Coordination between cell size and chromatin structure

Lucie Martin

Thesis submitted for the degree of Doctor of Philosophy
of Imperial College London
February 2022

Institute of Clinical Sciences,
Faculty of Medicine Imperial College London

Personal declaration

All the work presented in this dissertation is either mine or properly attributed to the corresponding authors.

Copyright declaration

Acknowledgements

First of all, I would like to acknowledge my two supervisors, Samuel Marguerat and Vahid Shahrezaei, who instigate the project and who trusted me to fulfil it. Without them, none of this work would have seen the light. Their complementary skills were a great source of reflection on my project. A special thanks to Samuel Marguerat, my main supervisor, for his amazing support and guidance throughout the PhD, from the bench to the analysis to the writing, even when things did not exactly go as planned. I will always respect his sense of precisions which takes him as close as possible to the truth.

I would like to thank Amalia Martínez Segura, a former member of the QGE lab, who initiated the project and who generated the relevant dataset I could use to complete my study. All the other members of the QGE lab, past and present, also contributed one way or another to this work, during lab meetings, discussions in the office or at lunch breaks. Xi-Ming, Amalia, Ben, Istvan, Wenhao gave me precious tips either at the bench or during the analysis and with whom I shared good laughs and refreshing breaks.

I would like to mention the people from the CRB building, Aubin, George, Julia, Cassandra, Jen, with whom I exchanged exciting scientific ideas or to whom I complained about massive experimental failures.

To Laura, Luke, Morgane, Ross, Miguel, thank you for poof-reading some pages of the thesis and correct the English.

My friends in London were also important during this challenging time; thank you Laura and Cecilia for the endless philosophical talks after a long day in the lab; George, Ifi, Julia, Vlad, Tatiana, Pedro for the gathering; the climbing crew Luke, Benji, Miguel, Joel for the awesome bouldering and rope climbing session, Aubin, JF, Amer, Laura for the camping trip across the UK; Morgane for her occasional visit in London.

To my friends from France, Mathilde, Jess, Louim, Alix, Eve for their kind encouragement.

Last but not least, to my partner, Armand, for his indefectible positivity, his inspiring mind and for his special vision of life. This journey was also his. To my parents, Chrystelle and Yves, who supported me throughout my studies, I am immensely grateful for their care

and help. To my brother Leo, and Romane for their crazy minds and their ability to make me disconnect from work. My in-laws, Jacques and Isabelle, for their kindness and their welcome.

Abstract

Cells change size during the cell cycle and in response to external conditions. In order to maintain biomolecule concentrations, gene expression is coordinated with cell size in a process called “scaling”. In fission yeast, scaling has been found to be associated with a genome-wide increase in transcription initiation rates and RNA polymerase II occupancy. However, the mechanistic details that underpin this global increase in transcription have not been defined yet. Since transcription initiation operates in the context of chromatin, we hypothesize that global transcriptional scaling is in part regulated by changes in chromatin state that reflect cell-size increase. To test this hypothesis, we used an analogue sensitive strain of the cyclin-dependent protein kinase Cdc2 (*cdc2-asM17*). Upon analogue treatment this strain exits the cell cycle in G2 and grows up to a size 3-4 times larger than wild-type after 6h of treatment. By mapping nucleosome occupancy in arrested *cdc2-asM17* cells of increasing size, I uncover the link between cell size and the global chromatin architecture. More precisely, I show that cell size increase is accompanied by an increasing occupancy at promoters. These factors occupying the promoter in large cells are MNase-sensitive factors, suggesting that they are not canonical nucleosomes. Moreover, I report that a marker of transcription initiation, H3K9ac, scales with cell size, suggesting a possible role of this post-translational modification in setting a favorable context for an increasing occupancy at promoter of larger cells. Finally, comparing the genome wide distribution of H3 between normal and large cells reveals that for most genes, the protein complex occupying the promoter of large cells is not composed of H3, suggesting that they are not nucleosomal particles, although some promoters indeed present an enrichment of H3 upon cell size increase. Altogether, this work provides insights into the mechanism that regulates scaling.

Table of content

List of Figures

List of Annexes

List of Tables

List of Abbreviations

CAGE-seq: Cap Analysis Gene Expression-sequencing

ChIP-seq: Chromatin Immuno-Precipitation- sequencing

CUT&RUN: Cleavage under targets and release using nuclease

FACT: Facilitates Chromatin Transcription

GRF: General Regulatory Factors

GTF: General Transcription Factor

HAT: Histone Acetyl Transferase

HDAC: Histone Deacetylase

MNase: Micrococcal Nuclease

NDR: Nucleosome Depleted Region

NRL: Nucleosome Repeat Length

PTM: Post-Translational Modification

RNAPII: RNA Polymerase II

RSC: Chromatin Structure Remodeling

TAF: TBP Associated Factor

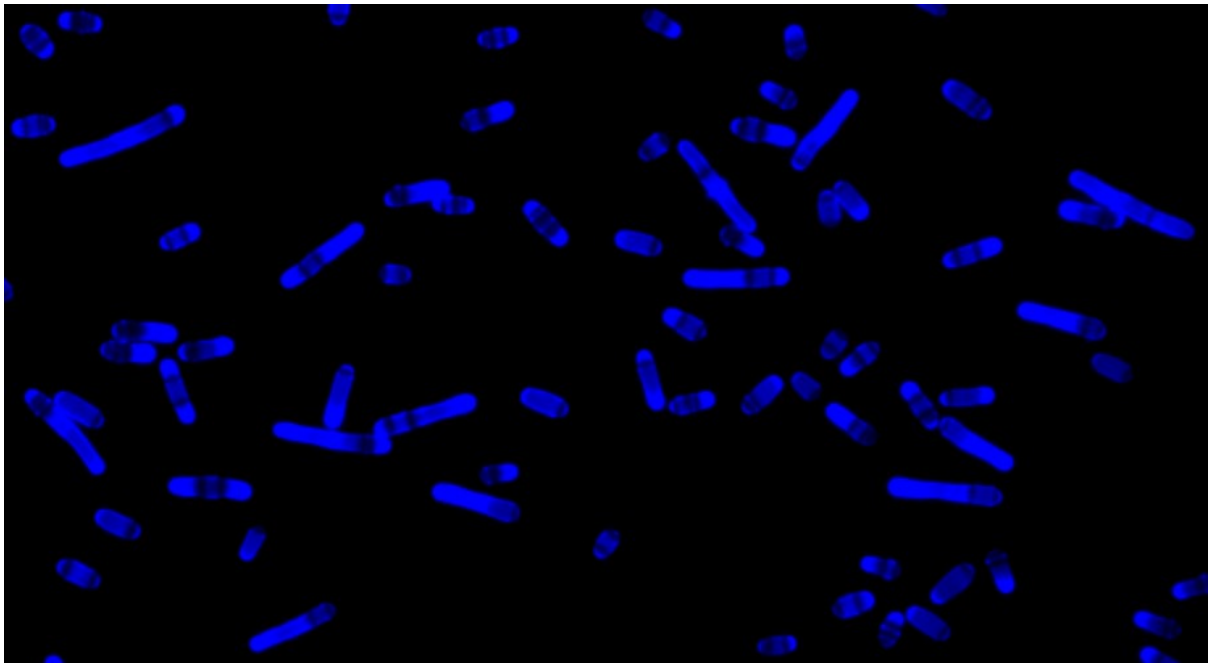
TBP: TATA-box binding protein

TCA: Citric Acid Cycle

TES: Transcription End Site

TF: Transcription Factor

TSS: Transcription Start Site



Schizosaccharomyces pombe mutant (cdc2-asM1) stained with calcofluor at different stages of elongation.

Chapter 1: Introduction

1.1. Size, shape and function

1.1.1. -Size is different depending on organisms

Living creatures populate the Earth in many forms and their shapes are so diverse that it would be impossible to cite them all. The size is perhaps the most prominent parameter: from the 90 meters tall sequoia redwood tree to the few millimeters' long ant, the difference is large! Thanks to the genius invention of the microscope, the characterization of biological systems goes deeper. In 1667, Robert Hook elucidated the building block of every living form: the cell. Consequently, we understood that differences in organisms' size is caused by a different total number of cells, rather than a difference in cell size (Marguerat and Bähler, 2012; Conlon and Raff, 1999). Nevertheless, if cell types belonging to a multicellular organism are placed under the microscope, one will witness striking and various shapes and sizes: the giant neuron, the elongated muscle cell, the tiny lymphocyte and so on.

"[…] Earth it self, which lyes so neer to us, under our feet, shews quite a new thing to us, and in every little particle of its matter, we now behold almost as great a variety of Creatures, as we were able before to reckon up in the whole Universe itself.", Hook wrote in his Micrographia in 1665

Four centuries later, thanks to modern technologies as well as the exploration of all corners of the planet, humanity has a dramatically broader knowledge of the incredible aspects that describe life forms.

1.1.2. Size can be linked to function

Size and shape of living forms do not only serve to identify them by biology lovers; it is a plastic trait that can be influenced by the environment, by ecological processes or even by physical constraints (Vermeij, 2016; Price and Hopkins, 2015; Payne et al., 2011; Smith and Lyons, 2011; Haldane, 1928) On the one hand, how the environment dictates the flexibility of animal size is

illustrated by macroevolution studies that reveal the existence of giant rodent in islands (Foster, 1964), dwarf elephants during the Late Cenozoic (Smith and Lyons 2011) or big dragonflies at the late Paleozoic (Calder, 1996). On the other hand, despite this flexibility, Haldane demonstrates that "for every animal there is a most convenient size" (Haldane, 1928). This statement reflects the idea that the size of an organism directly constrains its shape and structure (Thompson, 1942) and that size flexibility cannot exceed a certain range unless affecting the shape and function (Nicholson, 2020; Denny et al., 1985; Haldane, 1928). Altogether these observations convey the idea that size, structure, and function are in relation to one another.

One can hypothesize that this also holds true at the microscopic scale, in the context of cell biology and among eukaryotic as well as prokaryotic organisms. Let us take the example of the elongated muscle cell mentioned earlier. The muscle cell results from the fusion of multiple cells, creating a multinucleated cell that can reach several centimeters long. The striking length is a favorable condition to the formation of long chains of myofibril, responsible of the contraction of the muscle (Pollard et al., 1974). The blood cell must maintain a small size to pass through the capillaries (Amodeo and Skotheim, 2016). And the 1-meter neuron can transduce signals from the brain to the distanced limbs (Amodeo and Skotheim 2016).

Some studies reported a correlation between the specific dimension of a cell type and their fitness (Miettinen and Björklund, 2016; Monds et al., 2014). In a microevolution set up, scientists were able to observe that the largest bacteria from independent populations were rapidly selected after growth from a single clone (Lenski and Travisano, 1994). In accordance with these results, the measure of mass and growth rates of individual cells of *Bacillus subtilis*, *Escherichia coli*, *Saccharomyces cerevisiae* and mouse lymphoblasts indicates that heavier cells grow faster than lighter cells (Godin et al., 2010). On the contrary, in *Drosophila* larvae, reduced cell area seems to show better competitive advantages (Trotta et al., 2007). Another interesting example of varied sizes and shapes can be found in cancer cells that develop in multicellular organisms. How do some cancer cells manage to be larger, narrower, or smaller than normal cells? An elegant experiment on prostate cancer cells PC3 revealed the tendency of small PC3 cells to be more tumorigenic (Li et al., 2015b). Other studies show that solid tumors often contain giant cells that can survive anticancer therapy (Mirzayans et al., 2018).

From these observations, one could speculate that cell size influences tumor survival or development, although one could not exclude that size might simply reflect certain properties that help the tumor to survive.

1.1.3. Variation of size within a same cell population

To add up a layer of complexity, the size of an individual within the same cell type population can also be very variable, though it varies within the same order of magnitude. For the sake of simplicity, I will focus, from now on, only on cell size, rather than organism size.

The size of proliferating cells depends on the cell cycle stage, the DNA content as well as nutrients in unicellular organisms, or growth factors in multicellular systems. Indeed, cell size of proliferating cells is a balance between growth and division. The extent of growth varies with nutrient intake and the synthesis capacity of the cell which involves DNA, RNA and protein metabolism (Kleijn et al., 2021). In addition, cell size is also affected by random fluctuations operating at all levels of cellular processes, including gene expression, biochemical reactions or partitioning of proteins for division (Modi et al., 2017). Cycling cells are relatively small after division and reach their maximal size before division. For example, the length of *E. coli* increases around 3 µm between 2 division rounds (Campos et al., 2014b). The fission yeast *Schizosacharomyces pombe*, divides after reaching 14 µm (Wood and Nurse, 2015). On average, the size of symmetrically dividing cells doubles between each division (Jones et al., 2019). Moreover, it is now well established that cells grown in rich media are larger than cells cultivated in minimal media (Pérez-Hidalgo and Moreno, 2016). Data collected in our group (F. Bertaux), combined with others (Campos et al., 2014a), show that changing the nitrogen sources also affects growth rate and cell size.

However, despite the size variability within a given cell type, the distribution of size within a cell population in a given condition generally displays constant means and variances over time (Jones et al., 2019; Jun and Taheri-Araghi, 2015; Wood and Nurse, 2015). In a fully-grown human being, blood cells never exceed ~10 µm and neurons always span at least 1 meter (Amodeo and Skotheim 2016). This tendency of organisms to accurately maintain cell size is called size homeostasis (Amodeo and Skotheim, 2016; Wood and Nurse, 2015). Cell size

homeostasis is key for daughter cells to receive a suitable molecular material to sustain their mass and functions (e.g., such a control on cell size prevents the cell from losing their cellular content after several divisions). Beyond the empirical necessity of cell size homeostasis, evidence that this principle is fundamental in cell biology emerged from experiments wherein cell size is perturbed (Saucedo and Edgar, 2002). For instance, cell size defects in the brain caused by Pten deletion led to the human syndrome Lhermitte-Duclos (Backman et al., 2001; Kwon et al., 2001). In addition, in S6 kinase 1 defective mice, researchers observed a decrease in insulin secretion and in insulin content. Surprisingly, this phenotype of hypoinsulinaemia is not due to a problem in insulin production or glucose-sensing but instead, to a diminution of β cell size associated with a mass loss (Pende et al., 2000).

Altogether, these observations demonstrate the importance of cell size homeostasis. Most importantly, the existence of size homeostasis suggests that the cells can regulate, control, and potentially sense their size: how they do so has fascinated generations of researchers and is still an active field of research.

1.2. Size homeostasis implies a regulation of cell size

1.2.1. Three behaviors explaining cell size homeostasis

Cell size homeostasis necessarily implies a mechanism that sets the size at which a cell divides or commits to the cell cycle. Without such a mechanism, the size distribution of a cell population would change over time.

In theory, the control of cell size can originate from three different mechanisms: timer, adder and/or sizer (Vargas-Garcia et al., 2018) (**Fig 1.1**). A timer behavior relies on a constant amount of time between two divisions. Interestingly, a timer mechanism for exponentially growing cells cannot produce size homeostasis since the slightest asymmetry at division gives rise to cells which will deviate more and more from the average size. Besides, in such a scenario, cell size homeostasis is achieved without a sensing of cell size, but instead, with a sensing of how long they grew from birth to division. In the adder scenario, the cells grow of a constant amount of

mass before dividing. In the timer and adder strategies, cell size at birth is not involved for correcting any deviation from the standard size of the population (Fig 1.1). Therefore, cell size does not actively play a role in the size regulation in these two scenarios. On the contrary, cells relying on the sizer mechanism are able to directly sense their size; they trigger division after reaching a critical size (Wood and Nurse, 2015). (Fig 1.1). Although many efforts were deployed, assigning one of those mechanisms to each species is still a challenging task. One difficulty is that some species use a combination of distinct types of cell size regulation.

