progression are highly variable in different species, in organism within a species, in organs and tissues within an organism, leading to the conclusion that aging has a no universal cause or phenotype, except death.

Today aging is defined as complex multifactorial process shared by all living organisms, characterized by progressive decline in intrinsic physiological functions, leading to an increase of susceptibility to many diseases, including cancer.

Aging theories can be fall into two main categories: ***stochastic* and *non-stochastic theories***. The first group, also called *"error theories"*, consider aging as the result of environmental insults to living organisms that induce progressive damage at various levels (e.g., mitochondrial DNA damage, proteins or lipids oxidation).

In the second one, aging is thought to be depends on biological clocks regulating the timetable of the lifespan through the stages of growth, development, maturity, and old age: this regulation would depend on sequentially switching ON and OFF signals of genes expressed in nervous, endocrine, and immune systems responsible for maintenance of homeostasis and for activation of defense responses.

5.2 Stochastic theories of aging

Stochastic theories suggest that aging is the result of random accumulating changes that negatively affect biological systems. Aging could be the result of the accumulation of toxic by-products, damage due to nuclear radiation or other gradual deteriorative process. At this category belong many theories proposed such as "*wear and tear theory*", "*free radicals theory"* and "*somatic DNA damage theory"*.

Wear and tear theory

The so-valled "*wear and tear theory"*, asserts that the organism like a machine become damaged and eventually break down when utilized for a certain period. This theory propose that aging is simply the result of wear and tear, due to the deteriorating effect of processes such as oxidation, or other molecular damage due to ionizing radiation and toxic element uptake, or again other unavoidable natural processes. Aging is view as result of the accumulation of toxic by-products and, believing in this, the aging process is, theoretically impossible to revert. However, there are only few gerontologists that currently support the "wear and tear theories", mainly because this theory fails to explain enormous differences in lifespans between biochemically similar species. If aging is the result of fundamental limitations that presumably affect all organisms, why are lifespans of even very similar organisms so different (e.g naked mole-rats – 28.3 years vs nommon rock rats – 4.2 years)?

Free radical theory of aging (FRTA)

Originally described by Denham Harman *et al.,* the error theory of aging is one of the most prominent theories to explain aging. This theory proposes that the free radicals - continuously generated during the life of the cell – aren't counterbalance (in aged biological system) by cell antioxidant system causing damage. There is a great deal of experimental evidence in support of this theory. Old animals show a higher index of oxidation than young ones and indeed they accumulate oxidized proteins, DNA forms, and lipid (Stadtman - 1992).

Other experimental evidences support this theory, for instance, increasing antioxidant defense results in an extended life span, as demonstrated in experiments using flies. In fact, Tower (2000) show an increased life span

of transgenic flies expressing superoxide dismutase (SOD) indicating that free radical scavenging enzymes are sufficient to delay aging in *Drosophila*.

Extension of *C. elegans* life span, by synthetic small molecules that mimic catalase and/or SOD, demonstrates that antioxidant compounds can delay aging in worms (Melov *et al*., 2000).

These results indicates that free radical damages in short-lived organism opposes longevity, but in long-lived organisms such as mice models, a antioxidant diet can effectively reduce the accumulation of oxidized molecules but fail to extend lifespan. Calorie restriction is an intervention that prolongs the life span of nearly every organism to which it has been applied. In rodents, calorie restriction reduces generation of ROS from isolated mitochondrial preparations and attenuates the accumulation of oxidative damage.

The free radical theory argues that mutations in mitochondrial DNA accelerate free radical damage by introducing altered enzyme components into the electron transport chain. Faulty electron transport results in elevated free radical leakage and ultimately more mitochondrial DNA mutation and exacerbated oxidant production. This “vicious cycle” of mutation and oxidant production eventually leads to cellular catastrophe, organ failure and senescence. Another hypothesis also belonging to FRTA argue that radicals cause aging when oxidized proteins accumulate in cells. An age-dependent reduction in the ability to degrade oxidative proteins may contribute to the build-up of damaged and dysfunctional molecules in the cell. The free radical theory of aging could “interact” with *Somatic Mutation Theory of aging* where accumulation of genetic mutations in somatic cells are thought to be responsible of decrease in cellular functions and ROS damage may be an important source of these somatic mutations.

5.3 Non-stochastic theories of aging

Theories of aging based on accumulation of stochastic damage to macromolecules were widely support by gerontologists, even if it does not provide a convincing explanation for vastly different longevity of different cell types within the same individual, or individuals from different species living in the same environment. A series of pioneering experiment carried out in yeast, worm, flies, and mice have showed that lifespan could be strongly influenced by molecular pathway that have been conserved during evolution, leading new theories where aging is thought to be a "deterministic" programmed process.

In the early '80 years, Klass *et al.,* (1983) observed that *C. elegans* lifespan could be altered, but only in 1988 Friedman and Johnson showed that life extension of worm up to 65% was due to a mutation of gene that called *age-1*. This is the first *gerontogene* (refers to any such genetic elements that are involved in the regulation of aging and life span) discovered.

In the 1993, Kenyon C. showed that mutation in gene *daf-2 gene* cause fertile, active, adult *C. elegans* hermaphrodites to live more than twice as long as wild type.

Since then, dozens of genes were identified influencing lifespan supporting the "*longevity programmed theory"* where the presence of these genes is thought to be crucial to allow the organism survival in presence of environmental stresses.

Under stress condition such as long exposure at cold temperature or reducing calorie intake, getontogenes are silenced to allow to exceed the normal lifespan of organism by entering in "maintenance mode" (reverted with improvement of environmental conditions) characterized by

biological processes changes such as hypometabolism, high stress-resistance and low or no fertility. Notewhorty, studies in *C. elegans* show that genetic program actively promote longevity when animal models are expose at cold temperatures (Xiao *et al*., 2013).

Many mutations that extend life perturb endocrine signalling. In fact, *daf-2* gene encodes the only *insulin/IGF-1 receptor* expressed in worms (Fig. 11), while *age-1*gene encodes a homologue of mammalin *phosphoinositide 3-kinase* (*PI3K*) catalytic subunit (Morris *et al*., 1996; Kimura *et al*., 1997). Both proteins are involved in *insulin/IGF-1 pathway* (*IIS pathway*), a well conserved evolutionarily pathway across organism. The IIS pathway acts as a food and stress sensor during development. When food is abundant, worms develop rapidly and uninterrupted through four larval stages to reach adulthood. If worms develop in hot, food-limited or overcrowded conditions (stress conditions), they enter in an alternative long-lived larval state called *dauer* in which reproductive maturity is delayed and stress resistance increases.

IIS pathway is one of the best-understood molecular mechanisms able to influence the lifespan. Briefly, signalling through *daf-2* (*IGF-R* in human) activates *age-1* (*PI3K* in humans), which leads to the phosphorylation of *daf-16* and its inactivation by nuclear exclusion. In the absence of *IIS* activity or in *daf-2* or *age-1* mutants, *daf-16* enters the nucleus and enacts a transcriptional program that is thougth doubles worm life span (Lin *et al*., 1997; Lee *et al*., 2001 – Fig.11).

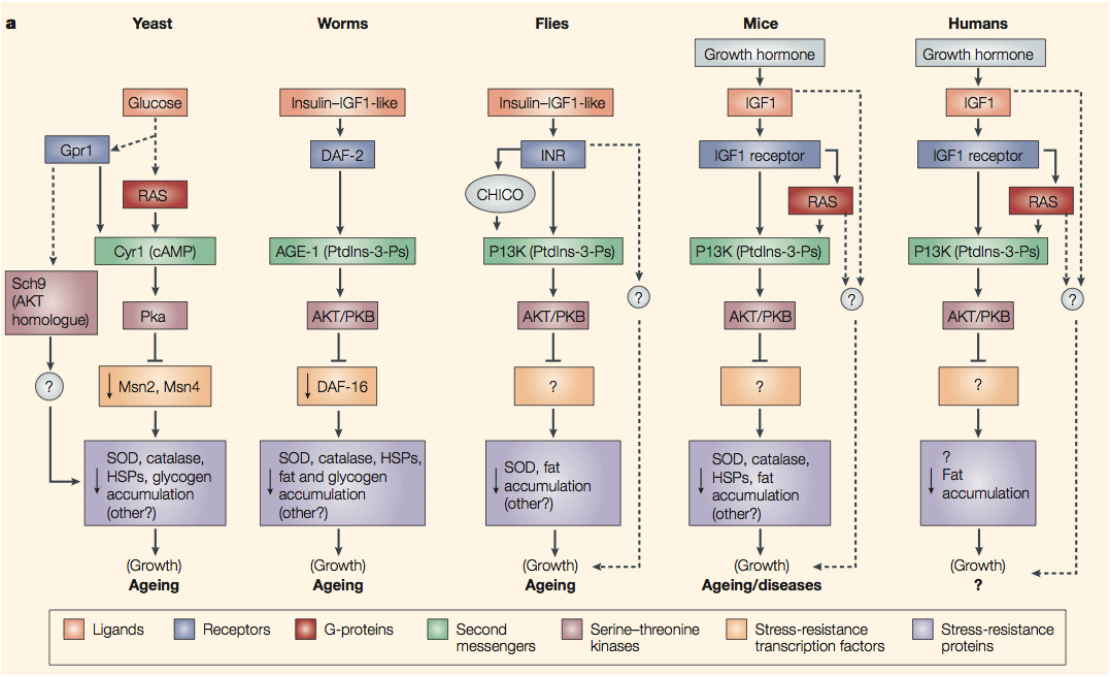


Figure 11. Longevity pathway in yeast, worms, flies mice (and human?). A series of pioneering experiment carried out in yeast, worm, flies, and mice have showed that lifespan can be strongly influenced by molecular pathway that have been conserved during evolution.

As mentioned before IIS is an evolutionarily conserved pathway that regulates life span across many organisms. *Drosophila melanogaster*, like *C. elegans*, has a single insulin-like receptor (*dInR*) that, when mutated, extends life span in a manner that is dependent on *dFOXO*, its *daf-16* homolog, (Tatar *et al.*, 2001). In mammals are present several *daf-2* homologs (*IGF-1R*), *insulin receptor A* and *B* (*IR-A* and *IR-B*) that display even greater complexity due their ability to form multiple homodimer and heterodimer pair. Despite these differences in insulin receptor expression, the functional consequences are similar; in fact, reduced IIS pathway activity results in extends life span in multiple mammalian species. Heterozygous *IGF-1R* knockout mice are long-lived (Holzenberger *et al.*, 2003) in addition IGF-1R mutations are highly represented in populations of centenarians (Suh *et al.*, 2008). Through

ChIP-seq and gene expression approach, genome wide analysis of *daf-16* in *C.elegans* was performed to extend our knowledge of how life extension is achieved in the worm. However, no single DAF-16 target that has been tested can recapitulate the long-lived phenotype of *daf-2* worms, suggesting that many disparate cellular pathways are regulated in concert to extend life span (Murphy *et al*., 2003; McElwee *et al*., 2004).

Similarly to IIS pathway, also the role of the TOR (target of rapamycin) pathway on aging is remarkably conserved. TOR is an evolutionarily conserved nutrient sensing protein kinase that regulates growth and metabolism in all eukaryotic cells. This kinase is inhibited by *rapamycin*, a by-product of the soil bacterium *Streptomyces hygroscopicus* (Vezina *et al*., 1975).

Rapamycin through *formyl peptide receptor 1* (FPR1), a peptidyl-prolyl cis-trans isomerase, regulate activities of two proteins identified as *TOR1* and *TOR2* in *S. cerevisae* (Heitman *et al*., 1991). However, in mammals there is only one gene encoding mammalian *TOR* (*mTOR*) and protein exists in two complexes:

- *TOR complex I* (TORC1 in yeast and mTORC1 in mammals), a rapamycin sensitive complex that controls temporal aspects of cellular growth mediated mostly through *S6 kinase 1* (*S6K1*);
- *TORC2* is rapamycin-insensitive and controls spatial aspects of growth within the cell and the effects are mostly mediated through *protein kinase B* (*PKB*/*Akt*) (Gonzalez *et al*., 2017).

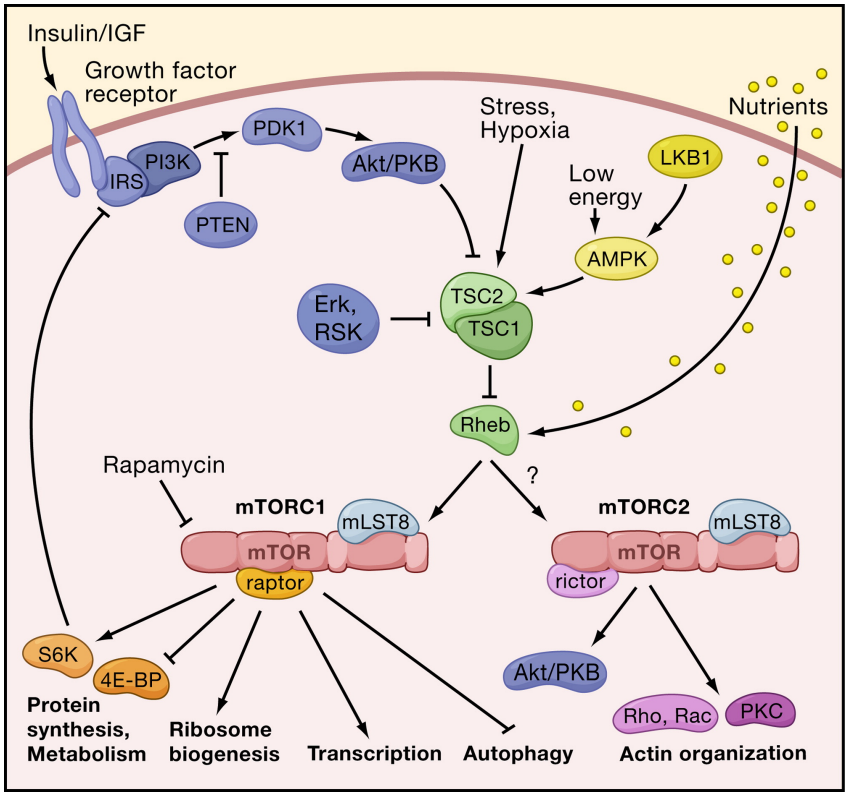


Figure 12. The TOR pathway and aging. mTOR is a key component of cellular metabolism that integrates nutrient sensing with cellular process that fuel cell growth and proliferation.

TORC1 is involved in many human diseases, including diabetes, obesity, heart disease, and cancer (Inoki and Guan, 2006). This pathway is also involved in lifespan control in multiple model systems; indeed, inhibition of TORC1 (hereafter referred to as TOR) extends lifespan in yeast, worms, flies and mice (Kapahi *et al*., 2010).

Large amounts of data show that signalling of TOR pathway can act both in parallel to but also interact with the *insulin/IGF-1 pathway* in flies (Kapahi and Zid, 2004). Combination of rapamycin and caloric restriction cause some additional longevity effects when compared to each single treatment (Bjedov *et al*., 2010) indicating that the effects of rapamycin

may not be identical to the mechanisms involved in caloric restriction-mediated extention of lifespan.

As mentioned before, the aging process is characterized by an increase of vulnerability to many diseases, due to an impairment of the different physiological systems, including the immune function.

In fact, it is well known that over the time there is a decrease in the resistance to infections and an increased incidence of autoimmune processes and cancer, indicating the presence of a less competent immune system.

6. Epigenetics of aging

Epigenetics is emerged as a major, nongenetic mechanism strongly related to the aging process is epigenetics. Epigenetic alterations affect all types of cells and tissues throughout life and have been linked to cancer development and are retained one of the major hallmarks of aging. During the aging process, the epigenome undergoes a progressive loss of its configuration that results in a significant change in genome integrity, chromatin architecture and gene expression pattern. This phenomenon is called "*epigenetic drift*".

However, to date the relationship between aging and epigenetic drift is controversial. Indeed, is not yet unclear whether changes in the activity of epigenetic enzyme (and thus the epigenome) influence the expression of critical longevity genes or whether alterations in the longevity genes are able to drive large-scale epigenetic changes in the genome.

As shown above, the impact of the genome on aging is well established and seems to account for only 25% of an individual's lifespan, while current studies indicate that epigenetic changes comprise a significant component of aging (Jones *et al.*, - 2015).
One of the well-studied epigenetic mechanisms in relation to aging is DNA methylation.

In general, aging-associated changes in DNA methylation include global hypomethylation and region-specific hypermethylation (Xiao *et al.*, - 2016). DNA hypomethylation also takes place in transposable DNA repetitive elements including Alu sequence and long interspersed nuclear elements (LINE-1), resulting in increased transposition activity and genomic instability (Wilson *et al.*, 2007).

On the other hand, promoter hyper-methylation affects the expression of certain transcription regulatory genes (Gentilini, *et al.,* 2013), apoptotic genes, development or differentiation regulatory genes. Promoter hyper-methylation has been also observed in several tumour suppressor genes such as *CDKN2A*, *LOX*, *RUNX3*, and *TIG1* (Waki *et al*., 2003) as well as, on estrogen receptor (*ER*) and insulin-like growth factor II (*IGF2*).

To date, evidence has shown that reduced global methylation can be caused by the different regulation in the expression of DNMTs as reported by Casillas and colleagues in 2003 where DNMT1 (the predominant maintenance methyltrasferase) expression was reduced contributing to the reduction of genomic methylation of cells. On the other hands, DNMT3b expression steadily increased with aging and has been proposed to be responsible of a paradoxical sporadic gene hypermethylation in aging cells. In accordance with age-dependent hypomethylation, several age-related diseases, including neurodegenerative disease, cardiovascular disease, and cancer, show close association with marked global methylation decrease (Baccarelli *et al.,* 2013 – Choulirias *et al.,* 2013).