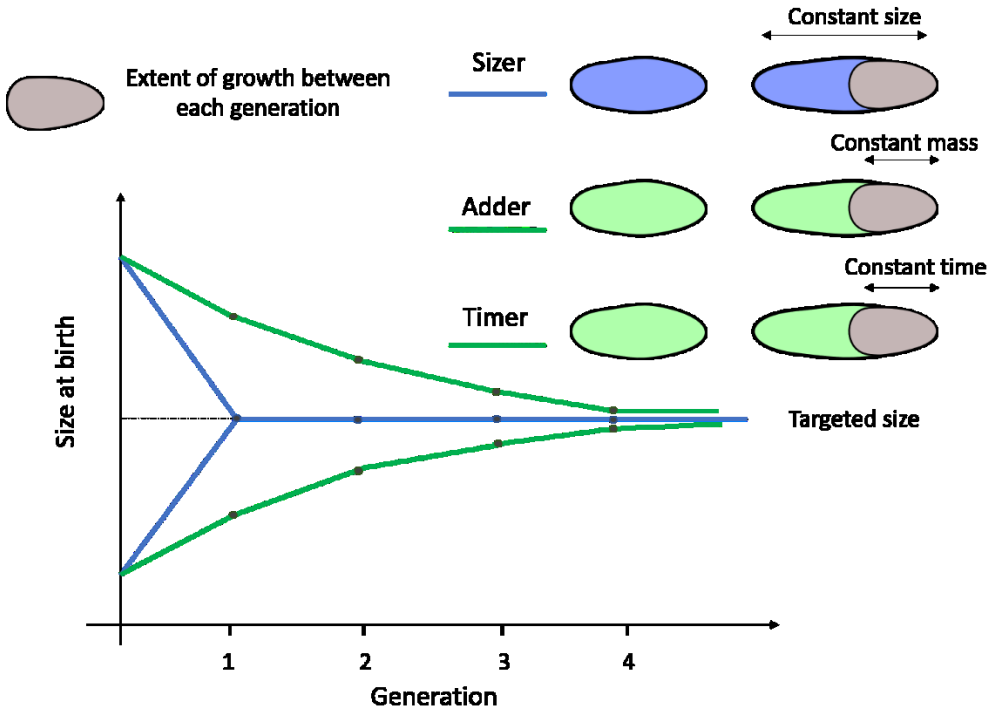


Figure 1.1: Different strategies for cell size homeostasis. *Cells following a sizer mechanism can correct their size within a cell cycle(Wood and Nurse, 2015). Timer and adder rely on a constant time or mass over generations to correct their size. In this cartoon, it is assumed that the cells having a timer are growing linearly. Adapted from Facchetti and colleagues (Facchetti et al., 2017).*

In the RAW 264.7 macrophage line and in rat leukemia cells, it was observed that a G1 sizer and a G1 adder could both operate depending on the birth size (Varsano et al., 2017). However, a previous investigation in proliferating rat Schwann cells revealed that size homeostasis is unlikely to be achieved thanks to a size threshold (Conlon and Raff, 2003), which complicates the model of cell size control. Further investigations with single-cell technologies showed that mammalian cells behave like adders (Cadart et al., 2018). Although the adder mechanism for mammalian cells seems to have prevailed in the literature, experiments for size control are

conducted on rapid proliferative cells which are exceedingly rare in adult mammals. Therefore, the use of unicellular organisms is straighter forward to explore the field of cell size control. Besides, their simpler cell geometry, their smaller regulatory pathways, and the availability of a great collection of mutants and genetic tools render them a perfect choice for tracking the coordination between growth and division.

In bacteria, it was first thought that cell size homeostasis followed a timer and a sizer mechanism. The timer emerges from the inference that the time between DNA replication and division is constant; the sizer because it was reported that the initiation of DNA replication always occurs at the same critical mass (Donachie, 1968). It is then the sizer that gained general approval in the field, especially thanks to mathematical modelling (Koch and Schaechter; Robert et al., 2014). However, the main criticism of these works is the systematic use of bacteria models that divide symmetrically (while many bacteria divide asymmetrically) and the fact that these bacteria models can adapt their size to nutrient availability, which can bias the conclusions. Thanks to the use of *Caulobacter crescentus* that divides asymmetrically and is insensitive to environmental fluctuations, as well as the tracking of a large number of cells of both *E. Coli* and *C. crescentus*, Campos and colleagues provide a convincing model for regulation of cell size through an adder mechanism (Campos et al., 2014b). They isolated the progeny of the cells after each cell cycle which enabled the measure of birth length and length at division without synchronization of the cell cycle. Hence, they could demonstrate that the cells elongate to a constant amount before dividing, irrespective of the birth length. Further measures of individual cells support this model of an adder strategy to maintain cell size in the bacterial world (Si et al., 2019; Taheri-Araghi et al., 2015).

In the yeasts *Schizosaccharomyces pombe* and *Sacharomyces cerevisiae*, the co-regulation of growth and division has also drawn particular attention. Early on, the fission yeast *S. pombe* has been described to follow a perfect size control (Fantes, 1977). Size of *S. pombe* at division varies extraordinarily little compared to other organisms with a coefficient of variation (standard deviation/mean) at ~0.06 vs ~0.17 in *S. cerevisiae*. Multiple experiments where the cell cycle was perturbed led to alteration of the cell size. Within two cycles, *S. pombe* manages to restore its normal size (Wood and Nurse, 2013; Turner et al., 2012; Fantes, 1977; Fantes and Nurse, 1977; Nurse, 1975; Mitchison and Creanor, 1971). Those works suggest the

presence of a size-sensing mechanism for size homeostasis in *S. pombe*: in a sizer scenario, size corrections for cells born either too small or too large can occur within few cycles, as opposed to a timer or adder strategy in which size corrections are not that immediate. *S. pombe* has two size checkpoints: one occurring at the G1/S transition and one at the G2/M transition. As a result, cells can pause cell-cycle progression until they have reached a size compatible with division or DNA replication, respectively. In normal conditions however, cells are always large enough to pass the G1/S size checkpoint (Wood and Nurse, 2015). Similarly, in the budding yeast *S. cerevisiae*, small cells spend more time in G1 and rapidly catch up with initially larger cells (Johnston et al., 1977; Hartwell, 1974). This observation in *S. cerevisiae* presents a compelling case for a size requirement to complete the G1 phase. However, the simple vision of a unique size prerequire occurring in G1 has been debated (Talia et al., 2007; Lord and Wheals, 1983) and a new theoretical model has been established in *S. cerevisiae*, wherein a sizer in G1 is followed by a timer in S/G2/M (Heldt et al., 2018). Hence, because of the consensus that prevailed in *S. pombe,* namely a sizer strategy for cell size control, this organism remains today a favored model to study cell size control through coordination of growth and cell cycle progression.

1.2.2. How regulation of the cell cycle is coordinated with cell size

How can the molecular components of the cell cycle regulation play a role in size homeostasis? In principle, regulators of the cell cycle can control, at specific timing of the cell cycle, whether a cell can move on the next step of the cell cycle (Wood and Nurse, 2015; Barnum and O'Connell, 2014; Turner et al., 2012; Jorgensen and Tyers, 2004). These processes are called checkpoints and they often take place at transitions between two cell cycle phases, like at the G1/S and G2/M transitions. In budding yeast, the main size checkpoint is believed to occur in late G1, at Start, a point at which the cell commits to the cell cycle (Barnum and O'Connell, 2014; Turner et al., 2012). In fission yeast, two size checkpoints were characterized, one at the G1/S transition and the other at the G2/M transition (Wood and Nurse, 2015). When a cell does not meet the size requirement at these checkpoints, namely if it is too small, cell cycle regulators prevent the cell from proceeding to the next phase. Therefore, it is through the tuning of the cell cycle progression that the cells can achieve size homeostasis, at least in certain systems such as budding and fission yeasts. I will thus briefly recall the main actors of

such a regulation in these two organisms, insisting particularly on fission yeast as it is the model I used throughout my PhD.

1.2.2.a. The events required for passage through the size checkpoints

- In *S. cerevisiae*

The progression through the main cell cycle transitions is coordinated by the cyclin-dependent kinase (CDK) Cdc28 which activity rises and falls throughout the cell cycle. The commitment to the cell cycle in late G1, at start, is induced by three cyclins, Cln3 at first and then Cln2 and Cln1, which all bind to and activate Cdc28 (Johnson and Skotheim, 2013; Koch et al., 1996). The complex Cdc28-Cln3 inhibits the transcriptional repressor and cell cycle inhibitor Whi5 (Costanzo et al., 2004). The action of Cdc28-Cln3 on Whi5 thus releases the inhibition of the G1/S transition. This results in the activation of the downstream cyclins Cln2 and Cln1 that promote further Whi5 inhibition, hence creating a positive feedback loop for irreversible activation of Start. Additional pathways and effectors were characterized in the G1/S transition, involving other cyclin-CDK such as Pcl-Pho85 and chromatin regulators like the histone deacetylases (HDAC) Hos3 and Rpd3 (Huang et al., 2009; Wang et al., 2009). Briefly, Whi5 forms a negative complex with HDAC which is disrupted by the action of the Cdc28-Cln3 and Pcl-Pho85.

- In *S. pombe*

Like *S. cerevisiae*, the two major steps from G1 to S and from G2 to M are orchestrated by the single cyclin-dependent-kinase (CDK) called Cdc2, homologue of the budding yeast Cdc28 and the metazoan CDK1 (Nurse and Bissett, 1981). Cdc2 level is constant throughout the cell cycle, but its activity increases thanks to its ability to form bipartite complexes with cyclins expressed in a specific cell cycle phase such as Cig2 or Cdc13. Cdc2 activity is also regulated by kinases like Rum1 or Wee1, or phosphatase like Cdc25 (**Fig 1.2**) which define the phosphorylation levels of Cdc2 and thus its activity (Wood and Nurse, 2015). Briefly, at the beginning of G1, Rum1 inhibits Cdc2; this interaction is essential for delaying G1 progression when the cell does not meet the size requirement for entry into S, in condition of nitrogen limitation for instance in which the cell is too small (Labib and Moreno, 1996; Moreno and Nurse, 1994). In late G1, Cig2 expression increases so that its association with Cdc2 leads to the onset of S-phase in which

DNA replication takes place (Mondesert et al., 1996) (**Fig 1.2**). After completion of DNA replication, in early G2, Cdc2 interacts with Cdc13 which expression rises to reach a maximum in late G2 (**Fig 1.2**). However, in early G2, the complex Cdc2-Cdc13 is inhibited by the kinase Wee1 (Wood and Nurse, 2015; Gould and Nurse, 1989; Russell and Nurse, 1987) (**Fig 1.2**). In late G2, the cdr2-cdr1 pathway, which was kept inhibited by the kinase Pom1 in early G2, inhibits Wee1, allowing the Cdc25 phosphatase to activate Cdc2 (Wood and Nurse, 2015) (**Fig 1.2**). Hence, a *cdc25* mutant exhibits a delay of entry into mitosis and consequently a size increase at division (Zhurinsky et al., 2010; Nurse, 1975), as opposed to the *wee1* mutant that divides at half the expected size. Interestingly, in the absence of canonical regulators of the cell cycle such as cig2 or wee1, expression of cdc2 fused with cdc13 under the control of cdc13 is sufficient to drive the separate phases of the cell cycle (Coudreuse and Nurse, 2010). This ingenious genetic system illustrates the extreme plasticity of the cell cycle where the major events are coordinated by multiple layers of regulation that are not all required for a functional cell cycle.

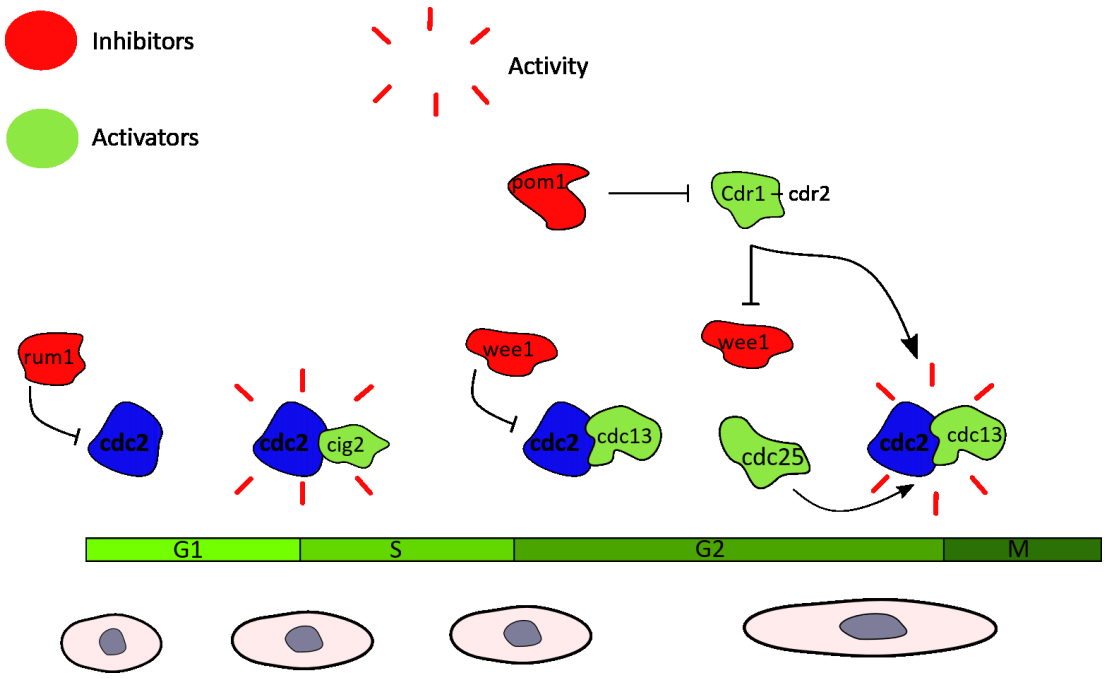


Figure 1.2: Simplified model of the key regulators of the cell cycle progression. *Cdc2 is the master regulator of the cell cycle and is required for the G1/S and G2/M transition. At the beginning of G1, Rum1 inhibits Cdc2 activity. The onset of S phase is permitted by the activity of the Cdc2-Cig2 complex in late G1. In early G2, Cdc2 forms a complex with Cdc13 which is inhibited by Wee1. Finally, in late G2, Wee1 is inhibited by the cdr1-cdr2 pathway, itself negatively regulated by Pom1, and Cdc25 activates Cdc2 for the G2/M transition (Wood and Nurse, 2015).*

Therefore, with several decades of investigation, the knowledge about cell cycle regulation, mainly at the size checkpoints, dramatically expanded. This constituted a solid basis to understand the coordination between cell size and cell cycle and, ultimately, the mechanism uncovering cell size control. Importantly, the cell cycle regulators were examined through a new angle: could they be used by the cell in order to sense its size?

1.2.2.b. The idea of a molecular size sensor

- In *S. cerevisiae*

The idea of a sensing mechanism for cell size control has caused much ink to flow (Wood and Nurse, 2015). In *S. cerevisiae*, different non-mutually exclusive models of size sensing for cell size homeostasis were developed. In the first model, the G1/S transition appears to rely on a protein dilution mechanism (Schmoller et al., 2015). More precisely, the amount of the cell cycle inhibitor Whi5 is cell size independent in a way that its concentration decreases as the cell grows. Therefore, small daughter cells have a higher Whi5 concentration than large daughter cells. Consequently, they need to grow more to pass the G1 checkpoint and enter S phase, thus maintaining cell size homeostasis over time. However, this model was challenged by the observation that Whi5 concentration does not decrease with size but instead stays constant (Dorsey et al., 2018). Another model describes the titration of the cell cycle activator Cln3 against nuclear sites and how this could control cell size. Specifically, the increasing Cln3 binds more and more to the promoter of its target genes. After all nuclear sites are filled, free Cln3 can phosphorylate and inhibits Whi5, which leads to commitment into the cell cycle (Heldt et al., 2018; Wang et al., 2009). However, this model alone cannot account for cell size homeostasis as deletion of both Cln3 and Whi5 does not lead to cell size homeostasis loss. A more recent study completes this view by swapping the promoters of whi5 and cln2 (Chen et al., 2020). By doing so, they finally managed to broaden the distribution of size in this population. Altogether, these data reveal possible mechanisms of cell size sensing through uncoupling of cell components (promoters, proteins) and cell size. Another convincing possibility is the induction of Start through different scaling of the proteome to cell size in G1 (Litsios et al., 2019). This differential scaling translates into large pulses of Cln3 translation that would lead to cell cycle commitment. All the models described here are relevant, but their

variety reflects the absence of a consensus for a mechanism through which the cell can measure its size. Therefore, the identity of a potential cell-size sensor remains elusive.

- In *S. pombe*

In *S. pombe*, two size checkpoints were identified at the G1/S and the G2/M transitions. Since cells are, in general, large enough to pass the G1/S size checkpoints, the major size corrections occur through G2 length adjustment, before the cells commit to mitosis (Wood and Nurse, 2015). But again, how do cells measure their size and respond accordingly by pausing or shortening the cell cycle? As mentioned above, the entry into mitosis is controlled by the balance between Cdc25 and Wee1 that activates or inhibits Cdc2, respectively. Then, a very straightforward model for cell size control comes from Keifenheim and co-workers, who characterize these two proteins, Wee1 and Cdc25, as key players of cell size control (Keifenheim et al., 2017) (**Fig 1.3**). The authors reported the size-dependent augmentation of Cdc25 concentration while Wee1 concentration stays constant. This means that small cells do not express enough Cdc25 to counteract the inhibitory effect of Wee1 on the G2/M transition. Hence, the inhibition of the mitotic entry is relieved only after the cells have grown passed a certain threshold, which provides a clever way to control cell size before division.