In the last few years, the role of histone modifications in aging has emerged, providing insights into epigenetic mechanisms of aging and lifespan regulation. The H3K4me3 histone modification, known to be a marker of active gene expression, is intimately involved in aging across model organisms. In *C. elegans*, deletion of the genes encoding any of the three Trithorax group proteins (WDR-5, SET-2 and ASH-2), that facilitate H3K4 trimethylation, results in decreased global levels of H3K4me3 and increased lifespan (Greer *et al.,* 2010). Similarly, in flies, deletion or RNA interference (RNAi) knockdown of drosophila little imaginal discs (Lid), a recently described histone demethylases, resulted in increased levels of H3K4me3 and reduced lifespan (Li *et al.,* 2010).

Lifespan regulation is influenced also by H3K27 methylation or acetylation, but in this case it seems to operate quite differently in different species.

In flies, loss of function mutations in genes encoding components of Polycomb Repressive Complex 2 (PRC2) - an H3K27me3 specific methyltransferase complex - result in decreased levels of H3K27me3 and increased lifespan (Siebold *et al.,* 2010). In worms, however, the opposite effects are observed (increased levels of H3K27me3 through knockdown of its demethylase UTX-1) and are associated with increased lifespan in an insulin dependent manner (Maures *et al.,* 2011).

In 2013 Liu and colleagues have shown that the repressive mark H3K27me3 increases in aged muscle quiescent stem cells and thus suggested to suppress stem functions in mice, while accumulation of H3K4me3 has also been observed in aged hematopoietic and muscle stem cells (Sun *et al.,* 2014). In the mouse liver, nucleosome occupancy (Bochkis *et al.,* 2014) and the overall content of different histone marks (Kawakami *et al,* 2009) were reported to change with age but genome-wide profiles of histone modifications during aging are not available. In humans, H3K4me3 distribution in prefrontal neurons from 11 individuals was the only genome-wide histone mark study reported throughout aging up to now.

7. Epigenetic, aging and caloric restriction

Environmental factors such as exercise, circadian rhythms and even sexual stimuli are shown to influence gene expression and longevity in different organisms. Nutrient availability and diet are to date the most thoroughly studied environmental factor to affect longevity; diet is known to significantly impact on the epigenome. This has brought the scientific community to focalize the attention on nutrition, to determine whether diet can affect chromatin structure and gene transcription through epigenetic changes. The viable yellow agouti mouse model (A^{vy}/a), where coat's colour variation and obesity is correlated to epigenetic marks established in early development, provided the earliest model for studying epigenetic inheritance in mammals. Wild-type murine Agouti gene encodes for agouti signalling protein (ASIP), a paracrine signalling molecule, whose binding to the melanocortin-1 receptor (MCR-1) prevents α-melanocyte-stimulating hormone signalling, thereby down-regulating synthesis of brown/black pigments (eumelanin) and increasing synthesis of yellow/red (pheomelanin) pigments during the mid-stage of the hair growth cycle.

In A^{vy}/a mice is present an inserted retrotransposon (an *intracisternal A particle* or IAP) is present upstream of the customary transcriptional start site of the wild type A gene. This insertion contains a cryptic promoter in its 5' long-term repeat (LTR) region, capable of driving agouti gene expression and overriding the control mechanisms that normally limit the production of ASIP to certain stages of the hair follicle cycle. In such mice, the fur is completely brown (pseudoagouti) when the cryptic promoter within the IAP is silent, but yellow when the promoter is fully active. This range of phenotypes (Fig. 13) is correlated with the degree of cytosine methylation of the 5′ -LTR of the IAP, with CpG island

hypermethylation associated with the pseudoagouti phenotype and hypomethylation with yellow fur and broad ectopic expression (Morgan *et al*., 1999). Moreover, A^{vy} mice are larger, obese, hyperinsulinemic, more susceptible to cancer and, on average, shorter lived than their non-yellow siblings. This phenotype is due to ectopic expression of ASIP outside hair follicle, most critically in the hypothalamus, where it antagonizes the melanocortin 4 receptor (MC4R).

Waterland and Jirtle (2003) have showed that diet methy-donor supplementation such as folic acid, vitamin B_{12}, choline, and betaine alter the phenotype of their A^{vy}/a offspring via increased CpG methylation at the A^{vy} locus. The observation that the coat colour of these mice and their associated degree of metabolic disease, can be modulated by diet provided to the dam during early development of her offspring, has also raised considerable medical interest, as it is becoming clear that both the quantity and quality of the food that a pregnant woman consumes during her pregnancy can either enhance or reduce the risks of her infant to develop diseases in adult life.

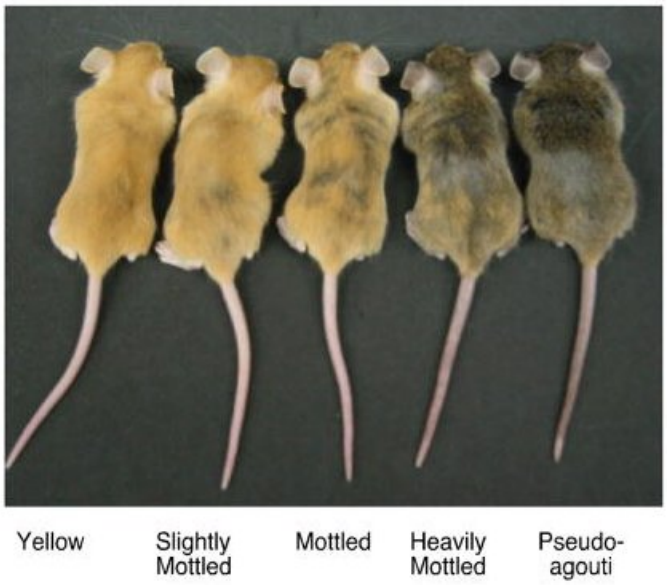


Figure 13. The viable yellow agouti (A^{vy}) mouse model. Genetically identical adult viable

yellow agouti (*Avy*) mice representing the five coat color phenotypes. Yellow mice are hypomethylated at the transposable element upstream of the *Agouti* gene allowing maximal ectopic expression, whereas hypermethylation of this site silences ectopic agouti expression in the pseudoagouti animals. Mice that are predominately yellow are also clearly more obese than brown mice.

Other evidences come from by epidemiology studies. Heijmans and colleagues studied individuals exposed to the "Dutch Hunger Winter", food rationing period in German-occupied Netherlands between '45 and '46, and have demonstrated, studying the descendants, that those who were exposed to famine in utero had very different methylation patterns in genes involved in growth and metabolic disease compared with controls. The studies reported a minor, but significant, decrease in DNA methylation at a differentially methylated region in the imprinted insulin-like growth factor 2 (*IGF2*) gene (Heijmans *et al*., 2008).

Further research identified persistent, small alterations in DNA methylation at other imprinted genes such as insulin (INS), guanine nucleotide binding protein a-stimulating (GNAS) and maternally expressed gene 3 (MEG3) (Tobi *et al*., 2009). Such alterations were also found at some loci that are involved in growth and metabolic disease, such as interleukin 10 (IL10), leptin (LEP) and ATP-binding cassette A1 (ABCA1). Moreover, our methylome can be influenced also by a modulation of caloric intake.

Different studies have shown that caloric restriction defined as moderate (normally, 20-40%) reduction in caloric intake - as compared with an *ad libitum* diet - without compromising the maintenance of all essential nutrients, can influence DNMT activity and lifespan.

7.1 Caloric restriction

The remarkable effects of caloric restriction (CR) on aging have been first described in 1935 in experimental animal models by McCay and colleagues. Since then, numerous research findings have revealed effects of CR on lifespan among diverse species, including yeast, worms, flies, fish and even mammals. CR has also been shown to delay a wide range of aging-associated diseases, such as cancer, diabetes, atherosclerosis, cardiovascular diseases and neurodegenerative diseases in higher mammals, such as nonhuman primates and humans (Omodei and Fontana 2011). Further evidence coming from epidemiological studies conducted in Japan found that urban Japanese peoples consumed more calories and have a higher incidence of cancer than peoples of rural Okinawa, where a regime of CR was adopted.

Sirtuins are probably the best-studied family of enzymes implicated in changing of the epigenome in response to caloric restriction.

Sirtuins are present in a variety of organisms, from yeast to mammals, and have been involved in the organization of global chromatin structure and in the epigenetic regulation of specific genes. Activation of sirtuins is observed in individuals subjected to CR, in relation with the elevated NAD^+ levels induced by the intervention (Kane and Sinclair – 2018). Among sirtuin, *SIRT1* has been the most extensively studied. More than a dozen substrates have already been described for *SIRT1* including members of the *FOXO* family (mammalian homologs of *DAF-16* of *C. elegans*), transcription factors that regulate the expression of genes involved in cell growth and in other cellular processes. In mammals these transcription factors, undergoing post-translational modifications, such as acetylation, de-acetylation, methylation, phosphorylation (Webb and Brunet - 2014) and can access the nucleus regulating the transcription of

genes implicated in various cellular processes, such as cell cycle arrest, resistance to oxidative stress and apoptosis. Kops *et al*., 2002 suppose that *FOXO3A*, by increasing the antioxidant capacity of the cell through the expression of MnSOD and catalase (Kops *et al*., 2002), can influence longevity avoiding the accumulation of ROS (reactive oxygen species) or highly reactive molecules capable of causing damage to biological structures.

Moreover, SIRT1 has been shown to deacetylate FOXO3A allowing this protein to remain inside the nucleus and regulate the transcription of genes that permits to the stem cell to maintain an undifferentiated and quiescent state (phase G0), a fundamental condition for the own capacity for self-regeneration (Miyamoto *et al*., 2007). Vaquero and colleagues in 2007 have shown that SIRT1 could suppress the expression of inflammatory genes by enhancing the activities of histone methyltrasferase SUV39H1, resulting in increased levels of H3K9me3 (Vaquero *et al*., 2007).

8. Aims of the project

PAT-ChIP procedure is a powerful approach to investigate the epigenomic landscape in FFPE tissues although some limitations are still present. During my Ph.D. activity, I focused my research activity mainly on two objectives:

- *The improvement of PAT-ChIP technology to allow epigenomic studies of "critical" biological models;*
- *The in vivo investigation of the epigenomic basis of aging focusing on the possible role of caloric restriction in the slowing down / revertion of the phenomenon.*

The first part of this project provides a detailed overview of all the procedure used to improve the original PAT-ChIP protocl while the second part describes the global characterization of aging-associated epigenetic landscapes and the effects induced by CR.

9. Epigenomic profiling of archived FFPE tissues by enhanced PAT-ChIP (EPAT-ChIP) technology

As already mentioned, chromatin immunoprecipitation coupled with massive parallel sequencing approach is a powerful technique to investigate the epigenomic landscape with an enormous potential to extend our knowledge on the influence of epigenetic alterations in several biological processes. Thanks to the introduction of PAT-ChIP, epigenomic studies have been extended to innumerable formalin-fixed paraffin embedded (FFPE) tissues stored in hospital archives and tissue bank around the world. However, due to the lack of standardization in FFPE tissues production, extensive tissues fixation is a frequent technical problem that could hinder the application of PAT-ChIP. In the light of these evidence, I have investigated the possibility to improve chromatin extraction efficiency from FFPE tissues, to facilitate the combination of the original PAT-ChIP protocol with NGS technology, allowing genome-wide studies also in clinical "complex" archival samples.

First attempts allowed to identify the best condition to increase the efficiency of chromatin extraction using normal colon tissues (chosen due to the availability of high quantities of this human material as scrap from colorectal surgeries). Once identified the best condition, antigen integrity was evaluated through immunofluorescence approach and then the new procedure, called Enhanced Pathology-Tissue Chromatin Immunoprecipitation (EPAT-ChIP), was applied on real archive breast human FFPE tissues.

9.1 Materials and methods

Materials

- 37% Formaldehyde solution (Sigma)
- 0.05 M Acetate Buffer
- Histolemon solution (Carlo Erba, Milan, Italy)
- Ethanol (VWR)
- Lysis buffer (10 mM Tris-HCl pH 7.4, 0.15 M NaCl, 3 mM $CaCl_2$, 2 mM $MgCl_2$, 0.5% Tween20, 1 mM PMSF, and 10 μg/mL RNase A—Roche, Mannheim, Germany)
- Fetal Bovine Serum (FBS)
- Digestion buffer (50 mM Tris-HCl, pH 7.4, 0.32 M sucrose, 4 mM $MgCl_2$, 1 mM $CaCl_2$, and 0.1 mM PMSF)
- Micrococcal nuclease (N.70196Y; USB)
- Extraction buffer (10 mM Tris-HCl pH 7.4, 0.15 M NaCl, 3 mM $CaCl_2$, 2 mM $MgCl_2$, 0.1% SDS)
- Sodium citrate buffer (10 mM sodium citrate, 0.05% Tween20, pH 6.0)
- Proteinase K (Roche, cat. no. 1092766)
- SYBR Gold (Invitrogen, Eugene, OR, USA)
- Incubation buffer (30 mM Tris-HCl pH 7.4, 50 mM NaCl, 5 mM Na_2EDTA, and 0.1 mM PMSF)
- Anti-H3K4me3 antibody (39159, Lot. 01609004; Active Motif, Carlsbad, CA, USA)
- Anti-H3K27me3 antibody (07-449, Lot. JBC1873477; Millipore, Temecula, CA, USA)

- Anti-H3K27ac antibody (ab 4729, Lot. GR254707-1; Abcam, Cambridge, UK)
- 50% v/v slurry rec-Protein G-Sepharose 4B Conjugate (Invitrogen, Frederick, MD, USA)
- Bovine Serum Albumin (BSA) 50 mg/ml (MilliporeSigma Calbiochem, cat. No.12659)
- Washing buffer A (20 mM Tris-HCl, pH 7.4, 1% (v/v) Triton X-100, 5 mM EDTA, 50 mM NaCl, 0.1 mM PMSF),
- Washing buffer B (50 mM Tris-HCl, pH 7.4, 1% (v/v) Triton X-100, 10 mM EDTA, 100 mM NaCl)
- Washing buffer C (50 mM Tris-HCl, pH 7.4, 1% (v/v) Triton X-100, 10 mM EDTA, 150 mM NaCl).
- Elution buffer (1X TE, 1% (w/v) SDS)
- dsDNA HS Assay Kit (Invitrogen, Eugene, OR, USA)
- Fluorochrome-conjugated secondary antibody (Donkey anti-rabbit IgG highly cross-adsorbed secondary antibody, Alexa Fluor 488— ThermoFisher Scientific, San Jose, CA, USA)
- DAPI nucleic acid stain (Invitrogen, cat. no. D3571)
- TopVision Agarose (Thermoscientific, cat. no R0492)
- SyBR® Gold nucleic acid gel stain (Invitrogen, cat. no. S11494)

Preparation of FFPE tissues

Specimens of normal colon tissue were obtained from a patient affected by colorectal cancer who underwent curative surgical resection at the Ospedale di Circolo, ASST Sette Laghi (Varese, Italy). After collection (10 cm distant from the tumor), the tissue was divided in three pieces of similar size and fixed for 24 h, 48 h, or 72 h. All tissues were fixed in

neutral-buffered formalin (4% w/v formaldehyde and 0.05 M acetate buffer) and routinely processed in paraffin wax using the automated tissue processor Donatello (Diapath, Bergamo, Italy).

Different attempts to increase chromatin extraction efficiency

For each sample, four FFPE tissue sections of 10 μm thickness were first deparaffinized by five sequential incubations, of 10 min each, in 1 ml of Histolemon at room temperature. When not specified, all centrifugations were performed at 17,860×*g* for 3 min at 4°C (in a refrigerated centrifuge). Samples were rehydrated by decreasing concentrations of ethanol starting from 100% through to 95%, 70%, 50%, 20% and at the end, water (10 min at room temperature for each step in 1 ml of each ethanol solution). Sections were then resuspended in 0.5 ml of lysis buffer and incubated 30 min at room temperature on a rotating platform. After centrifugation at 17,860×*g* for 3 min at 4°C, samples were resuspended in 300 µL of digestion buffer, partially fragmented through mild sonication (three times for 30 s - interrupted by 60 s pauses with an amplitude of 40%), using EpiShear sonicator mounting a 3.2 mm probe (Active Motif, Carlsbad, CA,USA). After then, different approachs were used to increase the chromatin extraction efficiency:

Modulation of sonication condition.

Once pre-fragmentation step are performed, three samples has been subjected to a controlled digestion for 1 min at 37 °C with micrococcal nuclease (MNase) followed by centrifugation at 17,860×*g* for 3 min at 4°C. The pellets were resuspended in 400 µL of extraction buffer and chromatin was extracted using a canonical sonicator and testing two

additional sonication conditions respect to those described in the original PAT-ChIP protocol (18 pulses of sonication for 5 s at 85% of amplitude).

Condition A: 54 pulses of sonication for 5 s at 75% of amplitude.
Condition B: 54 pulses of sonication for 5 s at 65% of amplitude).

Heat-mediated limited reversal of crosslinking (LRC)

The fourth sample, named *LRC* was centrifugated at 17,860×*g* for 3 min at + 4 °C after the pre-fragmentation step and the pellet was resuspended in 1 ml of sodium citrate buffer and incubated for 1 h at 80°C. After incubation, the sample was centrifugated and resuspended in 400 µL of extraction buffer and chromatin extraction was performed by 3 pulses of sonication for 30 s at 40% of amplitude.