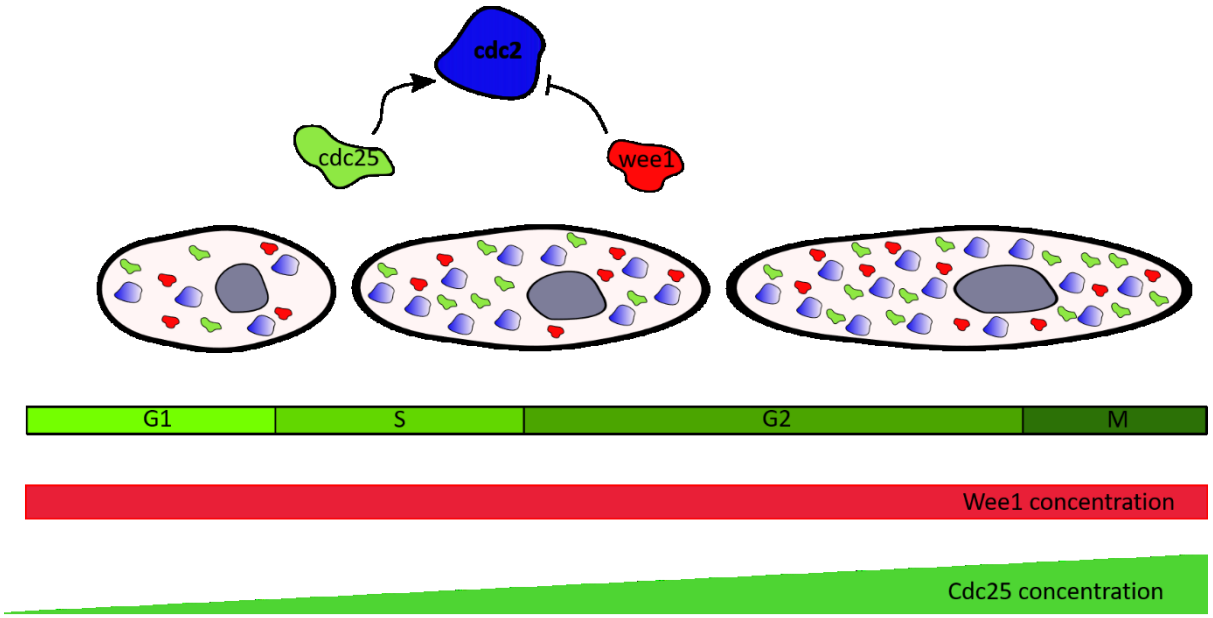


Figure 1.3: Model for cell size control at division by the size-dependant accumulation of cdc25, *adapted from Keifenheim and colleagues (Keifenheim et al., 2017).*

Notably, the initial search for a molecular sensor addressed effectors in the pathway downstream of Wee1 and Cdc25. It is the cell cycle inhibitor Pom1 that was first believed to measure the cell dimensions through a gradient mechanism (Martin and Berthelot-Grosjean, 2009; Moseley et al., 2009). Pom1 is an inhibitor of the G2/M transition as it negatively regulates Cdr1 and Cdr2, two kinases responsible for Wee1 inhibition. Moreover, Pom1 localizes at the tips of the cells (Bähler and Pringle, 1998). When the cell elongates, the concentration of Pom1 at the center of the cell decreases, enabling mitotic entry at a suitable size (Martin and Berthelot-Grosjean, 2009; Moseley et al., 2009). However, like in *S. cerevisiae*, this model has been challenged by the evidence that the deletion of pom1 does not affect size homeostasis, demonstrating the existence of a size control in this mutant (Novák, 2013; Wood and Nurse, 2013). But Pom1 and particularly its specific localization in the growing cells continued to seduce researchers and recent studies placed again pom1 as a main actor in coordinating the size with the entry into mitosis (Allard et al., 2019; Gerganova et al., 2019). Nevertheless, since the model for cell size sensing by Pom1 was disproved, other ideas proposed a size sensing strategy based on the local accumulation of other proteins. Pan and colleagues show that instead of Pom1, fission yeast cells monitor their size through Cdr2 (Pan et al., 2014) that, in addition to being a wee1 negative regulator, localizes at the membrane periphery (Martin and Berthelot-Grosjean, 2009) (Fig 1.4). As the cell increases in size, Cdr2 accumulates more and more in nodes or clusters at the medial cortex of the cell. The density of the nodal Cdr2 is thus a way for the cell to probe its size, more specifically its surface area (Pan et al., 2014). Later studies reveal that the Cdr1-Cdr2 contained at the nodes subsequently recruits and inactivates Wee1 in a burst-like manner (Allard et al., 2018) (Fig 1.4). In small cells, Pom1 suppresses the localization of Wee1 to the nodes. While the cell grows in size, the burst of Wee1 localization to the nodes increases, relieving the Cdc2 inhibition by Wee1. Upon the mutation of cdr2, the cell does not rely on its cell surface sensing anymore but either on its volume upon complete cdr2 deletion or on its length when the phosphorylation site for a Cdr2 activator is mutated (Facchetti et al., 2019). Therefore, the later work reveals the existence of redundant mechanisms for measuring cell size.

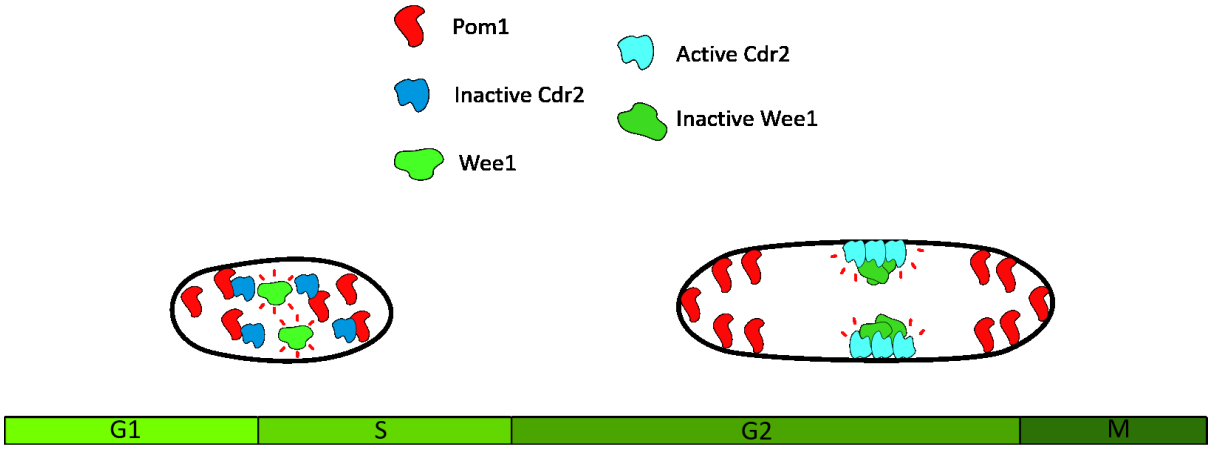


Figure 1.4: Cell size sensing and progression of the cell cycle with pom1. *In small cells, Pom1 negatively regulates Crd2 which allows Wee1 activity. Upon cell size increase, Pom1 localizes at the tip of the cell, relieving Cdr2 inhibition. Wee1 is therefore recruited at cortical nodes formed by Cdr2 which prevents its activity (Allard et al., 2018; Pan et al., 2014; Martin and Berthelot-Grosjean, 2009; Moseley et al., 2009).*

Like in budding yeast, the realm of cell size control in fission yeast is rich. Besides, as mentioned above, researchers succeeded in identifying the likely players of cell size coordination with entry into mitosis. The field describes a model wherein the induction of mitosis is signaled by a very specific spatial arrangement of these key players (Allard et al., 2018; Facchetti et al., 2019; Pan et al., 2014; Martin and Berthelot-Grosjean, 2009). Surely, because of the redundancy of the mechanisms controlling G2/M transition, the field still leaves space for new models. For instance, a screen of 82% of fission yeast non-essential genes identifies pathways involved in the G2/M transition and, consequently in the cell size control (Navarro and Nurse, 2012). This study thus provides compelling candidates for further research about cell size homeostasis. Among them, genes having a role in chromatin structure regulation, namely snf5, sol1, and sgf73, led to a reduced size at division when deleted. This last finding illustrates that even chromatin structure could be acting in cell size homeostasis.

1.3. How gene expression is coordinated with cell size

1.3.1. The concept of molecules scaling with cell size

Cell size varies but concentration of RNA and proteins are believed to remain constant (Marguerat and Bähler, 2012). Consequently, RNA and protein numbers also vary with size (Vargas-Garcia et al., 2018; Kempe et al., 2015; Marguerat and Bähler, 2012; Pritchard and Schubiger, 1996) (**Fig 1.5**). In large cells, the decoupling of cellular volume and protein synthesis impairs the good functioning of the cell and causes senescence (Lanz et al., 2021; Neurohr et al., 2019). On the contrary, an abnormal high protein concentration leads to subsequent "supergrowth" in the fission yeast (Knapp et al., 2019). This suggests that the homeostasis of gene-product concentration is essential for the stoichiometry of the biochemical reactions happening inside the cell.

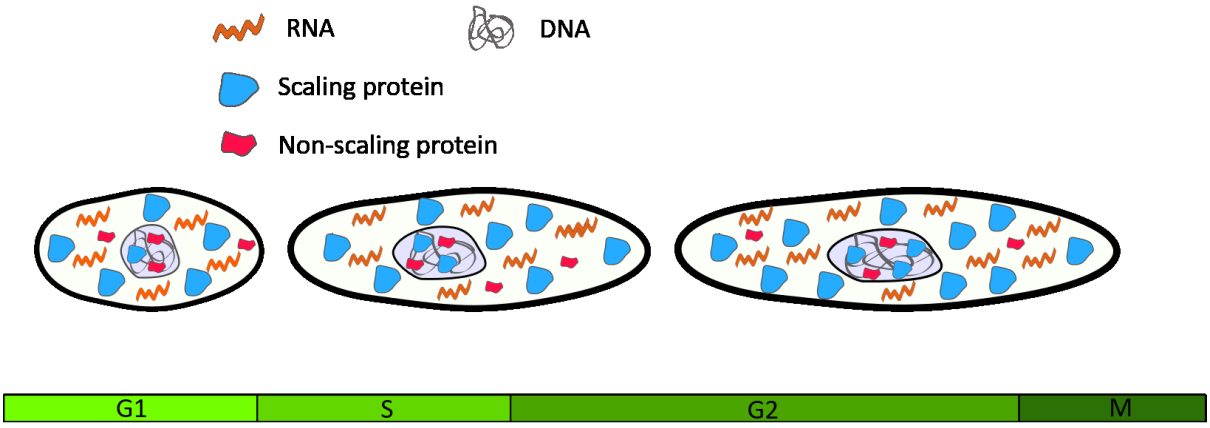


Figure 1.5: Most biomolecules scale with cell size *(from Marguerat and Bähler, 2012).*

In the context of cell size variation, this phenomenon of homeostasis of gene-product concentration is called the scaling of biomolecules to cell size (Marguerat and Bähler, 2012). Biomolecules scaling to cell size has been deeply examined in various eukaryotic species. To study how biomolecules such as RNAs were regulated as a function of cell size, experimentalists took advantage of the natural variation of the average cell size existing between rat tissues (Schmidt and Schibler, 1995). They witnessed that the RNA-to-DNA ratio correlated well with cell types. In unicellular eukaryotes, early studies explore the coordination of RNAs and

proteins with cell size, particularly in the context of the cell cycle (Elliott, 1983; Elliott and McLaughlin, 1978). More precisely, pulse labeling in synchronous cultures pointed to an increase in mRNA, rRNA and protein synthesis rate throughout the cell cycle (Elliott, 1983; Elliott and McLaughlin, 1978; Fraser and Nurse, 1978). Altogether, these examples suggest that cell size and production of RNA and proteins are linked. However, they are not tackling the genome-wide aspect of gene expression regulation with cell size: are all genes complying to the rule of scaling to cell size or is this phenomenon only targeting few genes? A key study in fission yeast revealed that the total gene expression of cells entering mitosis at smaller or larger size differ from wild type-sized cells' (Zhurinsky et al., 2010). Specifically, *cdc25* and *wee1* mutants, that divide into small and large size respectively, have been used to show that the average amount of most mRNAs in a population is proportional to cell size: lower than WT (Wild Type) in *wee1* and higher in *cdc25*. This illustrates the ability of the genome to respond globally to changes in cell size. More recent articles describe the scaling of mRNA, this time at the single cell level, using technologies permitting the observation of individual cells and molecules, namely single cell RNA-seq, microscopy and fluorescence *in situ* hybridization (FISH) that consists in visualizing molecules *in situ* by fluorescent probes. In mammalian cells, the use of single-molecule counting, and single-cell images illuminate the correlation between RNA number and cellular volume (Kempe et al., 2015; Padovan-Merhar et al., 2015). Similarly, FISH in the plant *Arabidopsis thaliana*, an elegant study shows that the regulation of the FLOWERING LOCUS C (FLC) expression is cell size dependent (Ietswaart et al., 2017). Finally, the scaling of transcripts to cell size in single cell was observed consistently in *S. pombe* (Sun et al., 2020). It is noteworthy that the nucleus scales with the total cell volumes (Sun et al., 2020).

The emergence of such a pervasive phenomenon, namely the global transcription scaling to cell size, attracted much attention and considerable scientific efforts are still deployed to uncover the molecular mechanisms behind it (Müller et al., 2021). Moreover, proteins that are not following this global mechanism and which their concentration changes with cell size are good candidates for acting in the control of cell size, showing that coordination of gene expression to cell size is intertwined with the quest of finding the regulation of cell size (Chen et al., 2020; Gu and Oliferenko, 2021; Martínez Segura, 2017).

1.3.2. The regulation of scaling through transcription rate

1.3.2.a. The potential origin of scaling

The coordination of RNA number with cell size can originate from the modulation of three components: an increase in the DNA content which is the template for RNA synthesis, a higher transcription rate and/or a slower RNA decay (Wilusz and Wilusz, 2004).

Firstly, genome size has been often considered as a direct constraint to cell size. Consistently, ploidies correlate with cell size in many organisms (Marguerat and Bähler, 2012; Galitski et al., 1999) as well as mRNA numbers (Sun et al., 2020; Wu et al., 2010; Galitski et al., 1999). In addition, in mice and human myocytes, the number of nuclei scales with cell volume in a way which suggests a role of DNA content in limiting cell size (Hansson et al., 2020). When fission yeast cells are arrested in G2 with a drug that targets a mutated Cdc2 protein (Aoi et al., 2014; Dischinger et al., 2008), they continue growing until they reach a plateau at approximately four times their WT size (Martínez Segura, 2017). In addition to cell size, the production of mRNA and proteins also reaches a limit. These data suggest that the genome is not able to sustain more growth through transcription unless it undergoes another round of replication (Martínez Segura, 2017; Zhurinsky et al., 2010).
These observations also show that cells possessing the same genome content can grow in size (e.g. the cdc2 mutant in G2) while maintaining RNA concentration. Likewise, in the *wee1* and *cdc25* mutants mentioned earlier, the mRNA homeostasis is achieved from the same genome size, by respectively increasing or decreasing the total RNA number. Moreover, the different cell types in metazoan (e.g., neurons vs lymphocytes) also provide an example of size and genomic content uncoupling: the 7 µm lymphocytes and the 1 m neuron both contain the same DNA content (Marguerat and Bähler, 2012). Altogether, these examples rule out the hypothesis whereby a change in DNA content is uniquely promoting scaling of RNA number to cell size.

Hence, transcription rate and RNA decay are the remaining candidates for driving such a cellular event. Indeed, both processes have been described to act in balance for controlling gene expression in accordance with the cellular needs (Ashworth et al., 2019).

On the one hand, RNA decay consists in the degradation of RNA by endo- or exonucleases. RNA decay can occur in the nucleus as well as in the cytoplasm and involves many pathways and components such as a super structure called the exosome (Garneau et al., 2007). On the other hand, the transcription rate is defined by the speed at which the transcription machinery is loaded onto the DNA and at which it transcribes RNA from DNA. The RNAP is a fascinating complex composed of twelve subunits in Eukaryotes. Although the number and composition of its subunits differ (e.g., five subunits in *E. Coli* vs 12 in *S. pombe*), the RNAP is found across all species and the process of transcription is widely conserved. Also, diverse types of RNAP exist depending on the RNA species to transcribe: mRNA and long non-coding transcripts are generated by the RNAP II whereas other non-coding RNAs such as rRNA and tRNA are transcribed by the RNAP I and III, respectively. Because proteins are produced from mRNAs, I will focus below only on the RNAPII.