After chromatin extraction, all samples have been cleared by centrifugation (9500×g for 5 min at room temperature), supernatants containing chromatin were saved and an aliquot of 40 μl (corresponding to the 10% of total isolated chromatin, named "input") was subjected to a complete de-crosslinking through overnight (16 h) incubation at 65°C in the presence of 0.2 M NaCl, followed by digestion with 0.1 mg/ml proteinase K (3 h at 45°C). DNA purification was carried out using the PCR Purification Kit (Qiagen, Hilden, Germany) following manufacturer's instructions and DNA was fluorimetrically quantified by Qubit (Invitrogen, Eugene, OR, USA) using the dsDNA HS Assay Kit (Invitrogen, Eugene, OR, USA) to estimate the total amount of chromatin present in the input fraction. Chromatin fragmentation was checked by

electrophoretic separation on a 1.3% agarose gel. DNA was stained with SYBR Gold stain.

Chromatin immunoprecipitation

Chromatin immunoprecipitation and DNA isolation were performed following the procedures already described by Fanelli *et al.,* 2011. Isolated chromatin was incubated in incubation buffer for 16 h at 4°C in a rotating platform, with follow antibodies: anti-H3K4me3 (2 ul of whole serum extract), anti-H3K27me3 (4 μg of protein A-purified antibody), anti-H3K27ac (1,25 μg of immunogen affinity-purified antibody). Forty microliters of 50% vol/vol slurry rec-Protein G-Sepharose 4B Conjugate (preincubated 16 h at 4 °C with 1 mg/mL of BSA in incubation buffer) were added to each ChIP assay and incubated for 3 h at 4°C. After centrifugation at 2,000×*g* for 5 min at 4°C, pellets were sequentially washed with 10 mL of washing buffer A, 10 mL of washing buffer B and 10 mL of washing buffer C. Elution was carried out by incubation for 40 min at RT, using a rotating platform, in 300 μL of elution buffer. After centrifugation (1270×*g* for 2 min at 4°C), the supernatant was saved and the elution repeated with only 50 μl of elution buffer as described above (and by vortexing 10 s at maximum speed) to obtain a final volume of 0.35 ml (named the "bound" fraction).

DNA isolation and locus-specific analysis by real-time PCR

Bound fractions, and an amount corresponding to 5% of input chromatin, were de-crosslinked through an overnight incubation at 65 °C with 0.2 M NaCl, followed by digestion with 0.1 mg/mL proteinase K (3 h at 45°C). DNA purification was carried out using the PCR Purification Kit

(Qiagen, Hilden, Germany) following manufacturer's instructions and DNA was fluorimetrically quantified by Qubit (Invitrogen, Eugene, OR, USA) using the dsDNA HS Assay Kit (Invitrogen, Eugene, OR, USA). The specificity of the immunoselection was preliminarily analyzed by real-time quantitative PCR (qPCR) using the Fast Start SYBR Green Master Mix and the Rotor-Gene 6000 robocycler (Corbett Life Science, Sydney, Australia). Amplifications of bound fraction and input were carried out using a set of primers pairs specific for known positive (enriched) and negative (not enriched) control regions as previously described (Amatori *et al*., 2014).

Library preparation and sequencing

Library preparation and sequencing were performed at the European Institute of Oncology (IEO - Milan) NGS facility. For this reason, only a brief description of the procedure used is reported. Bound fractions (from 2 up to 10 ng) were used for blunt-ending and phosphorylation reaction. A single 'A' nucleotide was added (A-tailing reaction) to the 3' ends of the fragments to allow the match with adapter that has a single-base 'T' overhangs. The ligation products were purified and accurately size-selected by agencourt AMPure XP beads. Purified DNA was PCR-amplified to enrich for fragments that have adapters on both ends. All step described were performed on the automation instrument Biomeck FX by Beckman Coulter. The final purified products were quantified and size checked on a Bioanalyzer 2100.

Libraries with distinct adapter indexes were multiplexed and after cluster generation on FlowCell were sequenced for 50 bases in the single read mode on a HiSeq 2000 sequencer.

Pipeline of ChIP-Seq analysis

Reads were aligned to hg19 using "bowtie" (version 0.6.2-r126). Unmapped reads, reads with a MAPQ smaller than 1, and duplicate reads were removed using "samtools" (version 0.1.18). Mapped sequence reads were extended to 200 bp, which was the estimated mean insert size targeted in the size selection step when preparing the libraries, using deepTools (version 2.5.4). Detection of peaks was performed using MACS2 software from Galaxy browser using the bdgpeakcall function to call narrow peaks from H3K4me3 and H3K27ac tracks and MACS2 bdgbroadcall function for broad peaks calling from H3K27me3. Intersections between genomic regions were performed using the specific function on Galaxy browser. The R/Bioconductor package ChIPseeker was used to annotate the genomic features of peaks, while data sets and peaks were visualized on UCSC Genome Browser from where snapshots were produced.

Immunofuorescence

HeLa cells were harvested and fixed with 1% FA in PBS for 10 min at 37°C or, to mimic an over-fixed condition, for 4 h with 4% of FA at 37°C. Immunofluorescence was performed following the same conditions used in the PAT-ChIP assay (buffers, incubation timing, and temperature) as previously described.

Briefly, after permeabilization with permeabilizing buffer, cells where spotted by citospin, while LRC samples were subjected to limited reversal of crosslinking step by heating 1 h at 80°C in sodium citrate buffer and then spotted. Cells were blocked with FBS (30 min at room temperature) and then incubated in a humidified chamber with the same

antibodies (anti-H3K4me3, anti-H3K27ac, and anti-H3K27me3), buffers and concentrations used in ChIP experiments. After washing with ice-cold washing buffer A, washing buffer B, and washing buffer C, cells were incubated with a fluorochrome-conjugated secondary antibody (Alexa Fluor 488), washed with PBS, and counterstained with DAPI as previously described by Fanelli M. and colleagues in 2010. Fluorescence signal was acquired using an Olympus BX51 microscope equipped with an Olympus F-View II digital camera and AnalySIS software (Soft Imaging System, GmbH).

9.2 Results

Chromatin extraction

As mentioned before, we have attempted to increase the efficiency of chromatin extraction of the Std PAT-ChIP procedure and to this end, using specimens of human normal colon tissue fixed for different time (24 h, 48 h, or 72 h), were used:

Summarizing, we have:

i. Modified sonication steps in terms of times and amplitudes.

ii. Performed a heat-mediated limited reversal of crosslinking (LRC) of the tissue with the intent to reduce the chromatin complexity and increase the isolation efficiency.

Different LRC conditions were tested, varying in temperature (from 65 to 95°C), time of incubation (from 10 min to 16 h), as well as buffer pH (from 6.0 to 9.0). At the end, the condition of 80°C for 1 h of incubation in sodium citrate buffer pH 6.0 was selected as the most efficient step to improve chromatin extraction (data not-show).

Once extracted, 40 ul of extracted chromatin was de-crosslinked and DNA was purified and fluorimetrically quantified by Qubit. Results are reported in term of amount of soluble chromatin extracted (Fig. 14a) and in term of extraction efficiency (Fig. 14b). As shown, modulation of sonication conditions did not allow a significant increase of extracted chromatin, while LRC significantly increased the efficiency of chromatin extraction from the human colon specimens in all the conditions of FA fixation tested

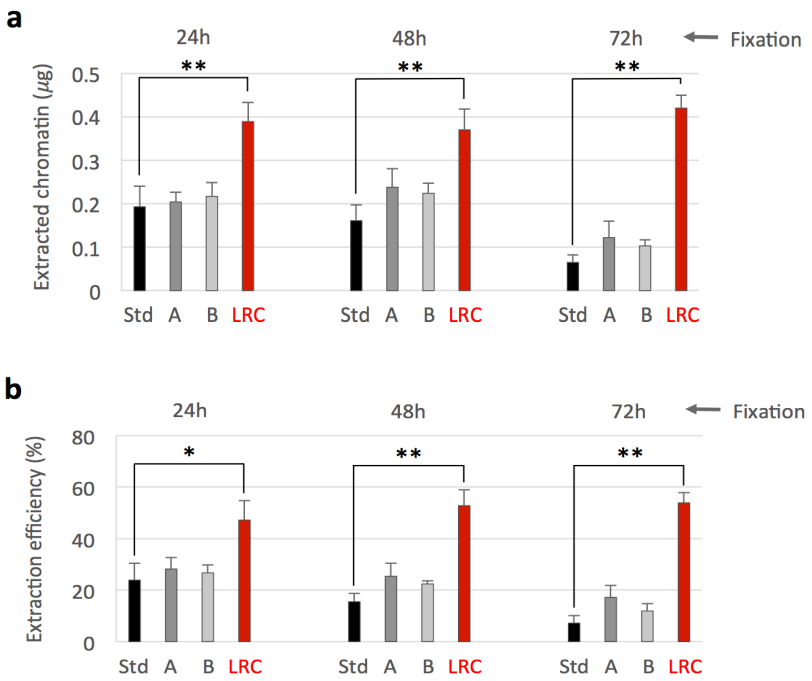


Figure 14. Different conditions of chromatin extraction. Normal colon FFPE tissues fixed at the times reported has been used to test different conditions of chromatin extraction. (**a**) Extracted chromatin was fluorimetrically quantified after chromatin de-crosslinking and DNA purification. (**b**) Extraction efficiency was calculated considering the amount of DNA of extracted chromatin compared to the total DNA present in the sample. **Std**: standard PAT-ChIP, 18 pulses of sonication of 5 s at 85% of amplitude; **A**: 54 pulses of sonication of 5 s at 75% of amplitude; **B**: 54 pulses of sonication of 5 s at 65% of amplitude; **LRC**: condition in which the sample was subjected to limited reversal of crosslinking, 3 pulses of sonication of 30 s at 40% of amplitude. *$P < 0.05$ with respect to standard condition for each time of fixation by one-way ANOVA with Tukey's HSD. **$P < 0.01$ with respect to standard condition for each time of fixation by one-way ANOVA with Tukey's HSD. All the experiments were conducted in triplicates.

For each condition tested, an evaluation of chromatin fragmentation by electrophoretic separation on 1.3% agarose gel electrophoresis (AGE) followed by SYBR Gold staining was performed. As show below, in all the conditions tested the levels of fragmentation are comparable to the one obtained using the Std conditions of sonication (Fig.15).

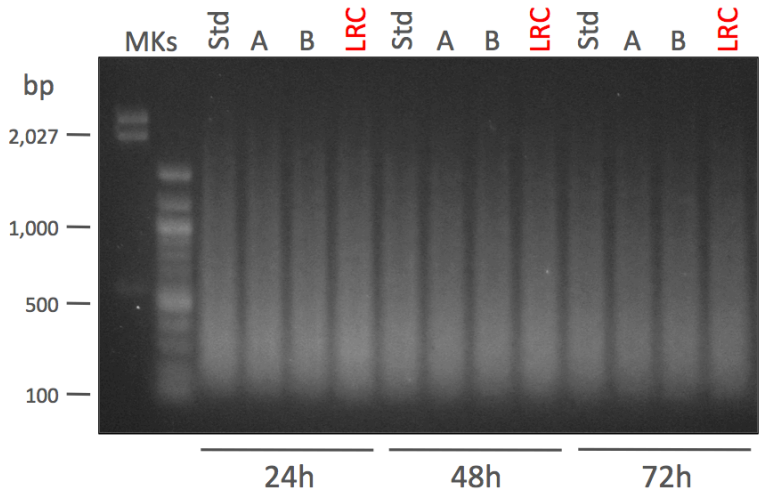


Figure 15. Evaluation of chromatin fragmentation by electrophoretic separation. Fragmentation level of extracted chromatin using different condition was evaluated on 1.3% agarose gel electrophoresis (AGE) followed by SYBR Gold staining.

Immunofluorescence

Immunofluorescence approach on HeLa cells has been exploited to assess if LRC step could interfere with the recognition of the epitope by the antibody. Immunofluorescence was conducted using HeLa cells fixed with 1% FA for 10 min at 37°C and using same condition used in PAT-ChIP protocol (buffer, temperatures and time of incubation) with antibodies previous mentioned. The signals of LRC samples (incubated 1 h at 80°C in sodium citrate buffer pH 6.0) were compared with control samples (not LRC-treated). DAPI staining of nuclei (blue) is also shown (Fig. 16).

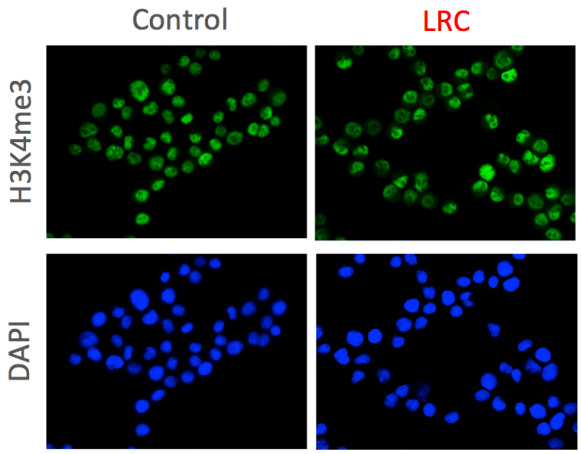


Figure 16. Evaluation of LRC treatment on epitope integrity. Compatibility between LRC and H3K4me3 immunoselection was evaluated in HeLa cells through an immunofluorescence approach. Cells were subjected to formaldehyde fixation and treated with LRC (right panel) and compared with untreated (left panel). Cells were stained by immunofluorescence with anti-H3K4me3 antibody (green, upper panels) following the same procedure described for the PAT-ChIP assay (buffers, timing, and temperature of incubations) and with DAPI to label nuclei (blue, lower panels).

Chromatin immunoprecipitation using colon FFPE tissues fixed for different time

Once extracted using the different conditions described above, chromatin was checked for immunoselection by the anti-H3K4me3 antibody. This HPTM was chosen for different reasons: *(i)* it represents a good "stress-test" for the procedure since immunoselection against H3K4me3 normally produces poor amounts of final DNA, especially when low quantities of input chromatin are used, increasing the chances of failure in the generation of NGS libraries; *(ii)* its close association with gene promoters and its narrow distribution allows a better measure of ChIP-Seq specificity and resolution.

Comparable efficiency of immunoselection (percentages of enrichment ranging between 1.12% and 1.92% - Fig. 17b) was obtained using the different conditions. However, the quantity of final DNA (Fig. 17a) was significantly greater in samples subjected to limited reversal of crosslinking (LRC) in consequence of the usage of higher amounts of input chromatin.

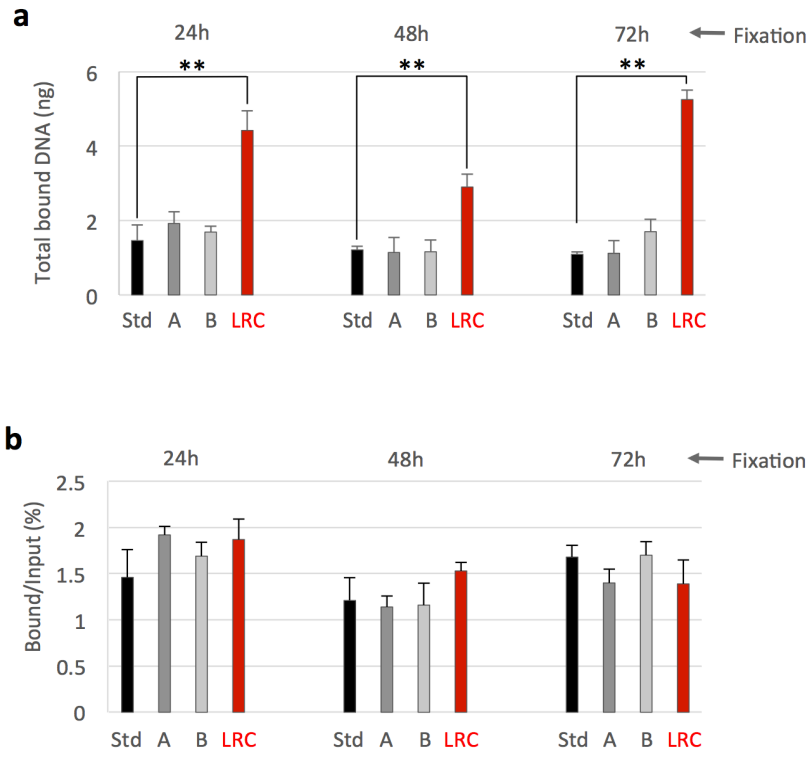


Figure 17. Quantitative analysis of immunoselected DNA. (**a**) Chromatin from normal colon FFPE samples at different times of fixation, extracted following the different strategies described above, was immunoprecipitated with an anti-H3K4me3 antibody. After immunoselection, immunoselected chromatin was de-crosslinked and the DNA purified and fluorimetrically quantified. (**b**) Input fractions were also purified and the percentage of enrichment by the antibody compared to the input was calculated. $^{**}P < 0.01$ with respect to standard condition for each time of fixation by one-way ANOVA with Tukey's HSD. All the experiments were conducted in triplicates

Fig. 18 shows the specificity of the immunoselection evaluated by real-time qPCR by using primer pairs able to amplify the promoter regions of two genes known to be ubiquitously expressed and enriched of H3K4me3 (vinculin —VCL and glyceraldehyde-3-phosphatase dehydrogenase —GAPDH) and two primers pairs able to amplify two genomic loci know to be silent andnot-enriched (hyaluronan and proteoglycan link protein 1—HAPLN1 and collagen, type II, alpha 1—COL2A1).

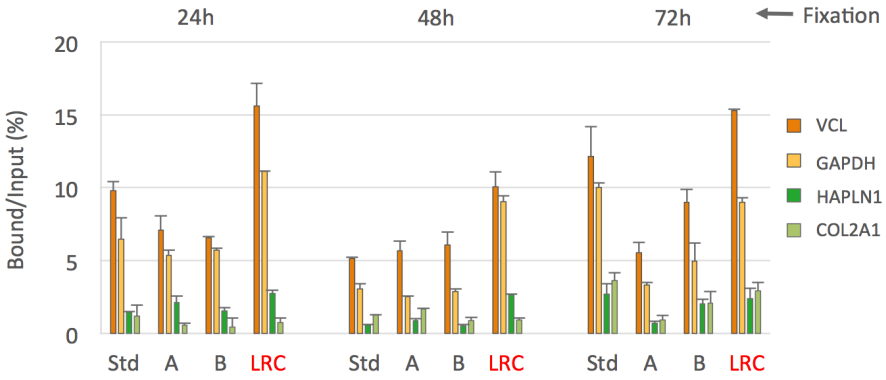


Figure 18. Real-time qPCR analysis. Transcriptionally active (VCL and GAPDH) and inactive (HAPLN1 and COL2A1) promoter regions were amplified by real-time qPCR to evaluate the specificity of the immunoselection. H3K4me3 enrichments are expressed as percentage of bound compared to the input.

Validation of the EPAT-ChIP protocol using real archival FFPE samples

The new PAT-ChIP procedure, with the introduction of LRC, was named Enhanced PAT-ChIP (EPAT-ChIP). EPAT-ChIP was then applied for validation using an archival invasive breast carcinoma (IBC) FFPE sample.

We have used 4 sections of 10 μm thickness each, with average area of 4 cm^2, for both standard PAT-ChIP and EPAT-ChIP protocols.

Also in this case, extracted chromatin from the sample subjected to LRC is higher then obtained with PAT-ChIP (Fig. 19a) while the average size in both case is compatible with ChIP-assay (Fig 19b).

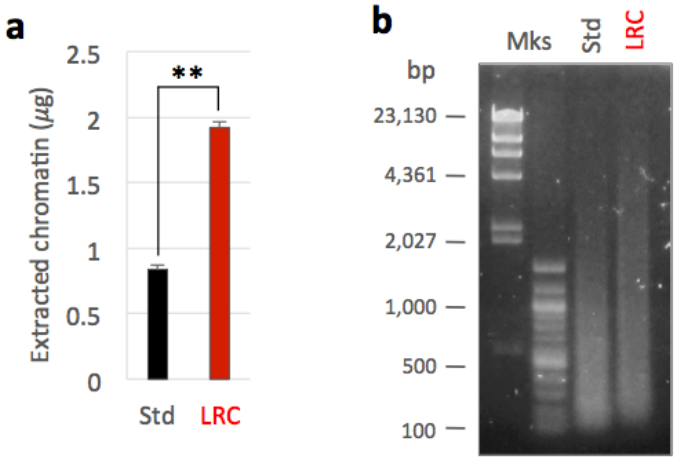


Figure 19. Application and validation of EPAT-ChIP protocol. Chromatin was extracted from an archival invasive breast carcinoma sample by both the standard PAT-ChIP procedure (Std) and the new LRC-based procedure (LRC). (**a**) The amount of extracted chromatin was estimated by fluorimetric quantitation of purified DNA after complete de-crosslinking. (**b**) Chromatin fragmentation was evaluated by electrophoretic separation on 1.3% AGE followed by SYBR Gold staining of purified input DNA.

After immunoselection with an H3K4me3 antibody, the bound fractions were decrosslinked and purified as described above and result show that LRC produced higher amounts of DNA (a mean of 7.6 ng), compared to standard procedure (mean of 1.95 ng — Fig. 20a) while the efficiency of immunoselection was comparable (Fig. 20b).

Even in this case, the enrichment of both active and silent gene promoters was analyzed using real-time qPCR and results are show below (Fig. 20c).

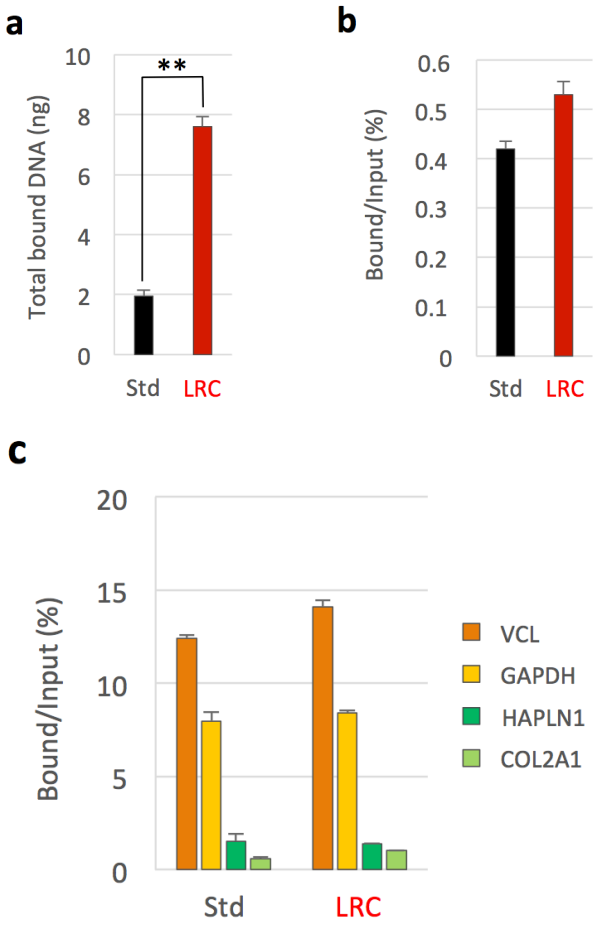


Figure 20. Quantitative and qualitative analysis of H3K4me3 immunoselection. (**a**) Chromatin extracted with both protocol was immunoprecipitated with an anti-H3K4me3 antibody, de-crosslinked, and the DNA purified and quantified. (**b**) Input fractions were also purified and the percentage of enrichment with respect to the input was calculated.

(**c**) Transcriptionally active (VCL and GAPDH) and inactive (HAPLN1 and COL2A1) promoter regions were amplified by real-time qPCR (each sample amplified in triplicate) to evaluate the specificity of the immunoselection. H3K4me3 enrichments are expressed as percentage of bound respect to the input. $^{**}P < 0.01$ with respect to standard condition by Student's *t* test.

EPAT-ChIP coupled with NGS for the epigenomic profiling of archival samples

The bound fractions obtained with both procedures are used for library preparation followed by massive parallel sequencing.

Reads are mapped on hg19 reference genome and 17,041 and 14,041 peaks from the Std PAT-ChIP and EPAT-ChIP data sets, respectively, were identified. ChIPseeker was used to annotate the genomic features of peaks and results show that both data sets have similar distribution of H3K4me3 enrichments (peaks) in correspondence of the promoter of active genes and associated CpG islands. Peaks obtained applying both protocols are mainly located in correspondence of gene promoters (85.95% for Std PAT-ChIP and 89.43% for EPAT-ChIP – Fig. 21a). However, while 94.29% of total promoter peaks from EPAT-ChIP are located within 1 Kb from TSS, only 87.77% peaks from canonical PAT-ChIP are located within the same region, suggesting that EPAT-ChIP signal resolution is higher compared to that obtained by the Std PAT-ChIP protocol (Fig. 21b).

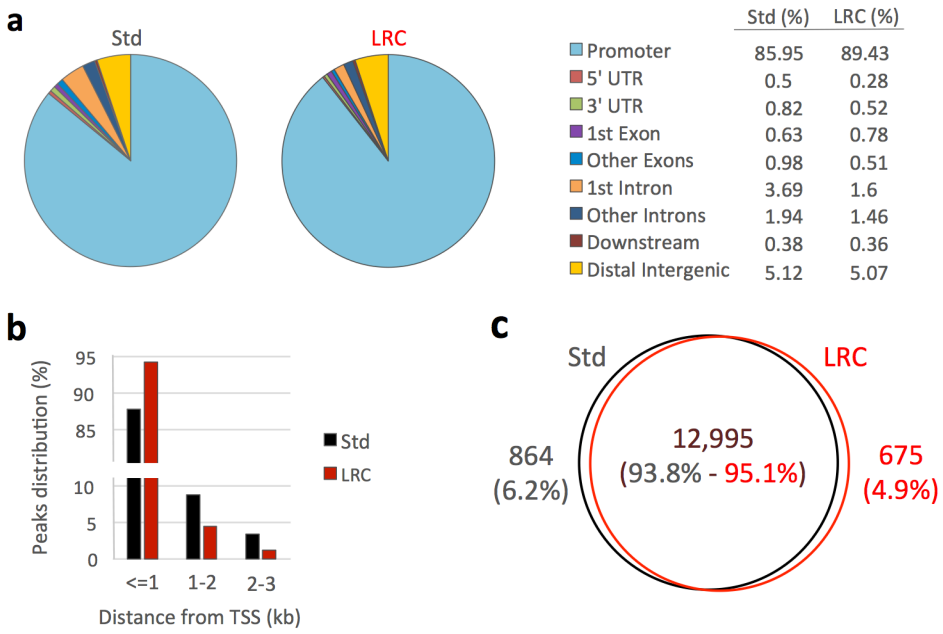


Figure 21. Analysis of global genomic annotation distribution of H3K4me3 by EPAT-ChIP. Immunoselected DNA was subjected to massive parallele sequencing. (**a**) Pie charts depicting the distribution across genomic features with relative percentage values shown on the right. Promoters are defined as – 3 Kb to + 3 Kb relative to the TSS, while Downstream as – 3 Kb relative to the end of 3′ UTR region. (**b**) Distribution of H3K4me3 promoter peaks relative to the TSS. (**c**)Venn diagram showing common and unique peak-containing promoters identified by standard PAT-ChIP and EPAT-ChIP.

The comparison of the identified peaks by the two protocols identified 93.8% and 95.1% of common gene promoters in standard and EPAT-ChIP protocols, respectively (Fig. 21c).

To further validate the obtained results, we have applied the same analytical pipeline on H3K4me3 data set obtained from human mammary epithelial cells (HMEC) available from UCSC Genome Browser (GEO accession number: GSM733712). Interestingly, peaks from the HMEC data set are tightly associated with TSS, showing a distribution comparable to that observed in EPAT-ChIP data set (Fig. 22b).

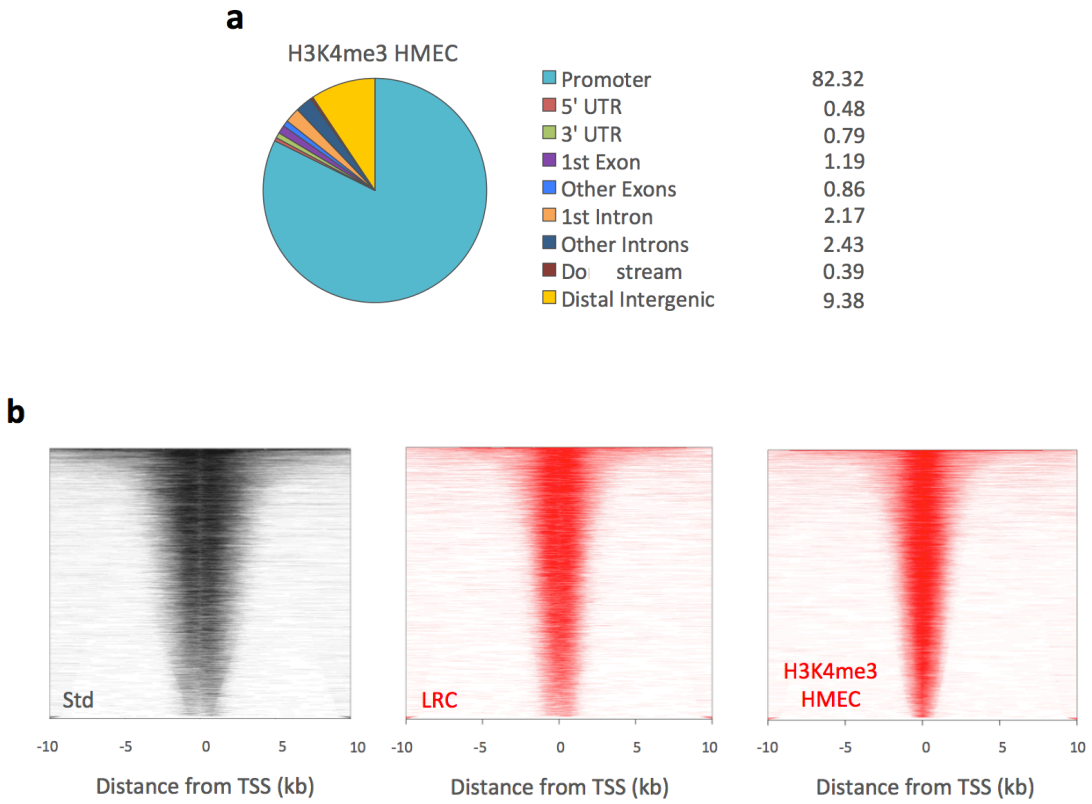


Figure 22. Analysis of H3K4me3 ChIP-Seq data from ENCODE project. H3K4me3 data from human mammary epidermal cells (HMEC) was taken from UCSC Genome Browser and analyzed following the same pipeline used for standard PAT-ChIP and EPAT-ChIP (LRC) data sets. (**a**) Pie charts depicting the distribution of peaks across genomic features with relative percentage values shown on the right. (**b**) Heatmaps illustrating H3K4me3 peak densities from – 10 Kb to + 10 Kb relative to the TSS.

Application of EPAT-ChIP to investigate the genome-wide distribution of other histone marks (H3K27me3 and H3K27ac) in archival samples

Application of EPAT-CHIP was extended to other HPTMs such as H3K27me3 and H3K27ac and, even in this case, both protocols were applied starting from four sections of archival invasive breast carcinoma (IBC) FFPE sample. After extraction, chromatin was immunoselected with H3K27ac and H3K27me3 antibodies and the bound fractions were decrosslinked, purified and analysed as discussed. Real-time qPCR results demonstrate that samples processed using both experimental procedures showed a similar behaviour (Fig. 23), with silent genes exhibiting high enrichments of this histone mark (as expected) compared to active genes (not-enriched for this HPTM).

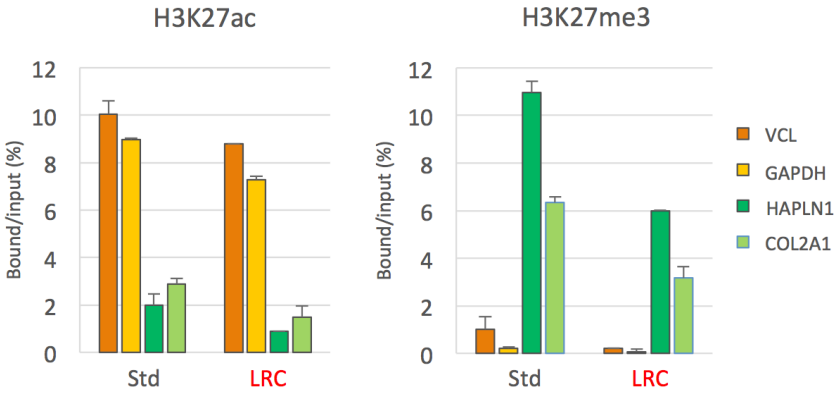


Figure 23. Qualitative analysis of H3K27ac and H3K27me3 immunoselection by real-time qPCR. Transcriptionally active (VCL and GAPDH) and inactive (HAPLN1 and COL2A1) promoter regions were amplified by real-time qPCR (each sample amplified in triplicates) to evaluate the specificity of the immunoselections. Enrichment is expressed as percentage of bound DNA with respect to the input.

Libraries were then successfully produced and sequenced, and the distribution of peaks was consistent between the two techniques and with what has been already described in the literature. H3K27me3 peaks are primarily located in distal intergenic regions (51.93% and 51.42% for Std procedure and EPAT-ChIP, respectively – Fig. 24).

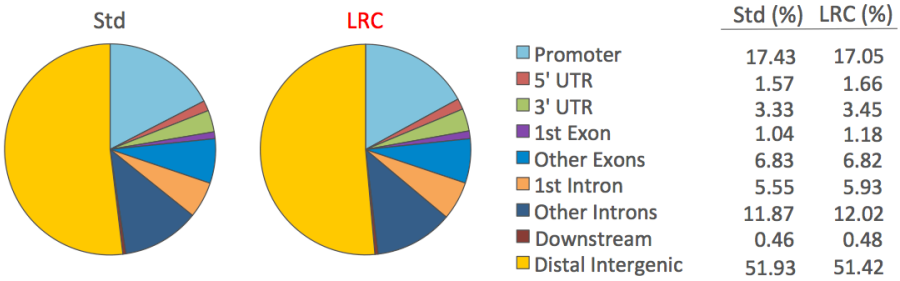


Figure 24. Analysis of global genomic annotation distribution of H3K27me3 obtained with both procedures. Pie charts depicting the distribution of peaks across genomic features with relative percentage values shown on the right. Promoters are defined as – 3 Kb to + 3 Kb relative to the TSS, while Downstream as – 3 Kb relative to the end of 3′ UTR region.

Even in this case, similar results were obtained applying the same analytical pipeline that we used for our data sets to a H3K27me3 data set obtained using HMEC available from UCSC Genome Browser (GEO accession number: GSM733722 – Fig. 25)

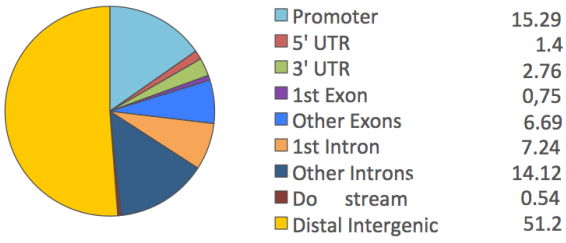


Figure 25. Analysis of global genomic annotation distribution of H3K27me3 ChIP-Seq data from ENCODE project. H3K27me3 data from human mammary epidermal cells (HMEC) was taken from UCSC Genome Browser and analyzed following the same pipeline

used for standard PAT-ChIP and EPAT-ChIP (LRC) data sets. Pie charts depicting the distribution of peaks across genomic features with relative percentage values shown on the right.