Different studies try to unravel the respective implication of transcriptional activity and RNA degradation in the scaling of mRNA numbers. The early studies mentioned above indicate that scaling involves mRNA production rather than RNA decay (Schmidt and Schibler, 1995; Elliott, 1983; Elliott and McLaughlin, 1978). Then, with incorporation of radiolabeled adenine in newly synthesized RNA and the observation of a constant decay between *wee1*, *cdc25* and WT strains after transcription inhibition, Zhurinsky and co-workers showed that the accumulation of RNA is likely to emerge from RNA synthesis rather than RNA decay (Zhurinsky et al., 2010). In line with this, they noticed that the RNAPII occupancy on the DNA also scales with cell size, being lower in *wee1* than *cdc25* mutant cells. This observation agrees with recent findings in budding yeast where RNAPII occupancy is proportional to cell size (Swaffer et al., 2021a). Back in *S. pombe*, a study using the *wee1* and *cdc25* strains combines mathematical modelling and the comparison of the single-molecule FISH signal between WT, *wee1* and *cdc25* strains after transcription inhibition (Sun et al., 2020). The authors noticed that the half-lives of each probed mRNA are similar in the three strains. In mammals, the same experiment, consisting in measuring RNA half-lives after transcription inhibition, shows that RNA half-lives are too short to match the increase in cell size (Padovan-Merhar et al., 2015). Similarly, a constant RNA decay has been identified in plants for a transcript that exhibits cell-to-cell variability according to cell size (Ietswaart et al., 2017). These findings are pointing towards a modulation of the

transcription rate for coordination of gene expression with cell size for various species. Therefore, these findings mean that RNA number scaling is regulated more at the level of RNA production rate rather than RNA decay. Other recent findings in human cells and budding yeast reveal that mRNA decay plays in mRNA amount scaling to cell size (Berry et al., 2021; Swaffer et al., 2021b). The authors uncover a mechanism wherein mRNA decay could indeed help the mRNA molecules to scale with cell size.

In the three organisms described above, the studies characterize the step of the transcription process that would be responsible for the transcription scaling to cell size (Sun et al., 2020; Ietswaart et al., 2017; Padovan-Merhar et al., 2015). Before exposing their discoveries, it would be helpful to remind the reader of the process of transcription, particularly the transcription initiation, and the different modes of transcription.

1.3.2.b. The process of transcription

The transcription process can be divided into three phases: initiation, elongation, and termination. Each of these phases is finely tuned and constitutes an important rate-limiting step of the transcription. For example, transcription initiation can be regulated through the recruitment of the RNAPII on the regulatory region upstream of the gene, the promoter (Sainsbury et al., 2015). This first interaction between the RNAPII and the DNA requires the general TFs (GTF) TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH that form with the RNAPII the pre-initiation complex (PIC) (Haberle and Stark, 2018; Louder et al., 2016). TFIID is an important complex containing proteins having specific functions. Hence, TFIID can form modules with the other GTFs, mainly IIA, IIB and IIC (Chen et al., 2021) thus stabilizing and assisting the loading of the PIC onto the DNA. Furthermore, the TFIID complex shares subunits with a well-known co-activator of transcription, the SAGA (Spt–Ada–Gcn5 acetyltransferase) complex (Timmers, 2021). Both of these complexes interact with the TBP (TATA-box binding protein) which recognizes a particular motif found in the core promoter: the TATA box. Historically, all genes were thought to possess this motif, but subsequent evidence reveals the unexpected diversity of the core promoters (Haberle and Stark, 2018; Li et al., 2015a; Juven-Gershon and Kadonaga, 2010). In fact, the TATA box is more likely to be found on the promoter of stress or inducible genes rather than growth or constitutive genes that exhibit instead a TATA-like motif (Haberle

and Stark, 2018; Rando and Winston, 2012; Basehoar et al., 2004). Consequently, these two categories of gene (having a TATA box or a TATA-like motif at the promoter) do not encompass the same mode of transcription initiation (Chen et al., 2021). Initially, it was believed that TATA-box promoters depended on the activator SAGA (Spt–Ada–Gcn5 acetyltransferase) whereas activation of TATA-less promoters mostly coincide with TFIID binding (Rhee and Pugh, 2012; Basehoar et al., 2004; Huisinga and Pugh, 2004). However, in budding yeast, deletion of components belonging to the SAGA or the TFIID complex highlight the contribution of both complexes in the expression of nearly all genes (Baptista et al., 2017; Warfield et al., 2017; Huisinga and Pugh, 2004). Furthermore, a study from Donczew and colleagues characterizes two sets of genes: one being sensitive to the rapid depletion of both TFIID and SAGA and the other being sensitive to the rapid depletion of TFIID alone (Donczew et al., 2020). Besides, none of these categories were particularly enriched in the TATA motif, showing the subtle nature of transcription initiation. In metazoan, the situation is even more complex since such organisms need to deal with spatially confined transcription in tissues or specific transcription for the development of those tissues (Lenhard et al., 2012). Therefore, they possess different co-activators such as TFs involved in tissue development like the hox genes for example (Jain et al., 2018). In addition, their cis-regulatory elements encompass larger genomic regions (Lenhard et al., 2012). Immense research was conducted to investigate the process of transcription initiation in Metazoan but, because this thesis is about fission yeast, I am not detailing the literature on this matter here.

Finally, to form a stable complex with the DNA, the transcription machinery also must face a highly organized structure found in the eukaryotic nucleus: the chromatin (Kornberg and Lorch, 2020; Bai and Morozov, 2010; Kornberg and Lorch, 1999). The chromatin, composed of specific proteins bound to the DNA, is responsible for the packaging of the DNA into the nucleus. Its role in transcription regulation is widely acknowledged, as I will report later in the introduction. In addition to TATA and TATA-less motif, other classifications of genes' promoter exist and are based, for example, on chromatin architecture, an aspect that will be discussed later in the thesis (Haberle and Stark, 2018; Kubik et al., 2015; Rach et al., 2011). Therefore, transcription initiation is regulated at several levels: through co-activators, cis element on the promoter and chromatin structure.

1.3.2.c. The concept of transcriptional noise

Although RNA transcription is subjected to substantial control, stochasticity in gene expression also finds its place in this complex process, as chemical reactions happening in the cells are also stochastic. This means the level of a single transcript varies from cell to cell (Raj and van Oudenaarden, 2008). Although transcription stochasticity, or noise, is controlled (Bahar Halpern et al., 2015; Battich et al., 2015; Raser and O'Shea, 2005; Rao et al., 2002), this cell-to-cell variability in gene expression can confer a competitive advantage to the overall population (Raj and van Oudenaarden, 2008; Shahrezaei and Swain, 2008; Kepler and Elston, 2001), reflected by growth itself (Shahrezaei and Marguerat, 2015). For example, in environments of fluctuating resources, cells having different metabolic statuses in the population could show a better adaptability thereby providing a fitness advantage to the population. In this regard, a complete work on fission yeast, combining single-cell RNA-seq and novel computational analysis, reveals the association of gene expression heterogeneity with distinct phenotypes such as cell size or growth (Saint et al., 2019). In addition, several works corroborate that some of the cell-to-cell variability in transcripts abundance can be explained by variation in cell size (Sun et al., 2020; Ietswaart et al., 2017; Kempe et al., 2015; Padovan-Merhar et al., 2015). This shows how noise in a population of isogenic cells is somehow intertwined to the scaling of transcripts to cell size. But cell size is not the only source of gene expression stochasticity in a cell population. Modelers, hand in hand with experimentalists, showed that many transcripts are produced in bursts, where the gene activation occurs at irregular intervals, alternating between an active and inactive state (Raj and van Oudenaarden, 2008; Golding et al., 2005). This contrasts the view of a steady-state or Poisson-like RNA synthesis which is theoretically less compatible with high transcriptional noise as opposed to transcriptional bursting (Raj and van Oudenaarden, 2008; Suter et al., 2011). However, the finding of "bursty" transcription does not exclude a Poissonian mode of transcription that still occurs for many genes (Sun et al., 2020; Zenklusen et al., 2008; Raj and van Oudenaarden, 2008).

1.3.2.d. The regulation of mRNA scaling

How does transcription rate mediate the scaling of mRNAs to cell size?
Firstly, the discovery of transcription scaling was followed by the characterization of the nature of transcription involved in this process: Poisson-like or in burst? With RNA intron probing by

FISH, Padovan-Merhar and colleagues confirm the burst mode of transcription for four genes (Padovan-Merhar et al., 2015). The transcription bursts can be characterized by their size (the number of mRNA produced during a single burst) or their frequency (how often the gene is active). Padovan-Merhar and co-workers observed that the intensity of the nascent transcripts was higher in larger cells. This suggests that the modulation of the transcription burst size leads to scaling of these mRNA to cellular volume. Using similar parameters for modelling transcription bursting (Padovan-Merhar et al., 2015; Shahrezaei and Swain, 2008), Ietswaart and colleagues describe the dynamics of the FLOWERING LOCUS C (FLC) expression in a cell population of heterogeneous size (Ietswaart et al., 2017). As opposed to what was found in mammals, they notice that the FISH-probed FLC mRNA did not fit a model of cell-to-cell variability through a "bursty" transcription. Instead, the size-dependent expression FLC transcripts is consistent with a Poissonian transcription. The recent work in *S. pombe* from Sun and colleagues recapitulates the previous finding (Sun et al., 2020). Indeed, by measuring different transcripts by single molecule (sm)FISH, they reveal that transcription scaling is mediated by a Poisson-like transcription initiation. Then, they infer the transcription mode by fitting five different models divided into two classes: either in the "bursty" transcription model or the Poisson-like one. They conclude that transcription rates of constitutive and cell-cycle specific genes are Poissonian and are scaling with cell size. However, genes that are expressed following certain stimuli are transcribed in bursts, although in this case, it remains unclear whether scaling is caused by the modulation of the burst size or the burst frequency. This bimodal transcription in *S. pombe* is reminiscent of what was found in budding yeast in which poissonian transcription limits the variation in the expression of certain genes whereas "bursty" transcription for other genes results in higher variation in their expression (Zenklusen et al., 2008). Last but not least, thanks to modelling and ChIP-seq (Chromatin Immuno-Precipitation and sequencing) experiments against different forms of the RNAPII, Sun and coworkers confirmed that an increase in RNAPII initiation rate couples transcription scaling to cell size, a finding supported by similar experiments in budding yeast (Swaffer et al., 2021b). As mentioned above, this recent study by Swaffer and colleagues, together with other works (Berry et al., 2021) also indicate a role in mRNA stability in adjusting the mRNA numbers to cell size in situations wherein transcription initiation would fail to do so, like when the proportion between RNAPII activity and cell size is not respected (Swaffer, et al. 2021).

Now, how can transcription initiation be adjusted to cell size?

Interestingly, DNA content does not follow the rule of homeostasis as the number of DNA molecules only doubles in S phase. Nevertheless, an appropriate range of DNA-to-cytoplasmic ratio must be maintained across species (Turner et al., 2012). Early work already suggested that cell size was reflected by the RNA:DNA ratio which is indeed higher in large cells (Schmidt and Schibler, 1995; Sato et al., 1994; Fraser and Nurse, 1979, 1978). Therefore, in a situation where cell size increases while DNA content remains constant (in G1 or G2 phase for instance), the DNA-to-cellular volume ratio could act as a signal for transcription initiation scaling. Padovan-Mehrar and colleagues explore this idea by fusing a large fibroblast cell to a small melanoma cell expressing a protein fused to GFP (Padovan-Merhar et al., 2015). Firstly, they observed that the expression of the GFP mRNA increased to match the larger volume of the fused cell. This demonstrates that the cell can integrate information about the cellular volume and respond accordingly by increasing the mRNA synthesis, rather than growing in response to the mRNA increase. Hence, the authors suggest the existence of a diffusible factor that can transmit information about cell volume to the DNA in order to adapt the transcription rate. Interestingly, the concentration of the GFP mRNA is 50% less in the heterokaryon cells compared to the single nucleus melanoma cells, suggesting that the scaling of the GFP mRNA relies on the two nuclei in the heterokaryon cells. Similarly, in both diploid *S. cerevisiae* and *S. pombe* cells, gene expression in growing cells relies on the two alleles of the genes to achieve maintenance of mRNA concentration (Sun et al., 2020; Schmoller et al., 2015). Altogether, these data show that the putative factor responsible for transcription scaling is able to integrate information about both cellular volume and DNA content.

It was previously hypothesized that this effect was produced by an increased concentration of transcription factors in the nucleus, which, when combined with a high DNA affinity, could act as a driving force to globally coordinate gene expression and cell size (Marguerat and Bähler, 2012). This mechanism, which could set the measures for transcription scaling, is highly reminiscent of the titration model whereby budding yeast achieve cell size control *via* the titration of Cln3 by the DNA (Heldt et al., 2018; Wang et al., 2009). However, rather than a TF, Sun and colleagues proposed that RNAPII itself as a potential volume sensor that can coordinate cell size and gene expression (Sun et al., 2020). Indeed, they found that RNAPII

concentration increases in the nucleus when cell size increases. Consequently, RNAPII binding sites are occupied in a cell size-dependent fashion which results in a likewise size-dependent transcription initiation rate. As it was discussed elsewhere (Lin and Amir, 2018; Padovan-Merhar et al., 2015), they also propose that the RNAPII is limiting to achieve such a task. Besides, this idea is supported by recent works from Swaffer and colleagues wherein the authors found that RNAPII alone and not the other PIC subunits are limiting for RNAPII recruitment to the promoters (Swaffer et al., 2021b). Although both findings could fit with the titration model whereby the limiting factor is titrated against the genome, Swaffer and colleagues imagine a more complex model of mass action dynamic equilibrium in which genome bound and unbound RNAPII coexist in the nucleus. More precisely, they show that RNAPII is not exactly proportional to cell size and cannot solely account for transcription scaling which also requires adjustment of mRNA stability as cell size increases.

To conclude, the mechanism underlying the coupling of cell size to transcription has proven to be complex. Nevertheless, thanks to deep research in this field, exciting findings have emerged. In summary, the homeostasis of mRNA concentration in a context of cell size variation is due to the adjustment of transcription initiation rate (Fig 1.6). This modulation might take place *via* a mechanism that involves the titration of a factor, expressed in a size-dependent manner. In budding and fission yeasts, this factor is the RNAPII itself: the RNAPII can transmit information about cellular volume to the genomic content that competes for size limited RNAPII (Sun et al., 2020).