Differently from what obtained analyzing H3K27me3, where distribution across genome of peaks obtained with the two protocols is very similar, EPAT-ChIP allows a strongly improvment of the results when H3K27ac was investigated. As shown in Fig. 23 left panel, real-time qPCR analysis of the H3K27ac immunoselected DNA show a similar behaviour between the two protocols, although the genome-wide distribution of peaks is quite different (Fig. 26a).

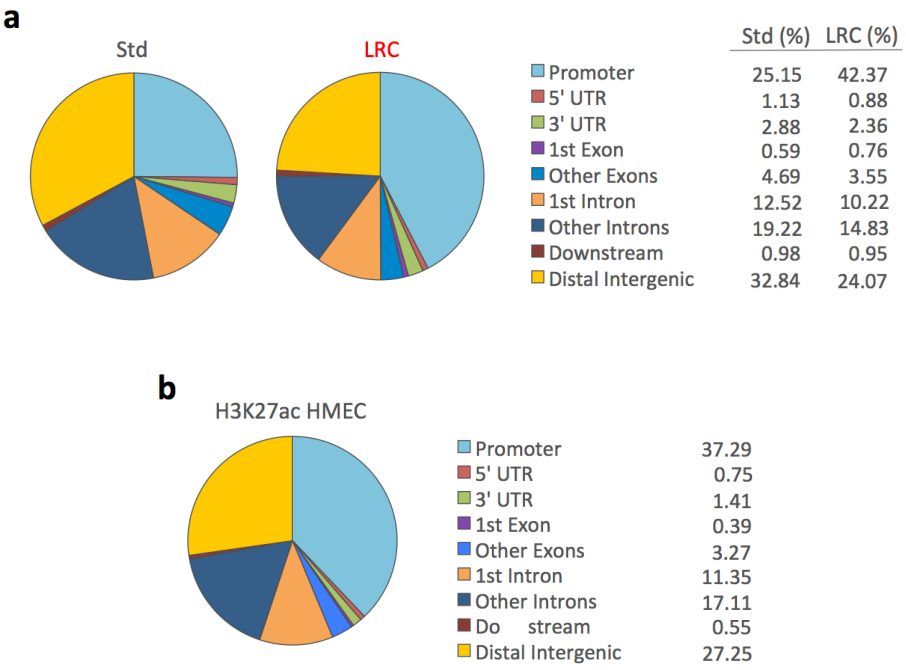


Figure 26. Analysis of global genomic annotation distribution of H3K27ac by Std PAT-ChIP, EPAT-ChIP and ChIP-Seq data from ENCODE project. (**a**) Pie charts depicting the distribution of peaks of H3K27ac across genomic features with relative percentage values shown on the right. (**b**) H3K27ac data from human mammary epidermal cells (HMEC) was taken from UCSC Genome Browser and analyzed following the same pipeline used for standard PAT-ChIP and EPAT-ChIP (LRC) data sets. Promoters are defined as – 3 Kb to + 3 Kb relative to the TSS, while Downstream as – 3 Kb relative to the end of 3′ UTR region.

Peaks obtained with both procedure show a distribution mainly focused on promoters and intergenic regions (Fig. 26a), however EPAT-ChIP was able to strongly improve the quality of H3K27ac profile (Fig. 26a) producing a distribution of peaks comparable to that obtained from a HMEC H3K27ac data set available from the UCSC Genome Browser (GEO accession number: GSM733660 – Fig. 26b).

9.3 Discussion

Formalin-fixation and paraffin embedding has become the standard preservation procedure for diagnostic surgical pathology. However, this method of preservation generates a dense network of crosslinked cellular biomolecules that could hamper the performances of several biological assays including chromatin immunoprecipitation. The introduction of PAT-ChIP technology opened the door to the study of archival FFPE samples, which represent an extraordinary source of epigenomic information, allowing to investigate the epigenetic basis of cancer and other diseases. Although to date the technique has been extensively used by many different groups of research, we found that its performances can be influenced by the crosslinking status of samples making necessary to improve the protocol in order to facilitate epigenomic studies in archival patient samples. The EPAT-ChIP protocol described here is characterized by the introduction of a controlled heat-mediated limited reversal of crosslinking (LRC) able to reduce tissue complexity, previously introduced by FA fixation, and thus to increase chromatin isolation efficiency from FFPE samples.

Taking advance from the large availability of normal human colon specimens, obtained from scrap from colorectal surgeries, different modifications of the original PAT-ChIP protocol have been evaluated. The reduced ethical implications allowed to fix specimens at different times (from 24 to 72 h), producing samples representative of what is normally found in FFPE archives. These tissues were used, first of all, to demonstrate that the extension of FA fixation progressively reduces the amount of chromatin that can be isolated from FFPE. We demonstrated that the introduction of LRC was able not only to overcome the effect of long times of fixation on chromatin extraction efficiency, but also to

increase the amount of chromatin extractable at all the times of FA fixation tested (Fig. 14) and, most importantly, it did not interfere with antigen recognition (Fig. 16).

Through the qualitative and quantitative analysis of bound fractions from anti-H3K4me3 IP (chosen as "stress-test" due its low abundance in the genome), I have demonstrated that the introduction of LRC was compatible with the immunoselection (Fig. 18). Then, an invasive breast carcinoma (IBC) sample was taken as representative of real archival FFPE samples, and processed with PAT-ChIP and the new EPAT-ChIP procedure. Even in this case I demonstrated that the amount of extracted chromatin is increased using EPAT-ChIP (Fig. 19a) and, after H3K4me3 immunoselection, we obtained overlapping results at both locus-specific (Fig. 20c) and genome-wide levels (Fig. 21a) between the two procedures.

These results were further validated applying the same bioinformatics analytical pipeline on H3K4me3 data set from HMEC publicly available from UCSC Genome Browser showing that, despite the different experimental conditions used, the data sets showed a consistent distribution of peaks with respect to genomic features (Fig. 22a).

Further confirmations of applicability and robustness of EPAT-ChIP were obtained when the protocol was applied on the archival IBC FFPE sample to investigate other two HPTMs, H3K27em3 and H3K27ac.

Locus-specific analysis by real-time qPCR showed the expected enrichments of gene promoters for both HPTMs (Fig. 23). H3K27me3 ChIP-Seq distribution profiles obtained using both techniques were comparable (Fig. 24) and this is probably due to the wide distribution of this HPTM that renders its immunoselection less critical respect to other HPTMs. In fact, for H3K27ac (an HPTM less distributed then H3K27me3) we observed a significant gain in the quality of the profile

obtained with EPAT-ChIP protocol compared to that of the standard procedure (Fig. 26a).

In conclusion, these results demonstrate that EPAT-CHIP procedure has the potential to facilitate the application of epigenomic studies in clinical FFPE archival samples even when complex tissues are produced.

This technique will allow genome-wide chromatin studies in pathology tissues, enabling its wide spread use and thus contributing to extend the current understanding of cancer epigenomes, the identification of novel tumor subtypes, and the development of new clinical biomarkers.

10. Epigenomic profiling of aging in murine models: the possible role of caloric restriction

As discuss above, epigenetic mechanisms are involved in all the biological phenomena in which they were investigated, and recent evidence identify epigenetic alterations as one of the most important hallmark of aging.

In this second part of my work, I describe the application of PAT-CHIP to characterize, *in vivo*, the epigenetic basis of aging and its possible modulation by caloric restriction (CR) using murine models.

Briefly, C57Bl/6 female mice were fed with standard diet for different times - 3, 6 and 12 months - while the fourth group (12 month of age) was kept in caloric restriction regimen starting from their third month from birth (to ensure an adequate development) up to twelfth month. At the end of treatment, mice were sacrificed and livers were collected, fixed in formaldehyde and embedded in paraffin (FFPE). This approach was chosen for different reasons:

i) FFPE tissues allows to study the real epigenetic landscape avoiding chromatin rearrangements during sample handling;

ii) to ensure an optimal and easy storage;

iii) to allow possible further analysis on DNA binding proteins (e.g. transcription factors).

Chromatin was extract starting from 4 FFPE tissue section of 10 μm thickness each and immunoprecipitated using antibodies that recognize the following histone post-translational modifications:

- H3K4me3 (active promoter genes)

- H3K27me3 (repressive promoter genes)
- H3K27ac (enhancer and promoter of active gene)
- H3K4me1 (enhanced)

Immunoselected DNA from each IP was used for deep sequencing by Illumina HiSeq 2000 and high-quality reads were aligned versus a reference mouse genome (UCSC mm10).

10.1 Material and methods

Mice colonies and samples collection

C57BL6 8 weeks old female mice were generated at the European Institute of Oncology (Milan) facility and divided in four groups. Three groups are fed standard diet (*ad libitum*) up to 3, 6 and 12 month of life, while the samples of fourth group are fed with CR from third month up to their 12th month.

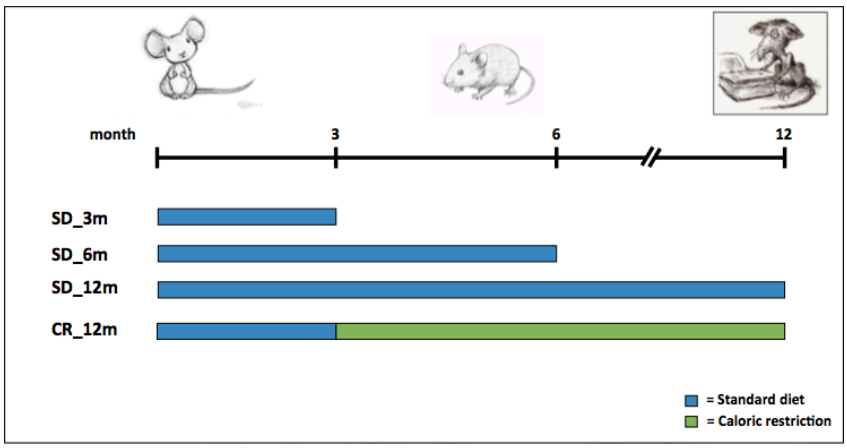


Figure 27. Mice colony and diet time treatment. C57BL6 8 weeks old female were divided in four groups and subjected to different diets for different times.

All the experiments with mice were performed in accordance with the Italian Laws (D.L.vo 116/92 and following additions), which enforces EU86/609 directives (Council Directive 86/609/EEC of 24 November 1986 on the approximation of law regulations and administrative, provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes).

Standard diet (2018S Tekland 18% Protein Rodent Diet, provided by Harlan Tekland, Madison, WI, USA) is a fixed formula, non-autoclavable diet; energy provided by the macronutrients was approximately 30% of proteins, 15% of fats and 55% of carbohydrates, for a total of 3.3 kcal g^{-1}.

Daily food intake *ad libitum* was measured in a subset of mice, by collecting and weighing all food remaining in the food hopper and cage at the same time each day for a week. Mice kept in CR were fed every day with 30% less of amount of food respect to the average daily food intake observed *ad libitum*.

In C57Bl/6 colony, all 3 months old females were fertile and plateaued at above 20 grams of body weight. Survival of these females was consistently 100% in the first year. The rare occurrence of injuries, deviant behavior and any external or internal, as determined upon necroscopy, signs of abnormalities including minimal alopecia, dermatitis or histiocytosis determined the exclusion of the individual from the study. Alterations of glycaemia or significant dyslipidemia or proteinuria were not detected in the females ranging 3-12 months of age fed SD. A barely significant drop of around 10% of systemic insulin sensitivity was appreciated in 12 vs 3 months old littermates of the mice fed SD included in the present epigenetic study whereas the 12 months old mice fed CR showed lower basal glycaemia and increased insulin sensitivity with respect to the 12 months old mice fed SD. Histological examinations of

the livers used to extract chromatin did not reveal structural differences in infiltrating cells or residue of hematopoiesis. Hepatocytes appeared comparable in size, ploidy and glycogen content, but early extent of steatosis was appreciable in all the 12 months old livers.

At the end of each treatment, mice were sacrificed by cervical dislocation and different organs were collected, rapidly washed in phosphate buffered saline (PBS) and incubate overnight at room-temperature (RT) in 4% formalin solution. Formalin-fixed samples were then dehydrated by increasing concentrations of ethanol, starting from 70% through to 80%, 90% and 100% (absolute ethanol), and subsequently included in paraffin by using a tissue processor.

Liver was chosen for different reasons: *i)* liver is a key metabolic organ *ii)* it is easy to manipulate because of the organ dimension *iii)* its histological structure is rather homogeneous. Altogether, 31 livers were collected from 3 months (n=8), 6 months (n=6), 12 months old mice fed SD (n=8) or CR (n=9).

Pathology-tissue chromatin immunoprecipitation (PAT-ChIP).

Material and methods

All reagents have been already listed in chapter 9.1 (Material and methods).

Chromatin preparation

Chromatin extraction was performed following the PAT-ChIP procedure (Fanelli *et al*., 2010; Fanelli *et al*., 2011) already described before in chapter 9.1 (Material and methods).

Immunoselection

Extracted chromatin was quantitated fluorimetrically by Qubit (Invitrogen) and immunoselection was carried out in incubation buffer using 200–400 ng of chromatin and antibodies previous described against four different HPTMs. All procedures used have been described in chapter 9.1 (Material and methods section).

After immunoselection, bound fractions were isolated, purified and analysed by real-time qPCR as previous described in chapter 9.1 – DNA isolation and locus-specific analysis by real-time PCR section. However, in this case, different primer pairs are used to evaluate the specificity of the immunoselection. All primers used are reported in table 1.

Gene	Forward primers sequence	Reverse primers sequence
Actb	TTCGCTCTCTCGTGGCTAGTACC	CAAAGAGTCTACACGCTAGGCGTTA
Col2a1	CCCCTGCCTACTTCCCTGACT	TGCTGCGACAGAGGTCCTTACT
Bact1	TTCCAGGCCCTCCCTCAT	GAACTTCCTGTCACAGTAGCAGGA
Bactin	TTCGCTCTCTCGTGGCTAGTACC	CAAAGAGTCTACACGCTAGGCGTA

Table 1. List and sequence of primers used to analyze the specificity of PAT-ChIP by real-time qPCR assay.

Library preparation and Hi-seq 2000 Illumina sequencing

Also in this case library preparation of each samples and sequencing were performed at IEO NGS facility (Milano). A brief description of the procedure is reported in chapter 9.1 (Material and methods – Library preparation and sequencing section).

Bioinformatic analysis

Processing of FASTQ files: Reads were aligned to the mm10 reference genome using 'bwa aln' (v0.6.2-r126). Unmapped reads, reads with a MAPQ smaller than 1, reads that mapped outside of chr1-19 and chrX, as well as duplicate reads were removed using samtools (v0.1.18).

Merging samples by group: For some analysis, the BAM files of the replicates of each group were merged using bamtools (v2.5.1), and subsequently indexed using samtools (v1.7).

Generation of bigWig files: BigWig files for each sample and group were generated from the respective BAM files using deepTools bamCoverage (v3.2.0) with a bin size of 10 bp, BPM-normalization (chrX was ignored for normalization), and reads are extended to 200 bp.

Data processing and plots: Down-stream analysis was executed inside a Docker/Singularity container. The Docker container used containing all tools, R and R packages used in the sections described below. If not specified otherwise, data was processed inside a Singularity (v2.6.0) container using R (v3.5.2) and plots were generated using the *ggplot2* R package (v3.1.0).

Genome-wide signal for UMAP and correlation: For each sample and group, the genome-wide signal was retrieved for consecutive, non-overlapping bins of 10 kb using deepTools multiBigwigSummary (v3.2.0). UMAP was calculated using the genome-wide signal of each sample in bins of 10 kb using the *umap* R package, with the "random_state" parameter set to 100.

Genome-wide correlation: Spearman correlation coefficients between samples and groups were calculated using deepTools plotCorrelation (v3.2.0) based on the genome-wide signal of 10 kb bins. The heatmap was generated using the *ComplexHeatmap* (v1.18.2) and *circlize* (v0.4.5) R packages. Statistical significance shown in the box plots were generated using the *ggsignif* (v0.4.0) R package using a Wilcoxon test.

TSS profiles, signal and ratio distribution: Annotation of TSS was taken from the curated RefSeq set, downloaded from UCSC Genome Browser on the 01.03.2018. The list of TSS includes the TSS of the longest transcript for each gene. The signal around TSS was calculated as mean signal in bins of 50 bp, with a range of +/- 5 kb around the TSS, using deepTools computeMatrix (v3.2.0). Missing data is treated as zero. The output was then plotted (profile) or summed up by TSS to generate the overall signal. For the ratio, summed signal values for H3K4me3 signal was divided by the H3K27me3 signal for each individual TSS.

Enhancer signal: Annotation of enhancers was taken from Y Shen et al. 2012). The signal around the three sets of enhancers were calculated as mean signal in bins of 200 bp, with a range of +/- 15 kb around the TSS, using deepTools computeMatrix (v3.2.0).

Chromatin state analysis: Chromatin state analysis was performed using ChromHMM (v1.17). Annotation of TSS +/- 2 kb represents the same list

of TSS as above with the range extended to +/- 2kb. Similarly, annotation of transcription end sites (TES) was taken from the curated RefSeq set, downloaded from UCSC Genome Browser. The list of TESs includes the TES of the longest transcript for each gene. Annotation of chrX was taken from the mm10.txt file provided with ChromHMM. LTR and LINE annotation was downloaded from the UCSC Genome Browser for mm9 and then converted to mm10 using the UCSC Genome Browser LiftOver feature. Annotation of constitutively lamina-associated domains (cLAD) was taken from Peric-Hupkes *et al*. (2010) and converted from mm9 to mm10 using the UCSC Genome Browser LiftOver feature (mm9 files can be found in the GSE17051 data set in the GEO database). First, merged BAM files of each group were binarized using the BinarizeBam command. Then, the chromatin state model was trained using the LearnModel command, settings the number of desired states to 8. Subsequently, the states were reordered using the Reorder command, then segmentation and enrichment with sets of genomic locations were generated using the MakeSegmentation and OverlapEnrichment commands. For the transition analysis, BED files of all groups containing the chromatin state for every location in the genome were cut into bins of 200 bp using awk (v1.3.3) and bedtools (v2.27.1). ChromHMM plots for definition of chromatin states and enrichment in sets of genomic locations were reproduced in *ggplot2* (v3.1.0). Alluvial plot was generated with the *ggalluvial* R package (v0.9.1).

10.2 Results

Chromatin extraction

As mentioned before, for each sample chromatin was extracted starting from four FFPE tissue sections of 10 µm and fluorimetrically quantified by Qubit (Invitrogen, Eugene, OR, USA) using the dsDNA HS Assay Kit (Invitrogen, Eugene, OR, USA). Being this work part of a larger project in collaboration with the European Institute of Oncology (Milan), only the chromatin and IP results of samples that I have processed are described here.

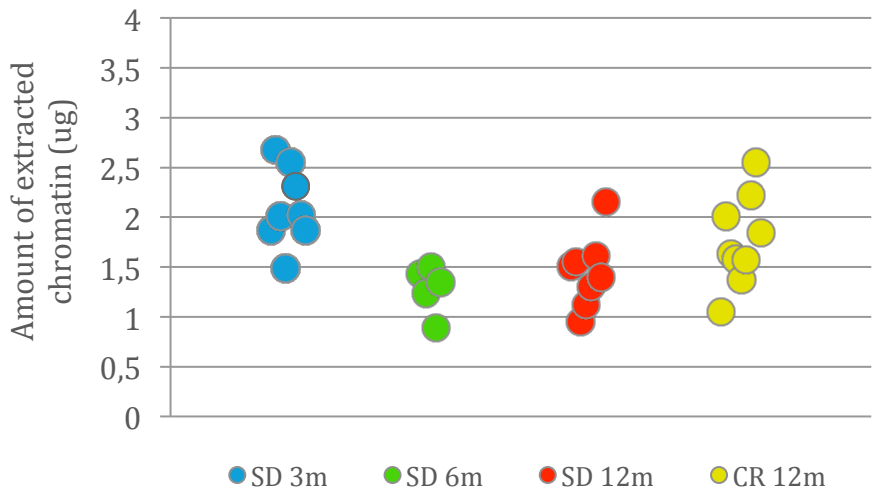


Figure 28. Amounts of extracted chromatin. Starting from four FFPE tissue sections of each sample, PAT-ChIP protocol was applied and extracted chromatin was fluorimetrically quantified by Qubit.

Immunoprecipitation and purification of "bound" fraction

Once extract, chromatin was immunoprecipitated with the different antibodies mentioned before. The DNA bound fractions were de-crosslinked and purified using the PCR Purification Kit (Qiagen, Hilden, Germany), and fluorimetrically quantified by Qubit. Results showed below are presented considering each histone modification separately (Figure 29).

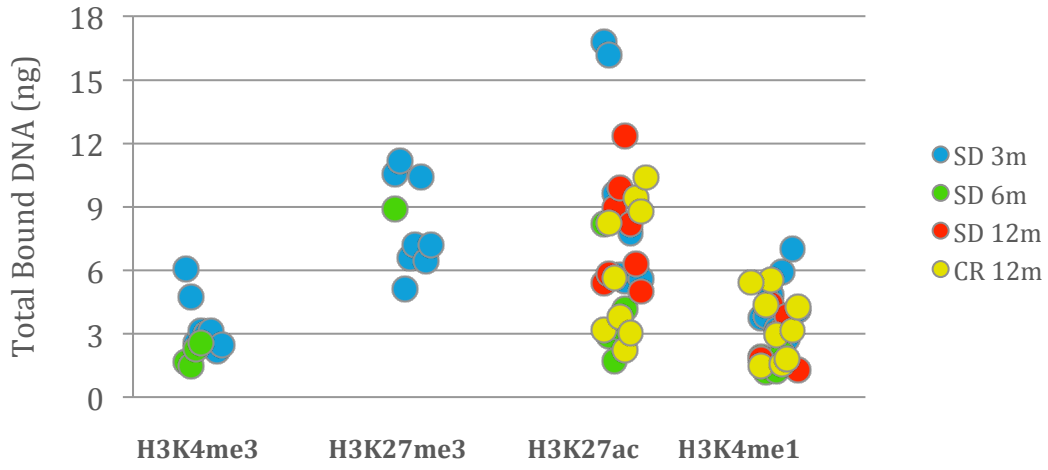


Figure 29. Total amounts of DNA immunoselected. Chromatin extracted was immunoprecipitated with anti-H3K4me3, anti-H3K27me3, anti-H3K27ac and anti-H3K4me1 antibodies. Immunoselected chromatin was de-crosslinked, and the DNA purified and fluorimetrically quantified.

Locus-specific analysis by real-time PCR

Specificity of the immunoselection was evaluated amplifying the purified DNA from bound and 5% input fractions by real-time quantitative PCR (qPCR) using the Fast Start SYBR Green Master Mix (Roche, Mannheim, Germany), the Rotor-Gene 6000 robocycler (Corbett Life

Science, Sydney, Australia) and using primers pairs able to specifically and quantitatively amplify the selected loci of interest (Table 1) know to be enriched of the HPTM investigated (Actb for both H3K4me3 and H3K27ac, Col2a1 for H3K27me3 and Bact1 for H3K4me1) and primer pairs able to amplify loci know to be not-enriched (Col2a1 for H3K4me3 and H3K27ac, Actb for H3K27me3, and Bactin for H3K4me1).

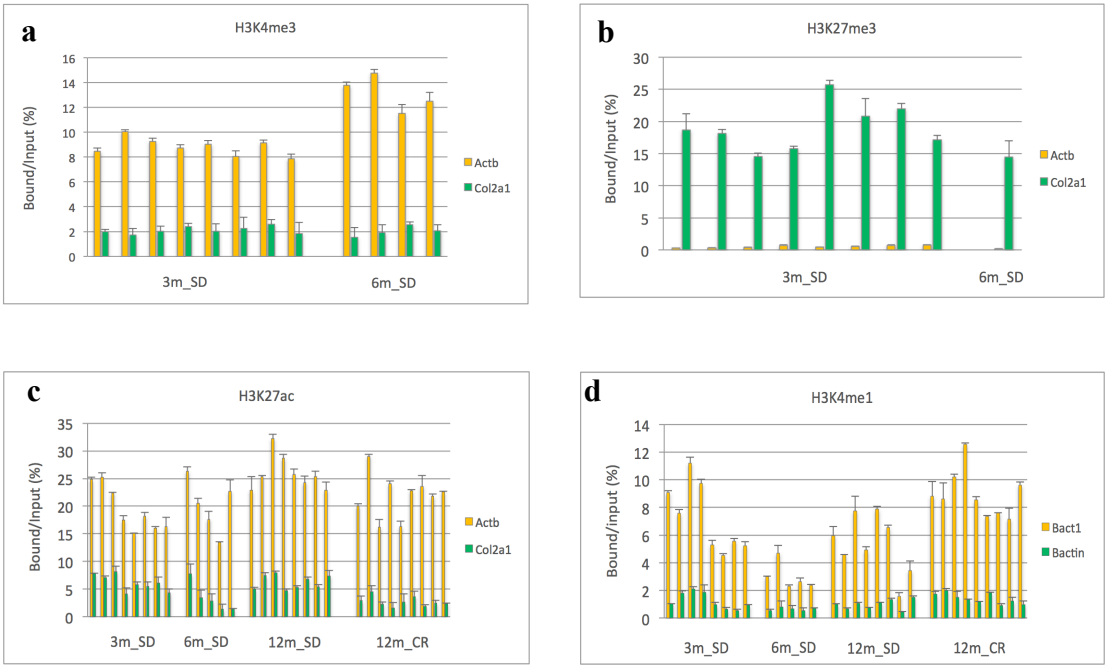
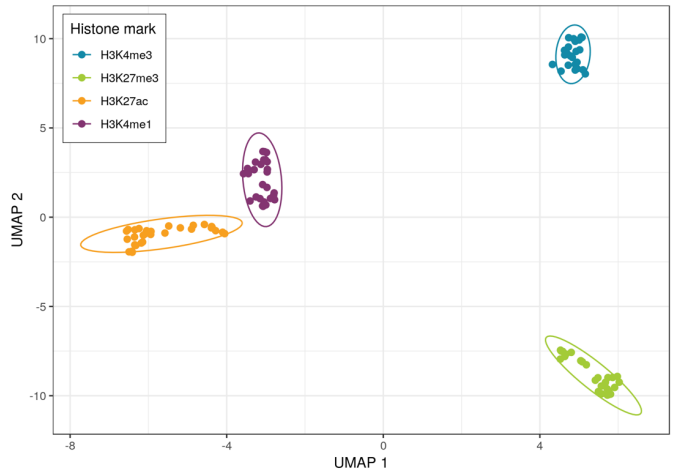


Figure 30. Qualitative analysis of immunoselection by real-time qPCR. Specificity of the immunoselection was evaluated amplifying the purified DNA from bound and 5% input fractions by real-time quantitative PCR.

Global analysis of histone mark profiles identified age and diet groups.

To establish whether differences among the histone marks, as well as age and diet, determine distinct clusters, all 108 profiles were analysed through a dimensionality reduction technique (uniform manifold approximation and projection, UMAP) using the genome-wide normalized signal of each respective histone mark in consecutive, non-overlapping bins of 10 kb. Samples clustered largely by histone mark, confirming that ChIP protocol have been conducted in homogenous manner (Fig. 31a) while the degree of heterogeneity among individuals is greater in older to younger samples for three histone marks (Fig. 31b). Comparison between profiles of same histone modification (Fig. 31b) allows distinguishing younger (blue dots) from older (yellow and red dots) in the case of H3K4me3 or H3K27me3, while for H3K27ac or H3K4m31 this distinction not occurs. Moreover only in H3K27ac profiles is possible to distinguish the CR profile (yellow dots) from the aged matched SD counterpart (red dots).

a



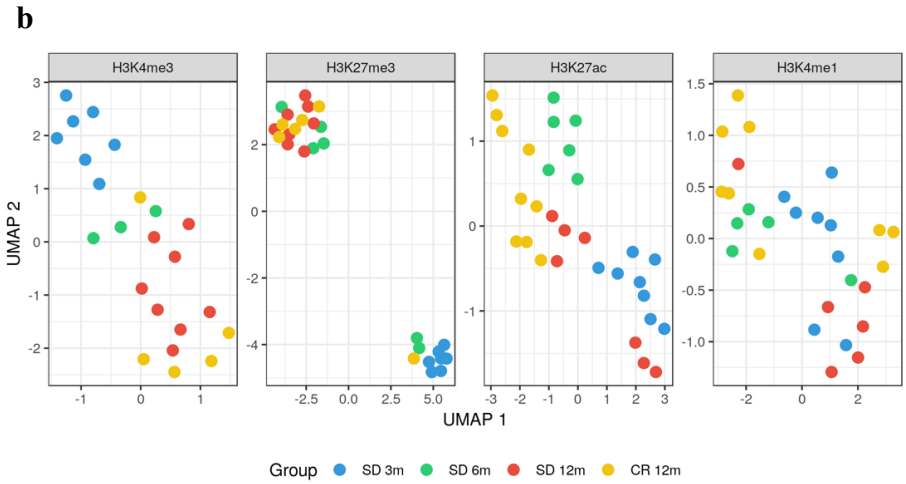


Figure 31. Global analysis of histone mark profiles. 108 profiles were analysed through UMAP technique. As show, samples clustered by histone mark while only H3K4me3 and H3K27me3 prifile allow to distinguish younger (blue dots) from older (yellow and red dots).

Merging the replicates within each group using genome-wide signal in consecutive, non-overlapping bins of 10 kb, Spearman's correlation coefficients were calculated, and differences are show below using a heatmap (Fig. 32). Results show that H3K4me1 and H3K27ac profiles are highly similar without differences among age or diet groups. On contrary, H3K4me3 and H3K27me3 have considerable differences between each other and a high internal correlation.

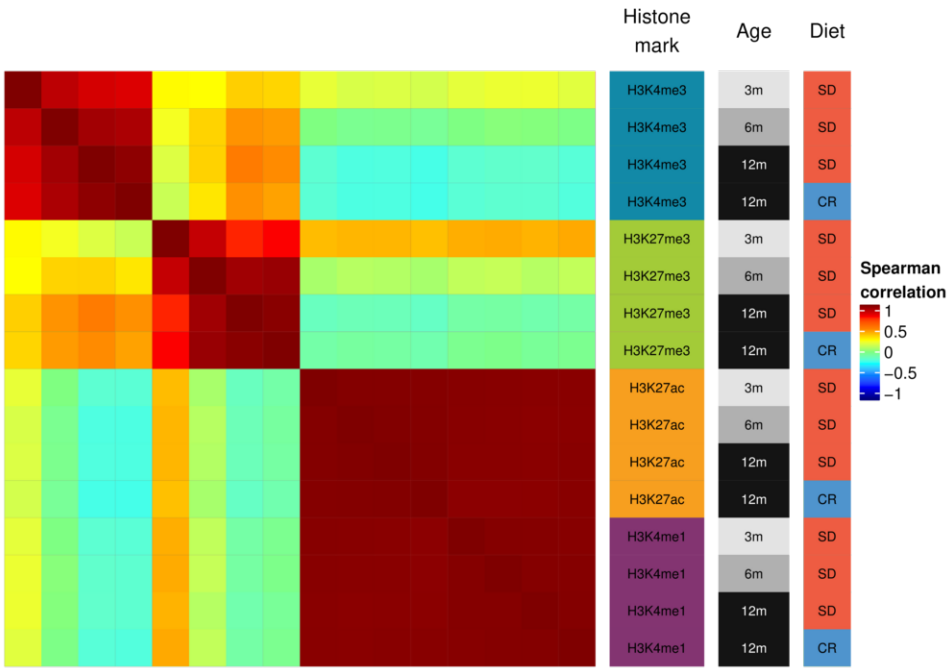


Figure 32. Heatmap representation of genome-wide correlation between HPTM and gourps. Spearman's correlation coefficients were calculated using non-overlapping bins of 10 kb. H3K4me1 and H3K27ac profiles are highly similar while, H3K4me3 and H3K27me3 have considerable differences.

Interestingly, we found that with age, the average correlation of H3K4me3 and H3K27me3 samples (left and central panel) to all H3K27ac and H3K4me1 samples significantly decreased from positive to negative values (anti-correlation) while, the correlation between H3K4me3 and H3K27me3 profiles increased with age (right panel), and 12 months samples differ significantly based on the diet they followed (CR vs SD – Fig.33). These findings suggest that, on a genome-wide scale, H3K4me3 and H3K27me3 are more strongly affected by aging than H3K27ac and H3K4me1 although this not excluded local age-related changes for these last histone marks.

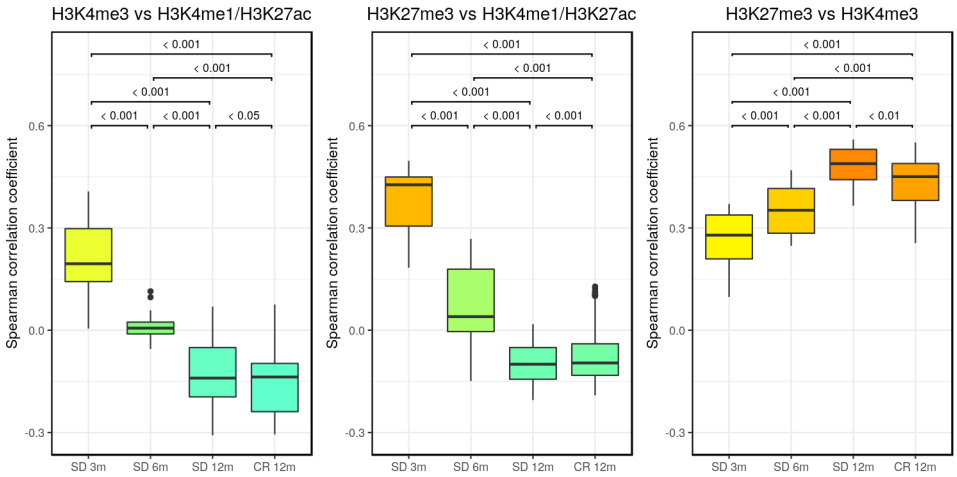


Figure 33. Box plot of Spearman correlation coefficients. With aging, correlation between H3K4me3 and H3K27me3 (first and second panel respectively) with H3K4me1/H3K27ac profiles decreases and are unaffected by CR, while correlation between H3K4me3 and H3K27me3 increase with age with a smaller effect in CR 12m group. P-value statistics: * < 0.05, ** < 0.01, *** < 0.001.

H3K4me3 and H3K27me3 signals around the TSS change with time.

H3K4me3 and H3K27me3 modifications are known to be enriched at the transcription start site (TSS) in a mutually exclusive manner (except for bivalent promoter). The analysis of the mean signals of H3K4me3 was performed and revealed consistently wider and higher peaks in the TSS regions of younger samples (Fig. 34 – left panel) and the drop was rescued by CR treatment that contrariwise not affect the loss of width.