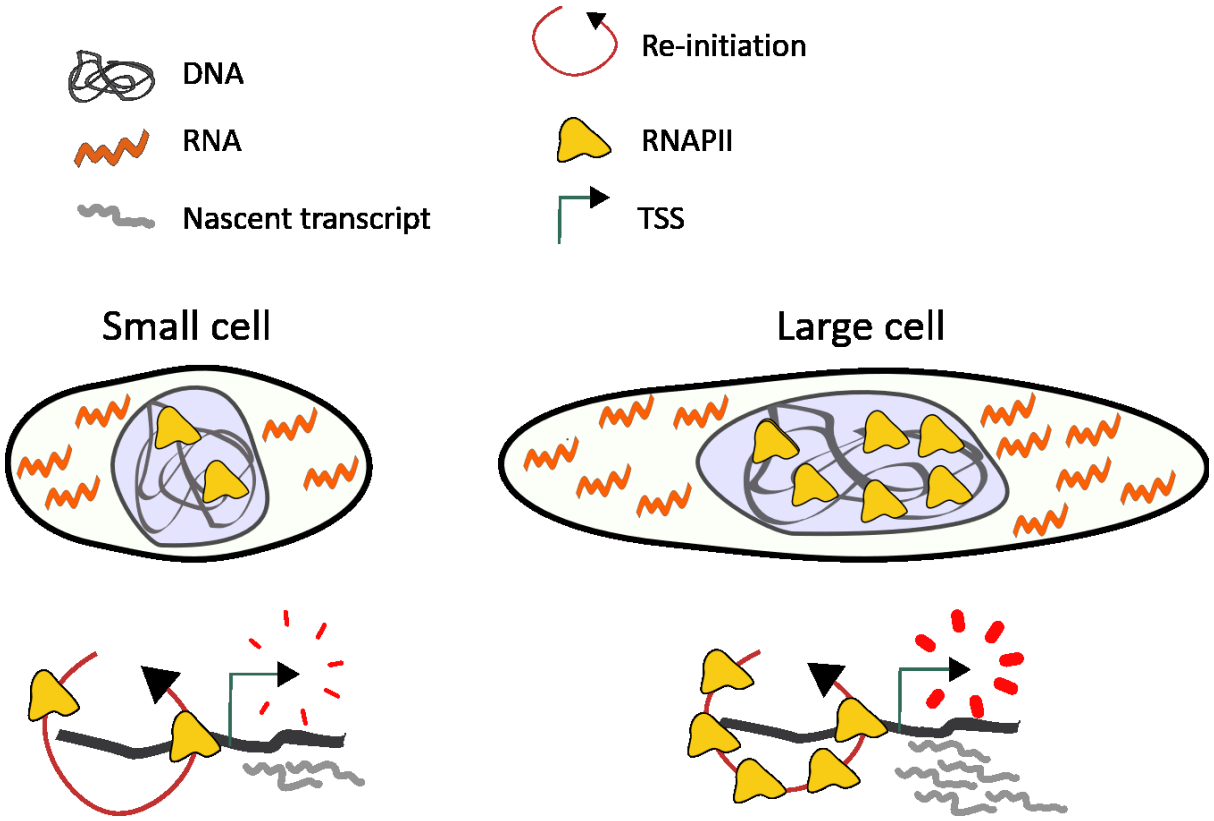


Figure 1.6: Increase in RNAPII initiation rates mediates scaling in large cells *(from Sun et al., 2020)*

As with DNA, some proteins/transcripts escape the global regulation that most genes undergo while cell size changes (Lanz et al., 2021; Chen et al., 2020; Keifenheim et al., 2017; Martínez Segura, 2017; Schmoller et al., 2015; Zhurinsky et al., 2010). These factors are drawing a lot of attention as they are potential cell size sensors acting for cell size homeostasis (Gu and Oliferenko, 2021). One category of such components has been identified multiple times: histones, the proteins forming the basic unit of chromatin (Claude et al., 2021; Lanz et al., 2021; Swaffer et al., 2021a; Martínez Segura, 2017). In fact, those proteins function in close relation with the DNA and since the DNA molecule is not scaling with cell size, it is not surprising that histones do not scale with cell size. Can it be that the chromatin structure is also participating in setting transcription scaling to cell size? This is an attracting hypothesis as the RNAPII directly cooperates with the chromatin to achieve the complex process of transcription. This raises a side question: can chromatin architecture explain how certain proteins escape the global regulation for scaling?

1.4. Chromatin implication in gene concentration homeostasis

In Eukaryotes, the DNA is wrapped around a set of well-known proteins, the histones (Kornberg, 1974). The histones are organized in octamers of four core histones: H2A, H2B, H3 and H4. Each histone octamer wraps ~150 bp of DNA with remarkably high affinity, thereby forming a globular structure called the nucleosome. The nucleosome, as well as other factors functioning in relation with the DNA, are forming the primary chromatin structure. It is noteworthy that chromatin can be structured into higher levels of organization with long-range contacts between and within chromosomes. The latter aspect will not be tackled in this thesis wherein the focus is on the primary structure of chromatin.

1.4.1. The nucleosome: the basic unit of chromatin

Strikingly, the nucleosomes are not randomly placed on the DNA. On the contrary, they are extremely organized, and this feature is widely conserved among species. Even in non-eukaryotic species such as archaea, evidence about the existence of organization of the DNA around histone-like proteins starts to emerge (Stevens et al., 2020; Hocher et al., 2019). The first hint for such a beautiful arrangement of the chromatin was obtained from X-ray diffraction of the DNA in complex with histones, which gives specific regular patterns (Wilkins et al., 1959). Subsequently, the visualization of nuclei from various species with electron microscopy revealed the existence of spherical particles that resemble “beads on a string” (Woodcock et al., 1976; Olins and Olins, 1974). Finally, this peculiar structure is also encountered when treating the nuclei with a nuclease, such as the MNase (micrococcal nuclease) that can be extracted from *Staphylococcus aureus* (Zaret, 1999; Luzzati and Nicolaïeff, 1963). This enzyme has an endo- and exonuclease activity and can thus digest all free DNA fragments that are not protected by the binding of chromatin proteins (Fig 1.7). After treating a cell population with MNase (refer to chapter 2 for detailed description of MNase) and analyzing the resulting DNA, experimentalists witness a striking ladder-like arrangement of the DNA (Baldi et al., 2020). Each ladder rung is a piece of DNA that reflects the linker DNA and the DNA protected by one nucleosome for the first rung, two nucleosomes for the second, three nucleosomes for the third and so forth (Fig 1.7). When increasing the MNase digestion time or when raising the MNase concentration, ~80% of the DNA is restrained in a ~150 bp band (Hewish and Burgoyne,

1973). This is explained by the cleavage of the DNA by the MNase on both sides of a nucleosome, namely the linker DNA, and at a frequency that depends on the genomic loci as well as on the extend of the digestion. Altogether, these early findings depict the arrangement of the chromatin into a constant repeating unit, the nucleosome.

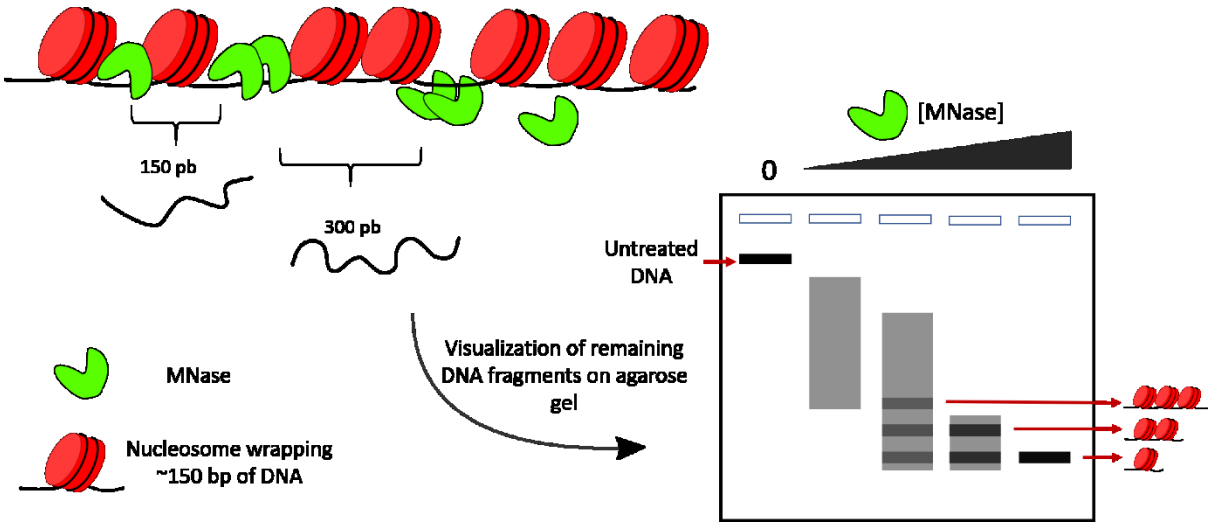


Figure 1.7: Size of DNA fragments after MNase digestion. *DNA is protected by nucleosomes, the chromatin unit wrapping ~150 bp of DNA. After DNA digestion by MNase, fragments of different sizes can be detected, depending on the MNase digestion as well as the state of chromatin (compact, open).*

1.4.2. The explosion of genome-wide nucleosome maps

Nowadays, the organization of the primary chromatin structure can be examined genome-wide through many different angles thanks to ever more powerful innovative technologies (Minnoye et al., 2021; Tsompana and Buck, 2014). These technologies consist in revealing either the DNA pieces that are accessible or the footprint on the DNA of the chromatin components, mainly the nucleosomes. In the first category, we can cite the DNAse-seq approach which consists in sequencing the DNA fragments that were accessible and consequently cut by a nuclease, the DNAse I; the ATAC-seq that uses the hypersensitive Tn5 transposase that simultaneously fragments and links adapters to the accessible DNA; the NOM-seq and dSMF that both uses methyltransferases that methylates the CpG or GpC when DNA is accessible (Krebs et al., 2017; Kelly et al., 2012). In the category of the foot printing method, the MNase-seq, a particularly appreciated technique, combines the above-mentioned MNase with a deep-sequencing of the resulting protected DNA (Tsompana and Buck, 2014). After

mapping the DNA fragments resulting from the digestion, an average nucleosome profile is generated by aligning the signal from all the genes on their TSS. Again, one will appreciate the regular and repetitive nature of the chromatin: a well-organized array of nucleosomes emerges from the average signal on the gene body. This is manifested by the appearance of oscillations starting from the TSS of the gene, leaving the promoter of the genes depleted in nucleosomes (Kelly et al., 2012; Lantermann et al., 2010; Mavrich et al., 2008a) (**Fig 1.8**). The height and width of the oscillation peaks respectively define the occupancy and the positioning of the nucleosomes (Baldi et al., 2020; Struhl and Segal, 2013) (**Fig 1.8**). On the one hand, the occupancy indicates the frequency to which a nucleosome is attached to a given locus. On the other hand, the positioning is related to the locus occupied by a nucleosome. A well-positioned nucleosome means that all the cells in a population exhibit a nucleosome at a specific locus. These two metrics have proven to be relevant in the race for describing the chromatin architecture in various species and conditions. An additional metric, the spacing of the peaks, reflecting the spacing of the nucleosomes from dyad to dyad (or nucleosome center), is also used when defining the property of nucleosome organization (**Fig 1.8**). The spacing is a direct measure of the nucleosome repeat length (NRL) which length differs depending on the cell type, the genomic location but also the species (Baldi et al., 2020; Lieleg et al., 2015). Fission yeast, for example, is known to possess a shorter DNA linker compared to budding yeast (Moyle-Heyrman et al., 2013). This might be due to the absence of the linker histone H1 in this organism.

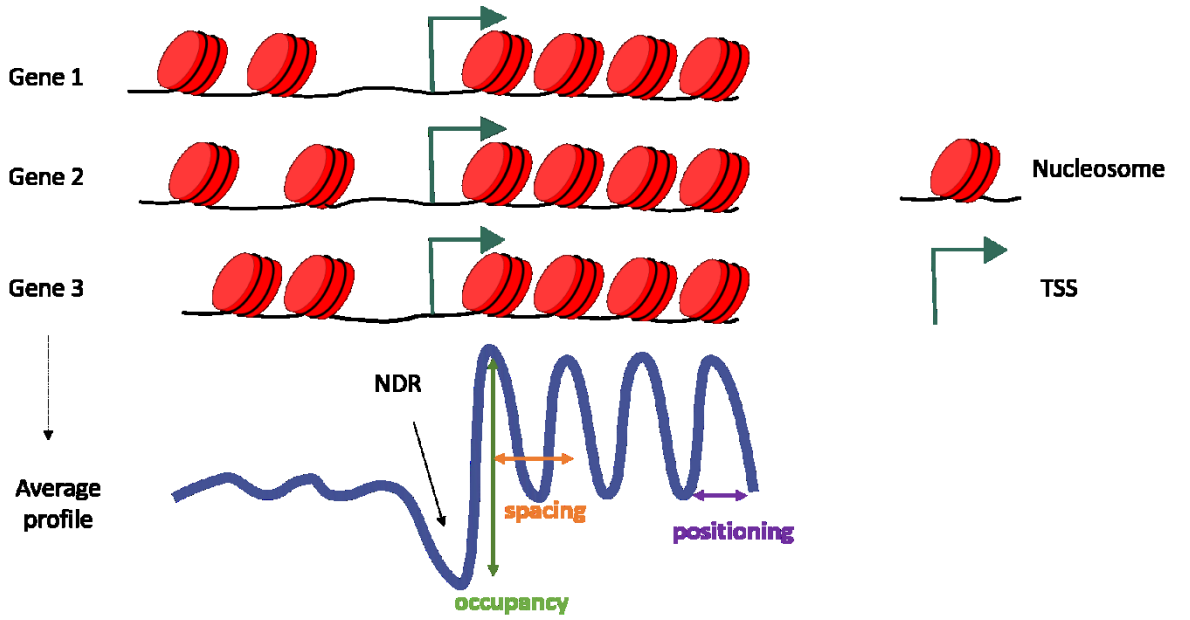


Figure 1.8: Pattern of nucleosome organization around the TSS. *After MNase digestion, the signal around the TSS reflects the presence of nucleosomes. The average signal of all the genes (bottom) can be calculated to visualize the nucleosome organisation of the average gene. Different metrics can be extracted from the average profile: occupancy (the frequency to which a nucleosome is bound to a given locus), spacing (the space between two nucleosome centre or dyad, defining the NRL) and the positioning (which DNA locus is occupied by the nucleosome).*

Therefore, the use of MNase has provided genome-wide nucleosome maps for many species and many conditions. Nevertheless, the correct interpretations of these maps have been a long-standing task as it requires a good grasp of the kinetics of the MNase reaction in order to find the best experimental conditions.

1.4.3. The kinetics of the MNase digestion

Historically, nucleosome maps were generated with the mononucleosome-sized fragments that represent most of the signal after a single MNase digestion (Lantermann et al., 2010; Rando, 2010; Shivaswamy et al., 2008; Barski et al., 2007; Yuan et al., 2005). At the time, it was granted that the selected mononucleosomal fragments represented almost the totality of the nucleosomes on the genome, thus being a fair sampling for the elaboration of an accurate nucleosome map. However, the presence of a ladder-like pattern after light MNase digestion (Fig 1.7) suggests that the genomic locations are not equally sensitive to the MNase. While some parts of the DNA are only fragmented into large fragments, others are reduced at a mononucleosome size: this is the reflection of the respective resistant and sensitive chromatin

at different loci. More importantly, high MNase digestion can further digest the resistant chromatin, leading to a majority of mononucleosomal fragments. In this scenario, the fragments from the previous sensitive nucleosomes are completely degraded and will thus never be sequenced (Chereji et al., 2019). On top of that, MNase prefers A/T rich fragments that will be destroyed faster than G/C fragments (Chereji et al., 2019; Hörz and Altenburger, 1981). Technical variability between nucleosome maps inevitably arose from the intrinsic nature of the chromatin as well as the MNase sequence specificity (Xi et al., 2011; Chung et al., 2010). Even a slight variation of MNase was enough to muddy the waters as they had considerable effects on the position and the occupancy of the nucleosomes. To answer this problem, different methods were developed to yield reproducibility (Chereji et al., 2019; Mieczkowski et al., 2016; Rizzo et al., 2012, 2011).

The new areas of research for improving the accuracy of nucleosome maps were unexpectedly fruitful as they paved the way for the discovery of new chromatin states. Indeed, some of these protocols propose to map the nucleosomes obtained after a light digestion. In this process, scientists started to see nucleosomes turning up on the genome at places that were initially considered nucleosome-depleted, such as at the promoter. These findings led to the controversial notion of "fragile nucleosomes" around the TSS (Kubik et al., 2017; Jeffers and Lieb, 2017; Chereji et al., 2016; Vera et al., 2014; Xi et al., 2011; Weiner et al., 2010; Henikoff et al., 2009; Lee et al., 2007). The debate includes two different points of view: one side states that the fragile nucleosomes are non nucleosomal particles like chromatin remodelers or TFs (Chereji et al., 2017, 2016; Vera et al., 2014), whereas the other side stipulates that they are instead composed of histone proteins forming unstable nucleosomes (Kubik et al., 2017, 2015; Xi et al., 2011; Rhee and Pugh, 2012; Jin et al., 2009). Although the question remains open, those two views can coexist in a model where those fragile nucleosomes represent sub-nucleosomes bound to GRFs (General Regulatory Factors) or the remodeling factors such as RSC (remodeling the structure of chromatin) (Brahma and Henikoff, 2019; Kubik et al., 2015; Pradhan et al., 2015; Floer et al., 2010) (Fig 1.9). This is a relevant concept for the present work and further details about these curious particles (the "fragile" nucleosomes) will be exposed throughout the thesis.

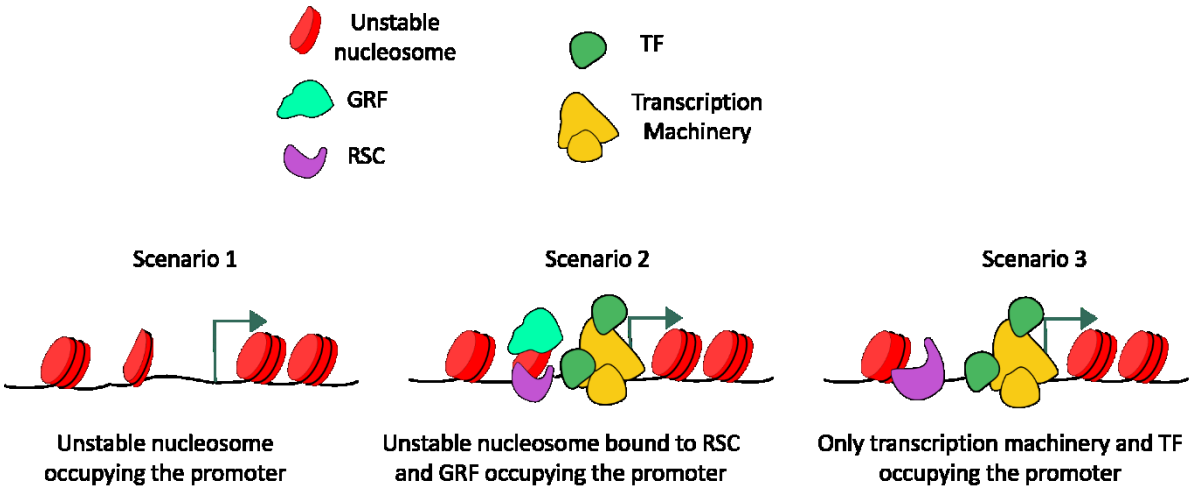


Figure 1.9: The concept of "fragile" nucleosome. *Scenario 1, 2 (Kubik et al., 2015; Brahma and Henikoff, 2019) and 3 (Chereji et. Al., 2017) each represent a conformation of promoter architecture that would generate signal upon mild MNase digestion.*

Hence, the use of MNase permits to go beyond the canonical nucleosome organization. By adapting the conditions of digestion, this technique can yield all proteins in interaction with the DNA and thus reveal different chromatin states at regulatory regions such as the promoter.

1.4.4. The regulation of the nucleosome organization

The multiple faces of the MNase digestion permitted to acquire more understanding of the nucleosome organization. Researchers hypothesized that such conserved and pervasive organization of the nucleosome on the DNA is established by redundant regulatory mechanisms, which would render the identification of the direct regulators challenging. A major *in vitro* study in budding yeast investigates the direct contributions of the ATP-dependent chromatin remodelers involved in the maintenance of the nucleosome-depleted state of the promoter, the positioning of the first nucleosome after the TSS, namely the+1 nucleosome and the regularity of the downstream array of nucleosomes from the 5' end of the gene (Krietenstein et al., 2016). Consistent with *in vivo* works (Brahma and Henikoff, 2019; Ganguli et al., 2014), the authors show that the NDR is formed by RSC that recognizes the poly(dA)/(dT) track on the promoter, a function that is conserved in fission yeast (Yague-Sanz et al., 2017). Moreover, RSC prevents the +1 nucleosome to occlude TBP binding sites, facilitating the loading of the PIC on accessible DNA (Kubik et al., 2018). Interestingly, a slight

downstream shift of the -1 nucleosome was previously noticed when chromatin structure was reconstituted *in vitro* without ATP-dependent factors, suggesting that a nucleosome is present on the promoter by default (Korber and Hörz, 2004). ISW2 and/or ISWI1 are responsible of the positioning of the +1 nucleosome (Krietenstein et al., 2016). These two events might be supported by general regulatory factors (GRF), such as Rap1 or Reb1, or even replaced by INO80 (Krietenstein et al., 2016; Bakel et al., 2013). Finally, ISW2 and INO80 form the regular downstream array and ISW1a adjust the spacing of the nucleosomes to set the canonical NRL (Krietenstein et al., 2016). These events are sufficient to generate the basic primary structure of chromatin around the TSS. However, it is noteworthy that they normally occur inside the nucleus in which a plethora of other components can exert redundancy and modulation of the nucleosome organization. For example, the regularity of the nucleosome can be perturbed, going from well-positioned or phased nucleosome to disorganized or fuzzy nucleosomes, depending on the genomic regions (Baldi et al., 2020). In addition, it is generally acknowledged that transcription can play a role in the arrangement of the nucleosomes onto the DNA (Jiang and Zhang, 2021; Mieczkowski et al., 2016; Lantermann et al., 2010; Weiner et al., 2010; Lee et al., 2007; Shivaswamy et al., 2008). For example, it was suggested that the directionality of the nucleosome array was set by the transcription (Lantermann et al., 2010; Zhang et al., 2009). More importantly, extensive research depicts the nucleosome as a key player in transcription regulation (Kornberg and Lorch, 2020; Rando and Winston, 2012; Li et al., 2007), as I will cover in the next part.

1.4.5. The nucleosome in the context of transcription

Nucleosomes locally occlude DNA and thus determine the spatial access to the genome by the transcription machinery (Kornberg and Lorch, 2020; Rando and Winston, 2012; Li et al., 2007). Early *in vitro* studies show that the presence of a nucleosome on the promoter prevents the transcription initiation by the RNAPII (Workman et al., 1988; Knezetic and Luse, 1986; Lorch et al., 1987). This was confirmed *in vivo* by the observation that nucleosome loss at the RNR3, PHO5, CYC1 and GAL1 promoters activates the corresponding genes (Zhang and Reese, 2007; Han and Grunstein, 1988). Consistently, upon artificial decrease of nucleosome occupancy at the promoter, a substantial number of genes is de-repressed, and their transcription increases (Gossett and Lieb, 2012). Today, numerous evidence corroborates the correlation between

transcription and DNA accessibility (or low nucleosome occupancy) at the promoter (Nguyen et al., 2021; Mieczkowski et al., 2016; Nabilsi et al., 2014; Kelly et al., 2012; Lantermann et al., 2010; Bryant et al., 2008; Boyle et al., 2008; Lee et al., 2007, 2004). Taken together, these data point toward an inhibitory role of the nucleosome at a crucial regulatory region of the gene, the promoter.

Additionally, the tumor suppressor MLH1 is abnormally silenced in various cancers, and this might be due to the formation of a nucleosome on its promoter and the subsequent loss of the NDR at this locus (Lin et al., 2007). Moreover, SNF5 and BRG1, two subunits of the ATP-dependent chromatin remodeler SWI/SNF, have tumor suppressor activity (Wilson and Roberts, 2011). Their inactivation in primary mouse cells leads to an overall decrease of nucleosome occupancy at promoters which apparently induce the over-expression of genes involved in the cell cycle progression (Tolstorukov et al., 2013). These last examples highlight that the chromatin organization at the promoter can, if deregulated, profoundly affect the health of the cell.

Since the NDR at the promoter is formed by chromatin remodeling complexes such as RSC (Brahma and Henikoff, 2019; Kubik et al., 2018; Krietenstein et al., 2016; Ganguli et al., 2014; Floer et al., 2010; Korber and Hörz, 2004), one could expect that all genes exhibit, by default, a nucleosome on their promoter which is destabilized by chromatin remodeling complexes like RSC to facilitate the PIC formation and subsequent transcription (Kornberg and Lorch, 2020; Kubik et al., 2018). This model is supported by the increased RSC occupancy at promoters of highly transcribed genes (Ganguli et al., 2014). Nevertheless, the promoters of specific genes such as the yeast RIO1 or AKY2, that comprises a particular cis-element distinct from the TATA-box motif, is be kept free of nucleosome without additional trans-factors (Angermayr et al., 2003, 2002). Such natural properties of the DNA sequence in the promoter enable a basal and constitutive transcription of these genes, uncovering a new class of TATA-less promoter for constitutive genes. In fact, chromatin architecture at the promoter of constitutive genes is different from chromatin architecture in stress genes. In yeast, the promoter of stress genes is often packaged into delocalized nucleosomes under repressive conditions whereas constitutive genes exhibit an obvious NDR on their promoter (Rando and Winston, 2012; Weiner et al., 2010; Ioshikhes et al., 2006). As mentioned in 1.3.2.b (“the process of transcription”), the two categories differ in their pattern of expression: while constitutive

genes are highly expressed with little cell-to-cell variation, stress genes are, on the contrary, inducible, expressed in specific time windows and in a noisier manner (Kornberg and Lorch, 2020). Curiously, the dichotomy between chromatin architecture at stress and constitutive genes is also associated with transcription plasticity (Sharon et al., 2014; Tirosh and Barkai, 2008). On the one hand, the transcription of constitutive genes can be modulated by TFs, yet within a limited range. On the other hand, a high occupancy of nucleosome at the promoter positively correlates with a high transcription plasticity, making the class of genes that possess such chromatin structure, namely the stress genes, more responsive to specific stimuli (Tirosh and Barkai, 2008). Moreover, it was suggested that the noisy expression of the stress genes was due to their particular chromatin structure on the promoter because of the competition between TFs and nucleosomes for binding the promoter (Sharon et al., 2014; Rando and Winston, 2012; Tirosh and Barkai, 2008). Therefore, the packaging of the promoter into a nucleosome does not only silence the gene; it is also a strategy for stress genes to be transcribed in a wide range of expression, as opposed to constitutive genes that possess a clear NDR at this locus and which are thus more continuously transcribed within a narrower range of expression.

Competition between GRF, such as Rap1, and nucleosome has been described as a main determinant of transcription regulation of growth and highly expressed genes (Mivelaz et al., 2020; Lickwar et al., 2012). More precisely, stable binding of Rap1 mediates efficient downstream transcription (Lickwar et al., 2012) and depletion of Rap1 lead to a nucleosome formation at its binding site (Kubik et al., 2015). Kubik and colleagues describe a model wherein GRF binding at the promoter of highly expressed and growth genes does not completely remove the nucleosome (Kubik et al., 2015; Knight et al., 2014). Instead, GRF bind to nucleosomal DNA, and, in collaboration with RSC, destabilizes the nucleosome, thereby establishing a “fragile” nucleosome at the promoter (Mivelaz et al., 2020; Brahma and Henikoff, 2019; Kubik et al., 2015; Knight et al., 2014). (Fig 1.10). Although histones were detected at the “fragile” nucleosome of both *S. pombe* and *S. cerevisiae* (Moyle-Heyrman et al., 2013; Xi et al., 2011), the precise composition of the “fragile” nucleosomes remains mysterious; whether they contain heavily modified histones, enrichment with the histone variant H2A.Z or whether they are missing one or several core histones, the debate is ongoing (Fig 1.10).

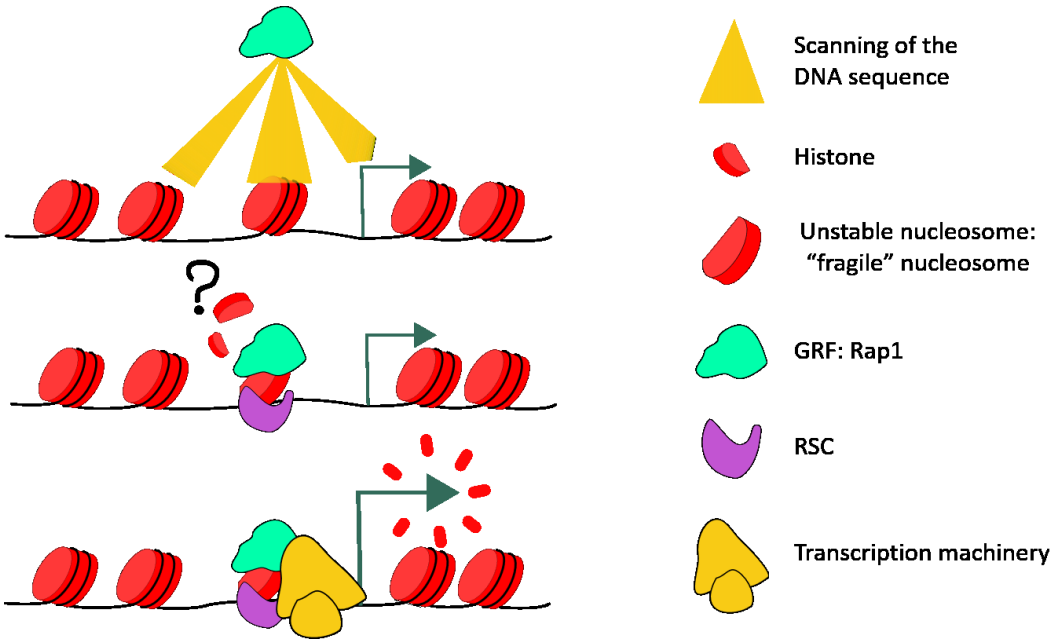


Figure 1.10: Establishment of a "fragile" nucleosome by the GRF Rap1 and chromatin remodeler RSC *(Mivelaz et al., 2020; Pradhan et al., 2015). GRFs such as Rap1 are able to scan the DNA at the surface of nucleosomes and engage their target site despite the presence of a nucleosome. This specific ability of Rap1, in cooperation with RSC activity, lead to the formation of a "fragile" non-canonical nucleosome. The exact composition of "fragile" nucleosomes is still unknown.*

Furthermore, GRF, chromatin remodeling complexes and histone chaperones also control transcription fidelity, by preventing aberrant transcription (Murawska et al., 2020; Challal et al., 2018; DeGennaro et al., 2013; Kaplan et al., 2003). For example, the depletion of Rap1 gives birth to inappropriate transcription initiation (Challal et al., 2018). In addition, nucleosome perturbation, induced genome wide by Spt6 mutation, leads to aberrant transcription, with increased levels of intragenic, intergenic, and antisense transcripts (DeGennaro et al., 2013; Kaplan et al., 2003). Likewise, a recent study in *S. pombe* shows that the histone chaperone FACT (facilitates chromatin transcription) controls the bona fide transcription of the genes by repressing their corresponding antisense transcription (Murawska et al., 2020). In budding yeast, FACT and Spt6 were shown to work together to prevent aberrant transcription by restraining the accumulation of the H2A variant H2A.Z into coding regions (Jeronimo et al., 2015). Hence, in addition to the negative regulation of the gene, nucleosomes also act in the accuracy of the transcription (Fig 1.11).

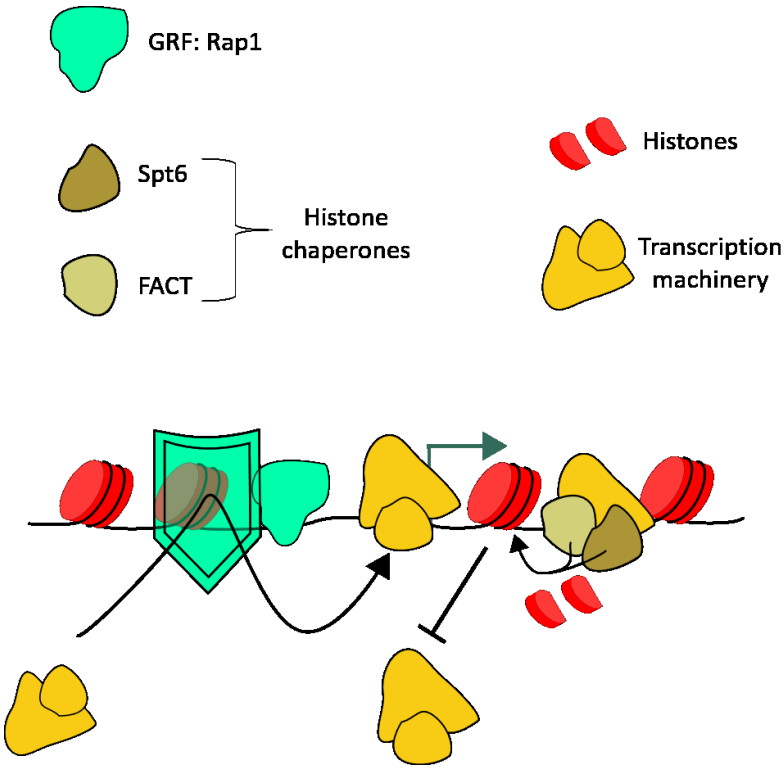


Figure 1.11: GRFs and histones chaperones work together to control transcription accuracy through nucleosomes *(Jeronimo 2015, Murawska 2020, DeGennero 2013, Challal 2018). GRFs, such as Rap1, prevent aberrant transcription from antisense strand and intergenic regions while histone chaperones like FACT and Spt6 reform the nucleosome in the wake of the RNAPII.*

In line with this, downstream of the promoter lays the well-positioned +1 nucleosome, which constitutes the first barrier encountered by RNAPII during transcription and which can compete with loading of the PIC (Nguyen et al., 2021; Rhee and Pugh, 2012). After incorporation of the H2A.Z is into the +1 nucleosome, the later becomes more permissive to RNAPII which can then overcome this barrier; this reveals the activating role of H2A.Z (Weber et al., 2014). However, other works associate incorporation of H2A.Z into the nucleosome with transcription repression (Giaimo et al., 2019), depicting the curious duality of H2A.Z functions. Furthermore, it was also described that the +1 nucleosome increased the ability of the RNAPII to transcribe the DNA template, compared to naked DNA (Kornberg and Lorch, 2020; Nagai et al., 2017), showing that chromatin can both potentiate and limit transcription. Consistently, the position of the +1 nucleosome right downstream of the TSS is key for TBP binding at the promoter and thus facilitate transcription initiation (Kubik et al., 2018).