Also for H3K27me3 modification, an average higher signal around TSS in younger than aged samples was recorded but, in this case, CR treatment did not rescue the loss of signals (Fig. 34 – right panel)

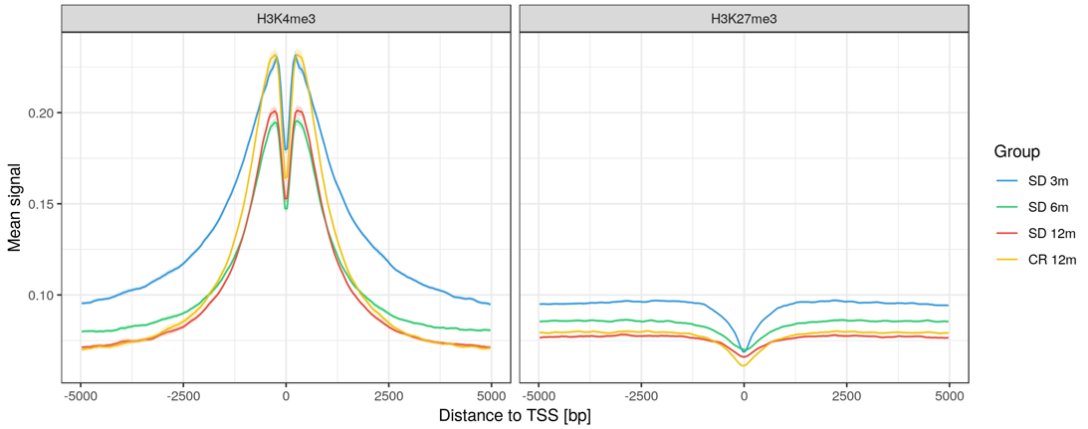


Figure 34. H3K4me3 and H3K27me3 signal at trasnscription start site (TSS). The signal around TSS was calculated as mean signal in bins of 50 bp, with a range of +/- 5 kb around the TSS.

We also investigated bivalency of H3K4me3 and H3K27me3 in all the 27,517 annotated TSS regions in younger and older samples. As shown below (Fig. 35), the ratio of H3K4me3-to-H3K27me3 signal revealed an almost unimodal distribution in 3 months old animals that dropped with age and became more bimodal in 12 months old mice, with an intermediate shape in the 6 months old mice. However also in this case, CR treatment not revert this trend.

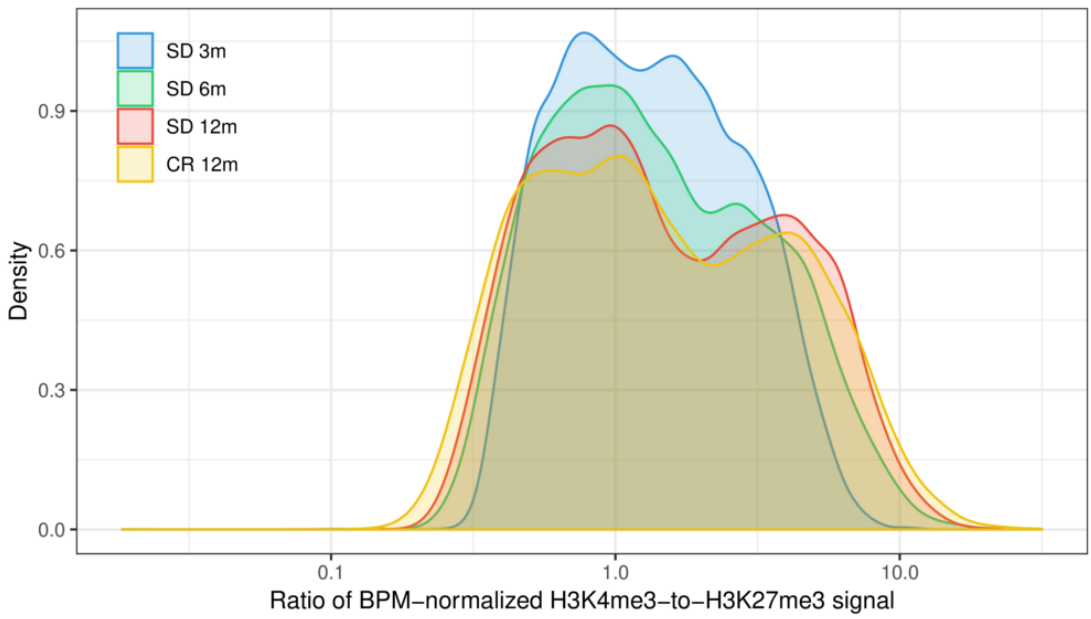


Figure 35. H3K4me3 and H3K27me3 signals ratio across TSS. Starting from unimodal distribution at 3 month, with age the signal ratio became bimodal and an intermediate shape in the 6 months old mice is recorded.

Further analysis, was focalized on TSS that appear to be targeted by the observed time- or diet-dependent changes. Results are summaryzed in two profiles showed in Fig. 36.

Fig. 36a contains the TSS regions of 405 genes in which H3K4me3 decreases with age and CR. These genes were significantly enriched in 28 GO terms involved in different processes, from immune response to import into cell, and response to stimuli. In Fig. 36b shows all TSS regions affected by a gradual loss of H3K4me3 through age, but with a partial rescue in CR.

These genes were significantly enriched in 94 GO terms and were classified as involved in cellular component maintenance, chromosome organization, chromatin dis-/assembly, nucleosome packaging, biosynthesis and metabolism

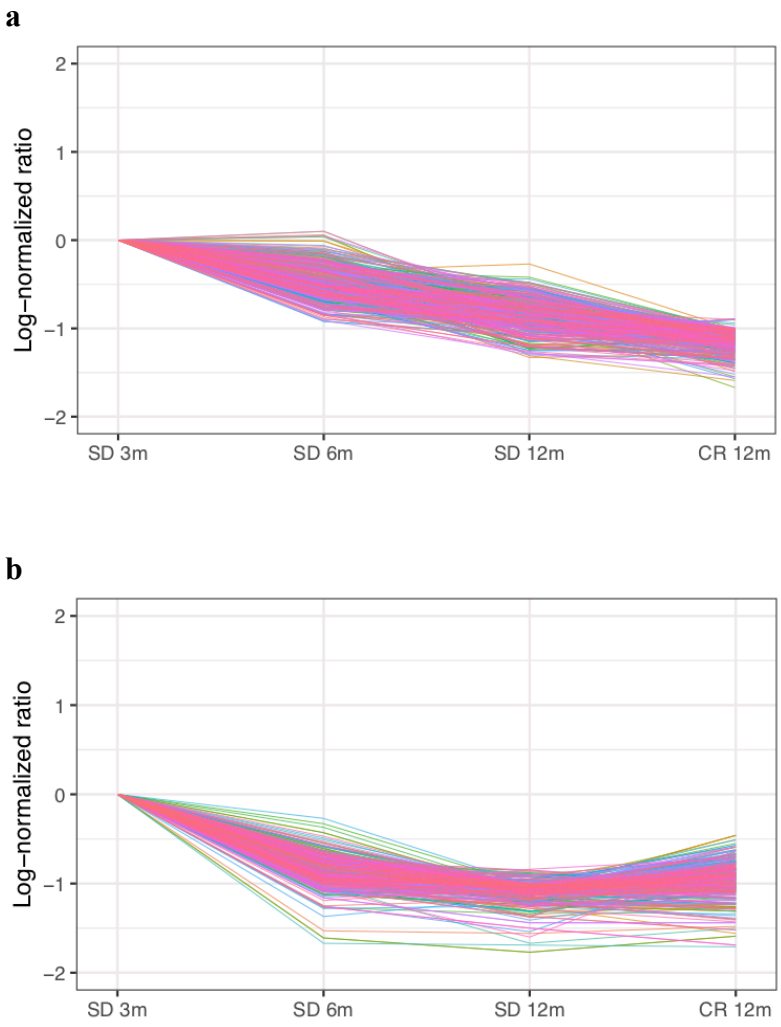


Figure 36. H3K4me3 signal across TSS that appears to be targeted by the observed time- or diet-dependent changes. (**a**) Log-ratio of TSS affected by a gradual loss of H3K4me3 signal through age and diet. (**b**) Log-ratio of TSS affected by a gradual loss of H3K4me3 signal through age but with a partial rescue in CR.

Time-dependent chromatin state transitions

The functional elements most affected by the observed time-dependent changes in histone modifications were determined through an integrated chromatin state analysis using all HM profile (ChromHMM). Chromatin states are defined as recurring combinations of the four histone marks investigated, which are then assigned to consecutive bins 200-bp across genome. To ensure maximum detail complexity with minimal state redundancy, 8 state models were set (Fig. 37 – left panel). Each state is characterized by specific genomic regions and functional elements (Fig. 37 – right panel).

As expected, given the histone modifications investigated, multiple TSS- (state 1 and 2) and enhancer-related (states 3-5) states were clearly identified while state 8 was arbitrarily labelled as quiescent chromatin because it contained none of the histone marks analysed (Fig. 37 – left panel). On contrary, state 7 presents a new signature of weak H3K4me3 and minimal H3K27me3 signal and the absence of the other two histone marks but, being state 7 not enriched in TSS region it is likely to not be the signature of bivalent promoters.

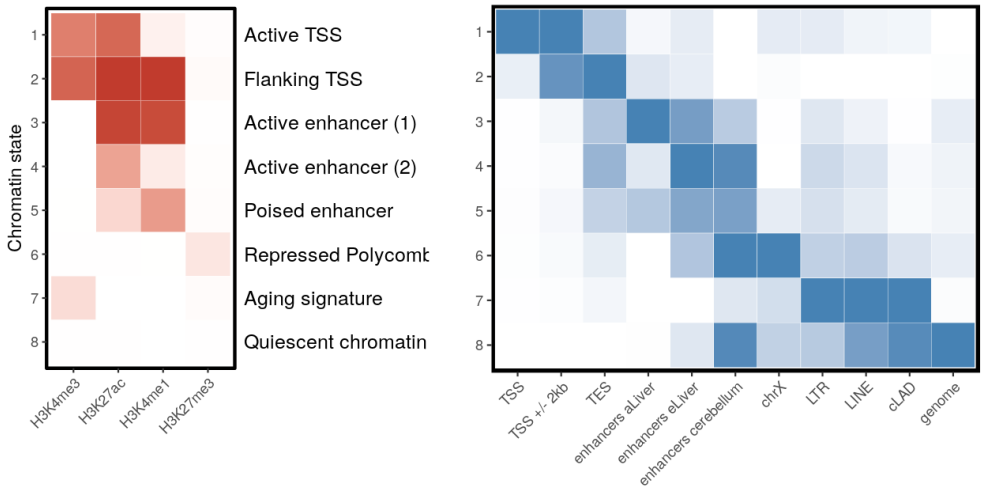


Figure 37. Chromatin state analysis. Chromatin states are defined as recurring combination of the four histone marks investigated. As expected, given the studied set of histone modifications, known promoter and enhancer signatures are found, as well as a new age-related chromatin state characterized by the presence of only H3K4me3.

As show in Fig. 38, where the percentages of genome coverage of each state are reported, the state 7 (enriched in LTR, LINE and constitutively lamina-associated domains – cLAD) is the state with the higher fold change between young and old mice kept in standard diet and, in function of this, this state was named "aging state".

Notably, mice kept in caloric restriction show only slightly elevated levels of state 7 compared to young mice. A similar trend can be observed for state 6, although with less drastic changes. The strong increase of state 7 in old mice kept in standard diet, compared to a very small increase in their littermates kept in caloric restriction indicates that state 7 is an effect of aging (or vice versa).

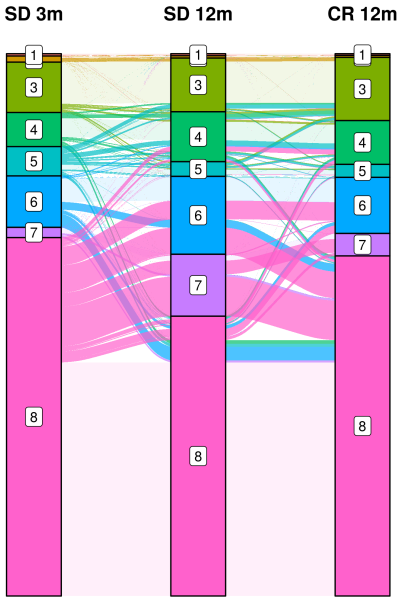


Figure 38. Chromatin state transition during aging. Alluvial plot showing the transition between chromatin states in the SD 3m, SD 12m, and CR 12m groups. Bar height represents genome coverage in percentages (the sum of all states gives 100%). Many, but not all, chromatin state transitions observed with age, especially the loss of quiescent chromatin, are reverted in the comparison with the CR 12m group.

Table 2 reports the percentage of the genome in which the chromatin states changed or not respect to the young group mice (SD 3m). Notably, the part of the genome in which chromatin state changes increases with age from about 20% in the first three months to more than 30% in the 12 months old mice. Moreover, the comparison with old mice kept in caloric restriction shows that only 24.18% of the genome changes.

	Fraction of genome with same chromatin state	**Fraction of genome with changing chromatin state**
SD 6m	**78.90%**	**21.10%**
SD 12m	**68.10%**	**31.90%**
CR 12m	**75.82%**	**24.18%**

Table 2. Percentage of the genome that has the same or changing chromatin states compared to the young group (SD 3m).

10.3 Discussion

Chromatin immunoprecipitation is considered one of the most powerful experimental approaches to investigate the epigenetic landscape in many biological models.

In this study PAT-ChIP-seq technology has been used to characterize the temporal patterns of four key histone H3 modifications in the liver of healthy C57Bl/6 mice within a year of age to determine whether and to what extend time impacts on chromatin states. Changes in covalent modifications of the histones tails as well as in histone variant incorporation and in DNA methylation have been extensively studied during the development and aging of different organisms, but a pure time-dependent changes in histone modifications have not been systematically assessed in mammals.

To shed light on this epigenetics changes, C57BL6 8 weeks old female mice were generated and divided in four groups. The first three groups have been fed with standard diet for up to 3, 6 and 12 month to investigate the age-associated chromatin remodeling, while the last group is kept in caloric restriction to investigate the potential effect of CR on the aging-associated HPTMs profiles.

Starting from four FFPE liver sections, PAT-ChIP-seq was successfully applied in terms of chromatin extraction (Fig. 28) and immunoprecipitation, as showed specificity by both locus specific analysis by real-time qPCR (Fig. 30) and genome-wide analysis by NGS (Fig. 37 left panel). DNA associated to each of the four HPTMs investigated, was analysed at the genome-wide level by NGS technology to obtain specific PTMs profiles of histone H3. The comparison of histone mark profiles obtained at 3, 6 and 12 months of age indicate that few months were sufficient to identify global changes in H3K4me3 and

H3K27me3 profiles allowing to distinguish young from old mice, although, genome-wide changes observed in this analysis were not prevented by CR (Fig. 31b).

Global analysis performed using UMAP technique, revealed that CR particularly affected the acetylation of H3K27 (Fig. 31b) and results are consistently with what has been reported on the effect of CR on acetylation of protein (Nakamura *et al*., 2013) or histone acetylation (Li *et al*. 2011; Shimazu *et al*., 2013; Gong *et al*. 2015).

At transcription start site (TSS) level, average signal of both H3K4 and K27 trimethylation show an almost linear temporal decay, rescue by CR that instead not affect the width loss of the H3K4me3 peak from 3 to 6 months of age (Fig. 34). Moreover, at the TSS, the H3K4me3 - H3K27me3 signal ratio changed from unimodal at 3 months to a bimodal pattern at 12 months (Fig. 35). The width of H3K4me3 peaks around the TSS has been linked to transcriptional consistency (Benayoun *et al*., 2014), while changes in H3 K4 and K27 tri-methylation bivalency may affect the stability of the promoter landscape.

In addition, over the time, the Spearman's coefficient correlation of H3K4me3 and H3K27me3 to H3K4me1/H3K27ac profiles decreased whereas the correlation between H3K4me3 and H3K27me3 profiles increased and appear to be affect by CR (Fig. 33)

Histone modifications are well known to be associated with different chromatin states, each characterized by specific genomic functions. In our analysis, the genome was divided in 8 different states (Fig. 37) allowing to investigate their behaviour among the different age and diet groups. Results show that, over the time, one forth of the genome changes chromatin state (Table 2). In particular, the states that underwent a massive transition are those assigned with quiescent chromatin (state 6,

state 7 and 8 – Fig. 38), changes that are largely prevented by CR diet as also show in Table 2.

Conclusion and future perspective

The data presented in this work show that pathology-tissue chromatin immunoprecipitation is a powerful approach to investigate the epigenomic landscape in FFPE tissues. Issues related to over-fixed FFPE samples indeed can be overcome using the new procedure called enhanced PAT-ChIP. The technique improves the efficiency of chromatin isolation from FFPE samples allowing the study of long time-fixed specimens (up to 72 h), as well as the investigation of low distributed epigenetic marks (e.g., H3K4me3) and the analysis of multiple histone marks from low amounts of starting material.

EPAT-ChIP will facilitate the application of chromatin studies to archived pathology samples, thus contributing to extend the current understanding of cancer epigenomes and enabling the identification of clinically useful tumor biomarkers.

In addition, PAT-ChIP was exploit to investigate the epigenomic profile of aging in murine models and the possible role of caloric restriction. The comparison of different age profiles of histone H3 marks, revealed that they change with time mainly in intergenic regions and near TSS and that some aging related effects observed are prevented by CR, suggesting that they may impact on aging phenotypes.