On the gene body, the correlation between nucleosome organization and transcription is rather poor (Rando and Winston, 2012; Lantermann et al., 2010). This might be due to the

existence of a large group of factors facilitating the RNAPII passage through the nucleosome, thus maintaining the embedding of the coding region into a phased array of nucleosomes (Kulaeva and Studitsky, 2010; Kulaeva et al., 2009; Li et al., 2007). Upon transcription, histone chaperones, such as Spt5 or FACT, handle histone exchanges to allow the passage of the RNAPII (Evrin et al., 2022; Farnung et al., 2021; Venkatesh and Workman, 2015). The nucleosomal structures are thus very dynamic and need to be replaced immediately to prevent transcriptional arrest or aberrant transcription (Venkatesh and Workman, 2015; Workman, 2006). In addition, a minimal exchange of histones is allowed by the formation of an intranucleosomal DNA loop which leads to progression of the RNAPII through a permissive nucleosome, in which only the H2A/H2B dimer is exchanged (Kulaeva et al. 2009a). In this way, nucleosome organization at coding regions can be maintained, thus preventing the production of aberrant transcripts (Venkatesh and Workman, 2015).

Altogether, extensive research about nucleosomes and transcription depicts the collaboration between nucleosomes and plentiful TFs to achieve the subtle tuning of transcription.

1.4.6. Histone modification and transcription

A very important part of the nucleosome is not in direct contact with the DNA: the histone tail (Kornberg and Lorch, 2020). Histone tails are subject to various post-translational modifications (PTM) that form an epigenetic code (Berger, 2007; Jenuwein and Allis, 2001). This code, on top of the well-known genetic code, can be deciphered by specific factors that in turn affect DNA metabolism such as transcription (Bannister and Kouzarides, 2011; Jenuwein and Allis, 2001). Unlike the genetic code that only operates with four chemical bases (adenine, cytosine, guanine, tyrosine), the epigenetic code comprises a vast collection of PTM which keeps broadening (Chan and Maze, 2020). Besides, the variety of PTM, including methylation, acetylation, acylation, ubiquitination or phosphorylation, can function alone or in combination, thus offering a complex network of possibilities (Taylor and Young, 2021; Chan and Maze, 2020). It is acknowledged that PTM can modify the properties of histones. For example, acetylation has been described as neutralizing the positive charges (Waterborg, 2011) that enhance the contact between the positively charged nucleosome and the negatively charged DNA (Hong et al., 1993; Norton et al., 1989). Early studies show that this can affect nucleosome

conformation (Norton et al., 1989; Allegra et al., 1987) and consequently, the access of the transcription machinery to the DNA (Vettese-Dadey et al., 1996; Lee et al., 1993). Indeed, the positive correlation between acetylation and transcription has been acknowledged for decades (Gates et al., 2017; Rando, 2007; Pokholok et al., 2005; Tjian and Maniatis, 1994; Hebbes et al., 1988).

The PTM on histone tails are not irreversible but they usually result from a competition between factors that establish these PTM (the "writers") and the ones that remove them ("erasers") (Chan and Maze, 2020; Strahl and Allis, 2000). Acetylation is deposited by histone acetyltransferases (HAT) and removed by histone de-acetyltransferases (HDAC). Gcn5 is a famous HAT as it belongs to the SAGA complex (Grant et al., 1997). Since SAGA is known to be an activator of transcription (Baptista et al., 2017; Warfield et al., 2017; Huisinga and Pugh, 2004), one can speculate the associated acetylation established by Gcn5 is enhancing transcription. In *S. cerevisiae*, Gcn5, along with the other proteins of the SAGA complex, targets the lysines 9, 14, 18 and 23 of histone 3 (H3K9-14-18-23ac) (Grant et al., 1999) and its deletion indeed causes a global diminution of the transcripts' levels (Huisinga and Pugh, 2004; Kristjuhan et al., 2002). Furthermore, H3K9ac and H3K14ac are found at the promoter of active genes (Rhee and Pugh, 2012; Pokholok et al., 2005). In *S. pombe*, H3K9ac, which also requires Gcn5, exhibits a similar pattern, peaking on the ATG of highly expressed genes (Sinha et al., 2006). This pattern seems to be conserved in multi-organisms (Gates et al., 2017; Karmodiya et al., 2012; Kratz et al., 2010; Zhou et al., 2010). In addition, Gcn5 is enriched over the coding regions of highly expressed genes in S. pombe (Johnsson, 2009). These observations hence convey the idea that the acetylation on H3K9 by Gcn5 is concomitant with transcription. Can we now establish a causality between the presence of H3K9ac and transcription? The fact that Gcn5 deletion leads to a diminution of global transcripts (Donczew et al., 2020; Warfield et al., 2017; Huisinga and Pugh, 2004), and that acetylation affects the nucleosome binding to the DNA (Norton et al., 1989) could point towards an enhancement of transcription initiation mediated by H3K9ac.

Moreover, after being written on the histone tail, the PTM are read by specific factors (**Fig 1.12**). In other terms, they have been examined as molecular platforms for the recruitment of numerous factors. In respect with the acetylation, *in vitro* and *in vivo* studies indicate that nuclear proteins containing bromodomains recognize acetylated histones, particularly on their

lysine (Kanno et al., 2004; Matangkasombut and Buratowski, 2003; Jacobson et al., 2000; Dhalluin et al., 1999). Subunits of the yeast's chromatin remodelers SWI/SNF contain a bromodomain able to recognize acetylated histones by SAGA (Strahl and Briggs, 2021), although in fission yeast, the SWI/SNF complex was described as both a repressor and activator (Monahan et al., 2008). Furthermore, the bromodomain proteins, such as BRD4 in mammals or bdf1 in yeast, function in close relation with transcription regulation (Donczew et al., 2020; Jang et al., 2005; Matangkasombut and Buratowski, 2003; Jacobson et al., 2000) and are known to exclusively recognize the acetylation motif (Filippakopoulos et al., 2012). In HeLaS3 cells, immunoprecipitation studies show the interaction between BRD4 and the P-TEFb complex, the latter being responsible for the pause release of the RNAPII at active genes (Jang et al., 2005). In yeast, Bdf1 has been found in complex with a TAF (TBP-associated factor) and suppression of this interaction leads to defect in gene expression (Matangkasombut et al., 2000). Another example in HeLa nuclear extracts depicts the release of the HAT p300 factor of the promoter due to an increasing acetylation level on the chromatin and on p300 itself. Since p300 competes with TFIID, p300 release leads to PIC assembly at the gene (Black et al., 2006). Again, these observations suggest that transcription activation could be mediated by acetylation on histone tails.

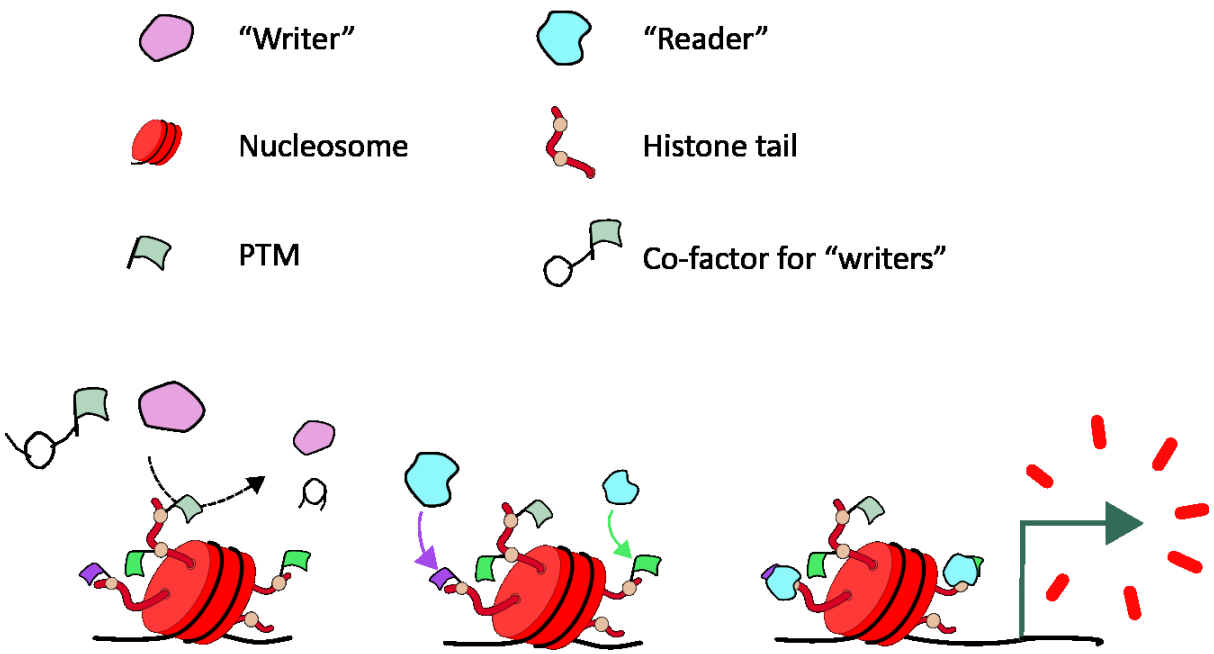


Figure 1.12: Example where PTMs on histone tails activate transcription. *"Writers" are complexes catalysing the PTM of histone tails while "readers" recognize the modified tails (Strahl and Allis, 2000).*

These last parts, dedicated to chromatin architecture, depict the extensive regulation of transcription by the basic unit of chromatin, the nucleosome. The link between chromatin and transcription is so strong that it seems impossible that one could operate without affecting the other. Before, I have presented the known evidence that transcription rate acts as a driving force for keeping the mRNA homeostasis in a growing cell. Hence, it is reasonable to place chromatin as a pivot in the process of transcription scaling.

1.4.7. Chromatin in the context of transcription scaling to size

What do we know about a potential cooperation between RNAPII and chromatin to support transcription scaling? It is fairly recent to incorporate chromatin in the scaling field; consequently, little is known on this matter.

On the one hand, one appealing hypothesis is that cell size increase is signaled to chromatin through metabolism, thereby adjusting transcription rate for transcription scaling. In fact, intermediates of the Tricarboxylic acid (TCA) cycle, such as acetyl-coA, are substrates for PTM writers like HAT (Santos, 2021; Reid et al., 2017). In line with this, progressive mitochondrial loss by artificial expression of a dominant-negative mtDNA polymerase leads to histone hypoacetylation in the nucleus (Martínez-Reyes et al., 2016). Importantly, supplying the cells with the genetic component of the TCA cycle restored acetylation, showing that establishment of histone acetylation requires the TCA cycle (Martínez-Reyes et al., 2016). Furthermore, a recent report indicates with the same system that mitochondrial loss is associated with significant lower H3K9ac enrichment at the promoter (Lozoya et al., 2019). Consistently, an elegant study depicts the activation of growth genes in response to carbon sources through histone acetylation with acetyl-coA (Cai et al., 2011). This means that acetyl-coA, being a substrate of Gcn5, can signal the metabolic state of the cell to the chromatin and induce the expression of the corresponding genes. Consistently, it has been observed that mitochondria proliferation occurred during *cdc2-asM17* elongation (Aoi et al., 2014). This might induce a higher production of ATP molecules in the cell. Since chromatin remodelers such as RSC or SWI/SNF use ATP to achieve local remodeling of the chromatin, one could hypothesize that the size increase might cause a higher activity in those chromatin remodeling complexes.

Altogether, these findings illustrate the link between growth, metabolic pathways, and the epigenome (Fig 1.13).

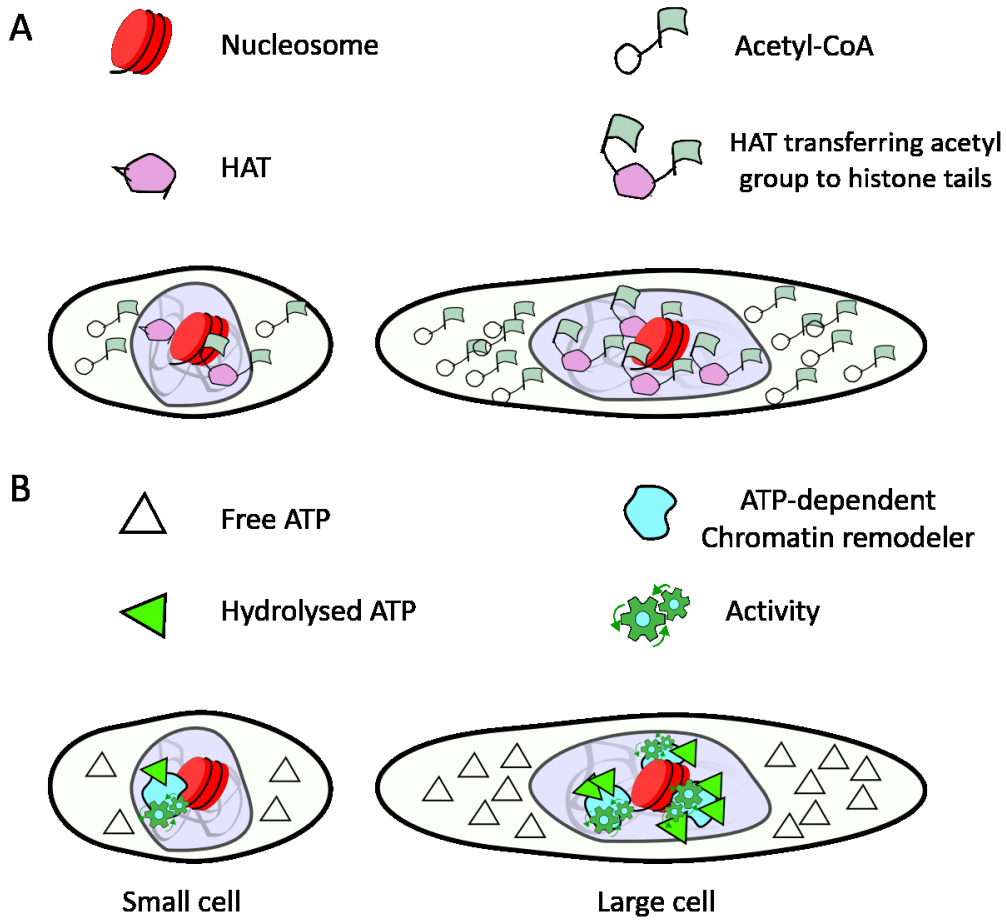


Figure 1.13: Large cells might signal their size to the genome through small metabolites such as ATP or acetyl-coA. A. *Acetyl-coA increases during growth (Cai et al., 2011) and can be used as co-factors by HAT for acetylation on histone tails (Michael A. Reid 2018, Janine Hertzog Santos 2021).* ***B.*** *Large* cdc2-asM17 show an increase in mitochondrial proliferation (Dischinger et al.,) which might be associated with a higher ATP production. ATP are hydrolyzed by chromatin remodelers providing the necessary energy for chromatin remodeling.

On the other hand, in a situation of size increase, it was shown that the number of histones, contrary to most proteins, were not coordinated with cell size (Claude et al., 2021; Swaffer et al., 2021a; Martínez Segura, 2017) but, instead, to DNA content (Claude et al., 2021). Consistently, histone proteins are usually produced during S phase, concomitantly to DNA replication (Eriksson et al., 2012; Takayama and Takahashi, 2007). In their work, Claude and co-workers analyze the promoters of the genes that produce the non-scaling histone transcripts. They conclude that the affinity of histone promoters for the limiting factor

responsible for transcription initiation is higher than other general promoters (Claude et al., 2021; Sun et al., 2020). Therefore, histone promoters rapidly reach saturation and are thus unable to respond to an increase in cell size, as opposed to the other general promoters. Beyond histone transcripts, this is an interesting model to understand how certain transcripts escape scaling to cell size. Besides, the link between non-scaling transcripts and promoters inspired other researchers who corroborate previous findings (Lin and Wang, 2021). More precisely, they show that transcripts like histones which do not increase with cell size, are enriched with specific motifs in their promoter that can affect the recruitment of RNAPII (Lin and Wang, 2021). Altogether, these results show that cell size does not affect the amount of the principal component of the chromatin, the histones. Consistently, a recent which used dSMF assay to probe chromatin accessibility suggests that the chromatin state mostly remains invariant as cell size changes (Swaffer et al., 2021b) and that the loading of the RNAPII to the genome depends mostly on the concentration of free RNAPII. The authors measured chromatin accessibility with a dSMF assay (Krebs et al., 2017) (mentioned in 1.4.2 "The explosion of nucleosome maps") and they did not notice significant changes. However, the four histone PTMs they measured, H3K4me2, H3K4me3, H3K79me3, H3K36me3, are increasing with cell size (Swaffer et al., 2021b). Amalia Martínez Segura from our group analyzed scaling versus non-scaling genes in the size-growing fission yeast strain *cdc2-asM17* (Martínez Segura 2017; Aoi et al. 2014). This mutant is sensitive to the addition of an ATP analogue (1NM-PP1) that impairs the functioning of the mutated cdc2. Consequently, with a specific concentration of 1NM-PP1, the cells arrest in G2, but they keep growing in size, leading to a synchronous population of increasing cell size (Fig 1.14). Interestingly, she observed that a cluster of genes containing prominent level of H3K9me2 were all positively non-scaling genes.