In conclusion, although our data do not establish if the epigenetic drift of H3 modifications is casually involved in aging, they shed light on the poor stability of epigenetic cell identity overtime identifying chromatin repression and remodelling in gene desert regions of the genome as putative basic phenomenon involved in the maintenance of genome functions in aging.

Bibliography

- Allfrey, V. G., Faulkner, R., & Mirsky, A. E. (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proceedings of the National Academy of Sciences*, *51*(5), 786-794.
- Altekruse, S. F., Kosary, C. L., Krapcho, M., Neyman, N., Aminou, R., Waldron, W., ... & Mariotto, A. (2010). SEER Cancer Statistics Review, 1975–2007. Bethesda, MD: National Cancer Institute; 2010. *Blagosklonny MV and Campisi J. Cancer and aging: more puzzles, more promises*, 2615-2618.
- Amatori, S., Ballarini, M., Faversani, A., Belloni, E., Fusar, F., Bosari, S., ... & Fanelli, M. (2014). PAT-ChIP coupled with laser microdissection allows the study of chromatin in selected cell populations from paraffin-embedded patient samples. *Epigenetics & chromatin*, *7*(1), 18.
- Amatori, S., Persico, G., Paolicelli, C., Hillje, R., Sahnane, N., Corini, F., ... & Pelicci, P. G. (2018). Epigenomic profiling of archived FFPE tissues by enhanced PAT-ChIP (EPAT-ChIP) technology. *Clinical epigenetics*, *10*(1), 143.
- An, W. (2007). Histone acetylation and methylation. In *Chromatin and Disease* (pp. 355-374). Springer, Dordrecht.
- Baccarelli, A., Wright, R., Bollati, V., Litonjua, A., Zanobetti, A., Tarantini, L., ... & Schwartz, J. (2010). Ischemic heart disease and stroke in relation to blood DNA methylation. *Epidemiology (Cambridge, Mass.)*, *21*(6), 819.
- Baker, M. (2011). Making sense of chromatin states.

- Bannister, A. J., & Kouzarides, T. (2005). Reversing histone methylation. *Nature*, *436*(7054), 1103.
- Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., ... & Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell*, *129*(4), 823-837.
- Bjedov, I., Toivonen, J. M., Kerr, F., Slack, C., Jacobson, J., Foley, A., & Partridge, L. (2010). Mechanisms of life span extension by rapamycin in the fruit fly Drosophila melanogaster. *Cell metabolism*, *11*(1), 35-46.
- Bochkis, I. M., Przybylski, D., Chen, J., & Regev, A. (2014). Changes in nucleosome occupancy associated with metabolic alterations in aged mammalian liver. *Cell reports*, *9*(3), 996-1006.
- Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., ... & Bell, G. W. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *nature*, *441*(7091), 349.
- Byvoet, P., Shepherd, G. R., Hardin, J. M., & Noland, B. J. (1972). The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. *Archives of biochemistry and biophysics*, *148*(2), 558-567.
- Casillas, M. A., Lopatina, N., Andrews, L. G., & Tollefsbol, T. O. (2003). Transcriptional control of the DNA methyltransferases is altered in aging and neoplastically-transformed human fibroblasts. *Molecular and cellular biochemistry*, *252*(1-2), 33-43
- Cedar, H., & Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews Genetics*, *10*(5), 295.

- Chouliaras, L., Mastroeni, D., Delvaux, E., Grover, A., Kenis, G., Hof, P. R., ... & van den Hove, D. L. (2013). Consistent decrease in global DNA methylation and hydroxymethylation in the hippocampus of Alzheimer's disease patients. *Neurobiology of aging*, *34*(9), 2091-2099.
- Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., ... & Boyer, L. A. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences*, *107*(50), 21931-21936.
- Fanelli, M., Amatori, S., Barozzi, I., & Minucci, S. (2011). Chromatin immunoprecipitation and high-throughput sequencing from paraffin-embedded pathology tissue. *Nature protocols*, *6*(12), 1905.
- Fanelli, M., Amatori, S., Barozzi, I., Soncini, M., Dal Zuffo, R., Bucci, G., ... & Alcalay, M. (2010). Pathology tissue–chromatin immunoprecipitation, coupled with high-throughput sequencing, allows the epigenetic profiling of patient samples. *Proceedings of the National Academy of Sciences*, *107*(50), 21535-21540.
- Fang M, Pak ML, Chamberlain L, Xing W, Yu H, Green MR. The CREB coactivator CRTC2 is a lymphoma tumor suppressor that preserves genome integrity through transcription of DNA mismatch repair genes. Cell Rep. 2015;11:1350–7.
- Felsenfeld, G., & Groudine, M. (2003). Controlling the double helix. *Nature*, *421*(6921), 448.ù

- Fischle, W., Wang, Y., & Allis, C. D. (2003). Histone and chromatin cross-talk. *Current opinion in cell biology*, *15*(2), 172-183
- Friedman, D. B., & Johnson, T. E. (1988). A mutation in the age-1 gene in Caenorhabditis elegans lengthens life and reduces hermaphrodite fertility. *Genetics*, *118*(1), 75-86.
- Gentilini, D., Mari, D., Castaldi, D., Remondini, D., Ogliari, G., Ostan, R., ... & Monti, D. (2013). Role of epigenetics in human aging and longevity: genome-wide DNA methylation profile in centenarians and centenarians' offspring. *Age*, *35*(5), 1961-1973.
- Ghosh, A. K., Majumder, M., Steele, R., White, R. A., & Ray, R. B. (2001). A novel 16-kilodalton cellular protein physically interacts with and antagonizes the functional activity of c-myc promoter-binding protein 1. *Molecular and cellular biology*, *21*(2), 655-662.
- Gilmour, D. S., & Lis, J. T. (1986). RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells. *Molecular and cellular biology*, *6*(11), 3984-3989.
- Gonzalez, S., & Rallis, C. (2017). The TOR signaling pathway in spatial and temporal control of cell size and growth. *Frontiers in cell and developmental biology*, *5*, 61.
- Greer, E. L., Maures, T. J., Hauswirth, A. G., Green, E. M., Leeman, D. S., Maro, G. S., ... & Brunet, A. (2010). Members of the H3K4 trimethylation complex regulate lifespan in a germline-dependent manner in C. elegans. *Nature*, *466*(7304), 383.
- Harikumar, A., & Meshorer, E. (2015). Chromatin remodeling and bivalent histone modifications in embryonic stem cells. *EMBO*

reports, *16*(12), 1609-1619.

- Harraan, D. (1955). Aging: a theory based on free radical and radiation chemistry.
- Heijmans, B. T., Tobi, E. W., Stein, A. D., Putter, H., Blauw, G. J., Susser, E. S., ... & Lumey, L. H. (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences*, *105*(44), 17046-17049.
- Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., ... & Wang, W. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature genetics*, *39*(3), 311.
- Heitman, J., Movva, N. R., & Hall, M. N. (1991). Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. *Science*, *253*(5022), 905-909.
- Hodawadekar, S. C., & Marmorstein, R. (2007). Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design. *Oncogene*, *26*(37), 5528
- Holzenberger, M., Dupont, J., Ducos, B., Leneuve, P., Géloën, A., Even, P. C., ... & Le Bouc, Y. (2003). IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature*, *421*(6919), 182.
- Inoki, K., & Guan, K. L. (2006). Complexity of the TOR signaling network. *Trends in cell biology*, *16*(4), 206-212.
- Jenuwein, T., & Allis, C. D. (2001). Translating the histone code. *Science*, *293*(5532), 1074-1080.

- Jones, M. J., Goodman, S. J., & Kobor, M. S. (2015). DNA methylation and healthy human aging. *Aging cell*, *14*(6), 924-932.
- Kane, A. E., & Sinclair, D. A. (2018). Sirtuins and NAD+ in the development and treatment of metabolic and cardiovascular diseases. *Circulation research*, *123*(7), 868-885.
- Kapahi, P., Chen, D., Rogers, A. N., Katewa, S. D., Li, P. W. L., Thomas, E. L., & Kockel, L. (2010). With TOR, less is more: a key role for the conserved nutrient-sensing TOR pathway in aging. *Cell metabolism*, *11*(6), 453-465
- Kapahi, P., Zid, B. M., Harper, T., Koslover, D., Sapin, V., & Benzer, S. (2004). Regulation of lifespan in Drosophila by modulation of genes in the TOR signaling pathway. *Current Biology*, *14*(10), 885-890.
- Kawakami, K., Nakamura, A., Ishigami, A., Goto, S., & Takahashi, R. (2009). Age-related difference of site-specific histone modifications in rat liver. *Biogerontology*, *10*(4), 415-421.
- Kenyon, C., Chang, J., Gensch, E., Rudner, A., & Tabtiang, R. (1993). A C. elegans mutant that lives twice as long as wild type. *Nature*, *366*(6454), 461.
- Kimura, K. D., Tissenbaum, H. A., Liu, Y., & Ruvkun, G. (1997). daf-2, an insulin receptor-like gene that regulates longevity and diapause in Caenorhabditis elegans. *Science*, *277*(5328), 942-946.
- Klass, M., Nguyen, P. N., & Dechavigny, A. (1983). Age-correlated changes in the DNA template in the nematode Caenorhabditis elegans. *Mechanisms of ageing and development*, *22*(3-4), 253-263.

- Klose, R. J., & Zhang, Y. (2007). Regulation of histone methylation by demethylimination and demethylation. *Nature reviews Molecular cell biology*, *8*(4), 307.
- Kops, G. J., Dansen, T. B., Polderman, P. E., Saarloos, I., Wirtz, K. W., Coffer, P. J., ... & Burgering, B. M. (2002). Forkhead transcription factor FOXO3a protects quiescent cells from oxidative stress. *Nature*, *419*(6904), 316.
- LaSalle, J. M., Powell, W. T., & Yasui, D. H. (2013). Epigenetic layers and players underlying neurodevelopment. *Trends in neurosciences*, *36*(8), 460-470.
- Lee, R. Y., Hench, J., & Ruvkun, G. (2001). Regulation of C. elegans DAF-16 and its human ortholog FKHRL1 by the daf-2 insulin-like signaling pathway. *Current Biology*, *11*(24), 1950-1957.
- ·Li, L., Greer, C., Eisenman, R. N., & Secombe, J. (2010). Essential functions of the histone demethylase lid. *PLoS genetics*, *6*(11), e1001221.
- Li, Y., Liu, L., & Tollefsbol, T. O. (2010). Glucose restriction can extend normal cell lifespan and impair precancerous cell growth through epigenetic control of hTERT and p16 expression. *The FASEB journal*, *24*(5), 1442-1453.
- Lin, K., Dorman, J. B., Rodan, A., & Kenyon, C. (1997). daf-16: An HNF-3/forkhead family member that can function to double the life-span of Caenorhabditis elegans. *Science*, *278*(5341), 1319-1322
- Liu, L., Cheung, T. H., Charville, G. W., Hurgo, B. M. C., Leavitt, T., Shih, J., ... & Rando, T. A. (2013). Chromatin modifications as determinants of muscle stem cell quiescence and chronological

aging. *Cell reports*, *4*(1), 189-204.

- Lyko, F. (2018). The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nature Reviews Genetics*, *19*(2), 81.
- McCay, C. M., Crowell, M. F., & Maynard, L. A. (1935). The effect of retarded growth upon the length of life span and upon the ultimate body size: one figure. *The journal of Nutrition*, *10*(1), 63-79.
- McElwee, J. J., Schuster, E., Blanc, E., Thomas, J. H., & Gems, D. (2004). Shared transcriptional signature in Caenorhabditis elegans Dauer larvae and long-lived daf-2 mutants implicates detoxification system in longevity assurance. *Journal of Biological Chemistry*, *279*(43), 44533-44543.
- Medvedev, Z. A. (1990). An attempt at a rational classification of theories of ageing. *Biological Reviews*, *65*(3), 375-398.
- Melov, S., Ravenscroft, J., Malik, S., Gill, M. S., Walker, D. W., Clayton, P. E., ... & Lithgow, G. J. (2000). Extension of life-span with superoxide dismutase/catalase mimetics. *Science*, *289*(5484), 1567-1569.
- Morgan, H. D., Sutherland, H. G., Martin, D. I., & Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. *Nature genetics*, *23*(3), 314.
- Morris, J. Z., Tissenbaum, H. A., & Ruvkun, G. (1996). A phosphatidylinositol-3-OH kinase family member regulating longevity and diapause in Caenorhabditis elegans. *Nature*, *382*(6591), 536.
- Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., ... & Kenyon, C. (2003). Genes that

act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. *Nature*, *424*(6946), 277

- Omodei, D., & Fontana, L. (2011). Calorie restriction and prevention of age‐associated chronic disease. *FEBS letters*, *585*(11), 1537-1542.
- Parthun, M. R. (2007). Hat1: the emerging cellular roles of a type B histone acetyltransferase. *Oncogene*, *26*(37), 5319.
- Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S. W., Solovei, I., Brugman, W., ... & Reinders, M. (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Molecular cell*, *38*(4), 603-613.
- Serra RW, Fang M, Park SM, Hutchinson L, Green MR. A KRAS-directed transcriptional silencing pathway that mediates the CpG island methylator phenotype. eLife. 2014;3:e02313.
- Seto, E., & Yoshida, M. (2014). Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harbor perspectives in biology*, *6*(4), a018713
- Sharakhov, I. V., & Sharakhova, M. V. (2015). Heterochromatin, histone modifications, and nuclear architecture in disease vectors. *Current opinion in insect science*, *10*, 110-117.
- Sharma, P., Knowell, A. E., Chinaranagari, S., Komaragiri, S., Nagappan, P., Patel, D., ... & Chaudhary, J. (2013). Id4 deficiency attenuates prostate development and promotes PIN-like lesions by regulating androgen receptor activity and expression of NKX3. 1 and PTEN. *Molecular cancer*, *12*(1), 67.
- Shen, Y., Yue, F., McCleary, D. F., Ye, Z., Edsall, L., Kuan, S., ... & Ren, B. (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature*, *488*(7409), 116.

- Siebold, A. P., Banerjee, R., Tie, F., Kiss, D. L., Moskowitz, J., & Harte, P. J. (2010). Polycomb Repressive Complex 2 and Trithorax modulate Drosophila longevity and stress resistance. *Proceedings of the National Academy of Sciences*, *107*(1), 169-174.
- Solomon, M. J., Larsen, P. L., & Varshavsky, A. (1988). Mapping proteinDNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell*, *53*(6), 937-947.
- Stadtman, E. R. (1992). Protein oxidation and aging. *Science*, *257*(5074), 1220-1224.
- Strahl, B. D., & Allis, C. D. (2000). The language of covalent histone modifications. *Nature*, *403*(6765), 41
- Suh, Y., Atzmon, G., Cho, M. O., Hwang, D., Liu, B., Leahy, D. J., ... & Cohen, P. (2008). Functionally significant insulin-like growth factor I receptor mutations in centenarians. *Proceedings of the National Academy of Sciences*, *105*(9), 3438-3442.
- Sullivan, B. A., & Karpen, G. H. (2004). Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nature structural & molecular biology*, *11*(11), 1076.
- Sun, D., Luo, M., Jeong, M., Rodriguez, B., Xia, Z., Hannah, R., ... & Gu, H. (2014). Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell stem cell*, *14*(5), 673-688
- Tatar, M., Kopelman, A., Epstein, D., Tu, M. P., Yin, C. M., & Garofalo, R. S. (2001). A mutant Drosophila insulin receptor homolog that extends life-span and impairs neuroendocrine function. *Science*, *292*(5514), 107-110

- Tobi, E. W., Lumey, L. H., Talens, R. P., Kremer, D., Putter, H., Stein, A. D., ... & Heijmans, B. T. (2009). DNA methylation differences after exposure to prenatal famine are common and timing-and sex-specific. *Human molecular genetics*, *18*(21), 4046-4053.
- Tower, J. (2000). Transgenic methods for increasing Drosophila life span. *Mechanisms of ageing and development*, *118*(1-2), 1-14
- Turner, B. (2001). ChIP with native chromatin: advatages and problems relative to methods using cross-linked materials.
- Vaquero, A., Scher, M., Erdjument-Bromage, H., Tempst, P., Serrano, L., & Reinberg, D. (2007). SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation. *Nature*, *450*(7168), 440
- Vezina, C., Kudelski, A., & Sehgal, S. N. (1975). Rapamycin (AY-22, 989), a new antifungal antibiotic. *The Journal of antibiotics*, *28*(10), 721-726.
- Waddington, C. H. (1939). An introduction to modern genetics. *An introduction to modern genetics*.
- Waki, T., Tamura, G., Sato, M., & Motoyama, T. (2003). Age-related methylation of tumor suppressor and tumor-related genes: an analysis of autopsy samples. *Oncogene*, *22*(26), 4128.
- Waterland, R. A., & Jirtle, R. L. (2003). Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Molecular and cellular biology*, *23*(15), 5293-5300.
- Webb, A. E., & Brunet, A. (2014). FOXO transcription factors: key regulators of cellular quality control. *Trends in biochemical sciences*, *39*(4), 159-169.

- Whetstine, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., ... & Shi, Y. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell*, *125*(3), 467-481.
- Wilson, A. S., Power, B. E., & Molloy, P. L. (2007). DNA hypomethylation and human diseases. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, *1775*(1), 138-162.
- Xiao, F. H., Kong, Q. P., Perry, B., & He, Y. H. (2016). Progress on the role of DNA methylation in aging and longevity. *Briefings in functional genomics*, elw009.
- Xiao, R., Zhang, B., Dong, Y., Gong, J., Xu, T., Liu, J., & Xu, X. S. (2013). A genetic program promotes C. elegans longevity at cold temperatures via a thermosensitive TRP channel. *Cell*, *152*(4), 806-817.
- Yang, X. J., & Seto, E. H. A. T. (2007). HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene*, *26*(37), 5310.