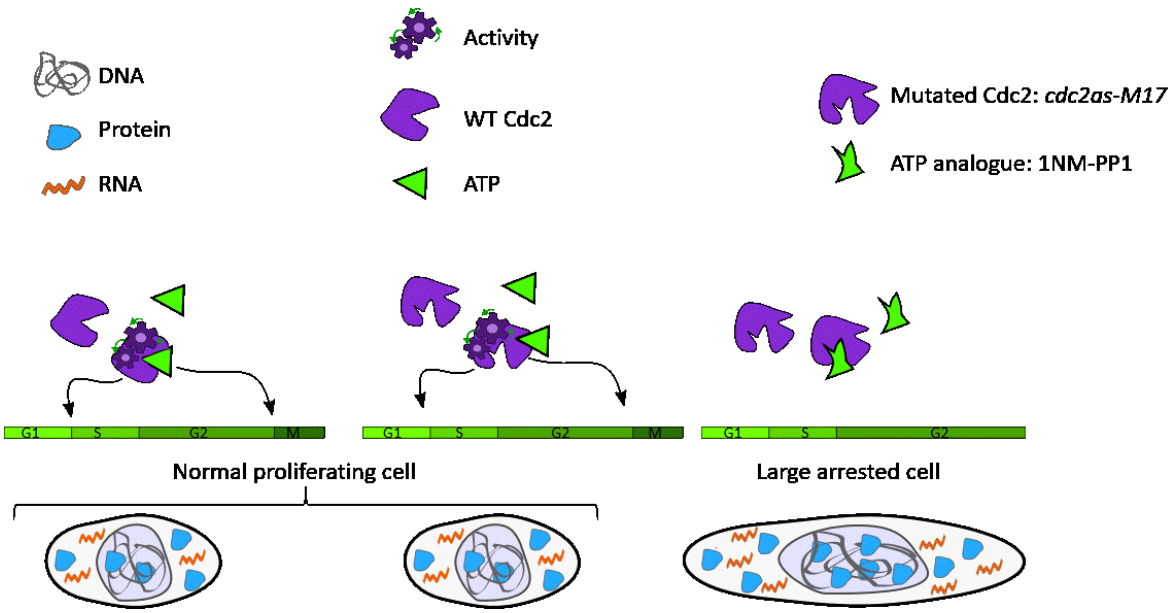


Figure 1.14: The* cdc2-asM17 *mutant (Aoi) is sensitive to the addition of 1NM-PP1 which targets Cdc2. *Addition of 1NM-PP1 in the culture leads to a cell population of increasing cell size.*

These last findings, together with those from Swaffer and co-workers (Swaffer et al., 2021b) raise the question about other size-dependent changes in histone PTMs, such as the previously described H3K9ac. If such chromatin marks change when size increases, this is likely to perturb chromatin architecture that in turn could become more permissive for transcription scaling.

In summary, research has recently shed light on the link between transcription scaling to cell size and histones, the core part of nucleosomes. However, how chromatin structure is changing with cell size is still unclear.

1.5. Aims of this thesis

My PhD studies answer to the following: "Does chromatin structure change with cell size and, if yes, how?". By exploring this question, I aim at deepening the understanding of the phenomena associated with cell size changes in fission yeast. In addition, this work will supply a new direction for future studies on transcription scaling regulation.

The aims of this thesis are:

- Define the global organization of chromatin structure in a fission yeast mutant growing in size (chapter 3).
- Characterize the chromatin on different set of genes in a WT situation and when cell size increases (chapter 4).
- Investigate the size-dependent expression of two histones (H3 and H2A) and one H3's PTM (H3K9ac) and the size-dependent occupancy of H3 on the DNA (chapter 5)

To this end, I am using several different approaches:

- The fission yeast mutant *cdc2-asM17* (Aoi et al. 2014) described hereafter in the method section.
- An MNase-seq approach to uncover the size-dependent changes in chromatin structure.
- A western blot and microscopy approach to measure the level of H3, H3K9ac and H2A abundances
- A ChIP-seq approach to measure the size-dependent enrichment of H3 on the DNA
- A computational approach to analyze the MNase-seq and the ChIP-seq data

Chapter 2: Material, methods, and characterization of the elongating cdc2-asM17

2.1. Experiments in the laboratory

2.1.1. Cell culture and strain

Strains were revived from glycerol frozen stocks on solid yeast extract agar (YE agar), or YE agar supplemented with 25 mg l-1 adenine, L-histidine, L-leucine, uracil, L-lysine (Sigma), and with or without antibiotics for selection. YE agar plates were incubated for approximately 72h at 32°C in a static incubator until visible large colonies could be observed. Single colonies were transferred into liquid yeast extract medium (YE) or in YE supplemented as above (YES) depending on the strain for a 20 ml C1 culture incubated overnight. The following morning, a 50 ml C2 culture is prepared in Edinburgh minimal medium (EMM) or YES for cdcasM17 and SBM320, respectively. At the end of the day, the optical density of the C2 culture should be at 600 nm (OD600) of ≈ 0.5. The C2IS then diluted to a volume that depends on the experiment, into either EMM (for *cdc2-asM17*) or EMMS (for SBM320): this is the C3 culture. Finally, the C3 culture is incubated overnight and harvested the following morning at around OD600 = 0.8.

- *cdc2-asM17*

The *cdc2-asM17* strain was optimized by Aoi and colleagues from a first version engineered by Dischinger and co-workers (Aoi et al. 2014; Dischinger et al. 2008). This mutant has a substitution of a single amino acid in the adenosine triphosphate (ATP)-binding pocket (called gatekeeper residue) of the Cdc2 protein and is thus sensitive to ATP-analogue molecules that are fitting into the active site of the mutated kinase. Since cdc2 is essential to the G2/M transition, *cdc2-asM17* cells cultivated with an ATP analogue are arrested in G2 but keep growing in size (**Fig1.14**). Our lab received it from Masamitsu Sato (MJ1358). At the time of reception, the *cdc2-asM17* had the following genotype: *h90 cdc2-asM17::bsdR ade6-M216 leu1-32 ura4-D18 sf1-CFP::NatMX*. The strain was later modified by Amalia Segura Martin and has now this genotype: *h- cdc2-asM17::bsdR*. During my PhD, the *cdc2-asM17* was grown in Edinburgh Minimal Media (EMM) at 32^{o}C overnight between OD600 = 0.8 and OD600=1. To create a pool of normal-sized cells, a volume of culture, depending on the experiment, was pelleted. Then, 2 µM of the ATP analogue 1NM-PP1 was added to the remaining culture and the rest of the culture was harvested every hour for 6 hours (**Fig 2.1a, c, and d**). Cells were

pelleted 3 minutes at 3000 rpm, washed and kept frozen at -80°C. The culture for the control was completed the same way except that after 0h, the remaining culture was washed with PBS and resuspended in glucose depleted EMM (Fig 2.1b and c). To equilibrate the number of cells for the experiments, a hemocytometer was used to count the cells. Finally, 1 ml of culture at each time-point and condition was stored to image the cells and measure their dimension.

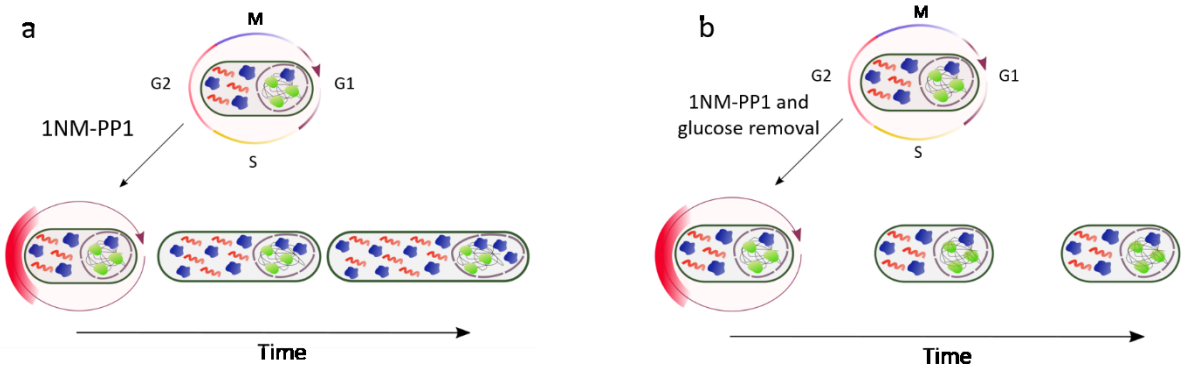


Figure 2.1: cdc2-M17 arrests in G2 phase upon addition of 1NM-PP1. A*. In a normal EMM culture, cells arrest after 1NM-PP1 addition but keep growing in size.* ***B*** *When glucose is removed from the medium and 1NM-PP1 is added, the cells are not growing in size. This population will to control the possible biases due to 1NM-PP1.*

- SBM320

To create the SBM320 used for the quantification of H2A-GFP, two colonies of *h90 hta2-GFP-HA::KanMX ade6-216 leu1-32 lys1-131 ura4-D18*, named SBM69 in our inventory (Sun et al. 2020a) and *h-cdc2-asM17::bsdR* (the *cdc2-asM17*) were mixed in 10 µl of sterile water. The 10 µl were then deposited on the center of a MEA plate and dried 15 minutes under the sterile hood. The cells were incubated three days at 25°C until spore-containing asci are visible under the microscope: they reflect the mating of the two strains. The cells are then scrapped and transferred into one microcentrifuge tube containing 1 ml of sterile water reaching approximately OD600 = 0.1. The tube is placed in a shaker/heater at 55º and 900rpm, for 30 minutes. This step kills most cells that are not spores. The remaining living cells are spread into a YES plate containing G418 and blasticidine antibiotics. The cells which are growing on this plate should be *cdc2-asM17::bsdR hta2-GFP::KanMX*. Each parent strain (*cdc2-asM17* and SBM69) is plated separately on this plate containing the two antibiotics. Because they are both missing a resistant cassette to one of the antibiotics, no colonies were able to grow on this plate, showing the efficiency of the two antibiotics G418 and blasticidin. To confirm that the

colonies from the cross are indeed GFP positive, the plate are placed under an epi-fluorescent stereomicroscope (Nikon smz1000) illuminating the plate with the laser for GFP excitation.
Then the new *hta2-GFP-HA::KanMX ade6-216 leu1-32 lys1-131 ura4-D18 cdc2-asM17::bsdR* strain, called SBM319 hereafter, can be either h-, h+ or h90, since the SBM233 is h90. Luckily, we have two strains h- and h+ both having the protein uch2 fused to the mcherry. Those strains are *h- ade6-21x leu1-32 uch2-mcherry* and *h+ ade6- M216 uch2-mCherry*, named SBM183 and SBM188, respectively. That is why two crosses are done as previously: SBM183 with SBM319 and SBM188 with SBM319. After the crosses, the condition giving asci is kept, namely SBM319 crossed with SBM188. Since these two strains do not have any antibiotic resistant cassette but are auxotrophic for adenine, the cells are selected with the same plates as previously, namely YES plates containing G481 and blasticidin. A negative control consists in growing cells without adenine in the medium to check that they do not grow (data not shown). Finally, the positive colonies are confirmed with the stereomicroscope (Nikon smz1000) the laser for mcherry excitation. SBM320 is then grown as explained above, first in YES for the C1 and the C2 and finally in EMMS (EMM with supplements) for the C3. The harvest of the cells, as well as the arrest with 1NM-PP1, are done the same way as for *cdc2-asM17.*

2.1.2. Staining the cells with calcofluor white

Before imaging the elongating *cdc2-asM17*, one crucial step is to stain them with calcofluor white. This dye is known to stain fission yeast cell wall and septum and is used in my work for two purposes: visualize the wall to precisely measure the length of the cells and count the cells having a septum in the population of *cdc2-asM17* arrested with 1NM-PP1.
Firstly, 50 µl of the culture at each condition is transferred into a microcentrifuge tube. Then, 2 µl µl of the calcofluor white solution is added and mixed to the culture 50 µl. µl Finally, a small volume is deposited on a glass slide and covered by a coverslip for imaging the stained cells. Because calcofluor white is excited between 340 and 360 nm wavelength and its emission peak is around 400–440 nm, the same filter as for detection of DAPI staining can be used.

2.1.3. Flow cytometry to control the DNA content of the cdc2-asM17 strain at different sizes

The *cdc2-asM17* strain was engineered by Aoi and colleagues (Aoi et al. 2014). To confirm that the strain behaves similarly as in the study, the DNA content of the *cdc2-asM17* grown without and 6h with 1NM-PP1 was measured by flow cytometer. Different populations of *cdc2-asM17* were tested: cultivated without 1NM-PP1, cultivated 6h with 1NM-PP1 and cultivated without 1NM-PP1 and without NH4Cl.

The protocol for preparing the cells before analyzing their DNA content by flow cytometry can be found in the *S. pombe* guide "Analysis of the Fission Yeast *Schizosaccharomyces pombe* Cell Cycle" by Eliana B. Gómez and Susan L. Forsburg (Gómez and Forsburg, 2004). The cultures mentioned above were transferred in a tube that was centrifuged 2000g to pellet the cells. To fix the cells, 1 ml of cold 70% ethanol was added to the pellet. After that, 0.3 ml from each tube was transferred in a 5 ml FACS (Fluorescence Activated Cell Sorting) tube containing 3ml 50 mM sodium citrate. The samples were centrifuged once more at 2000g for 5 min. After removing the supernatant, the pellet was resuspended in 0.5 ml of 50 mM sodium containing 0.1 mg/ml RNase A (Merck, 9001-99-4), the sample were incubated at 37°C for 1–2 h. Then, the DNA of the three populations was stained with propidium iodide (PI): 0.5 mL 50 mM sodium citrate containing 8 µg/mL PI was added to each sample so that the final concentration is 4 µg /ml of PI. Finally, stained cells were processed using a BD LSR II flow cytometer and the results were analyzed using the Flowjo software.

Because the G2 phase is longer than the other phases of the cell cycle in *S. pombe*, a normal population (cultivated without 1NM-PP1) shows a 2C DNA content, namely a replicated DNA molecule (**Fig 2.2a**). A population arrested in G1 because of deprivation in NH4Cl indeed shows most cells having half less signal than cells in G2 (**Fig 2.2c**). The result from cells cultivated 6h with 1NM-PP1 shows exactly what is expected for the *cdc2-asM17* were higher signal on the x-axis corresponds to mitochondrial DNA that replicates in large cells (**Fig 2.2b**) (Seel et al., 2021; Aoi et al., 2014; Dischinger et al., 2008). Therefore, the results acquired on our *cdc2-asM17* confirm to previous results in terms of its DNA content after size increase (Aoi et al., 2014; Dischinger et al., 2008).

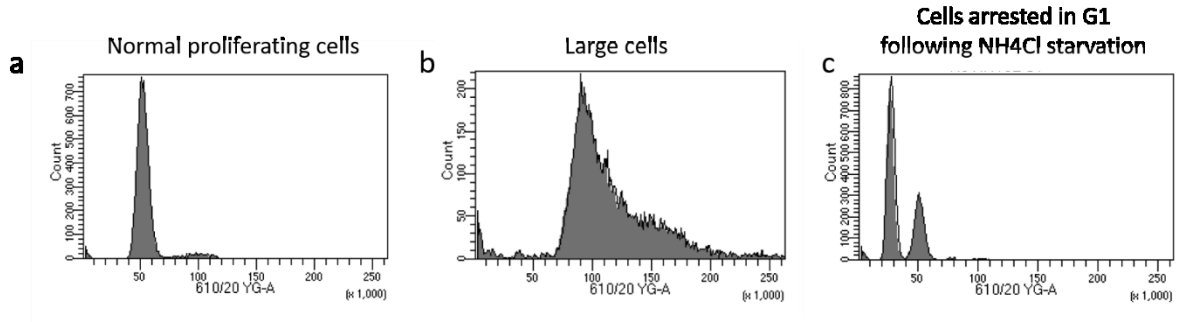


Figure 2.2: DNA content for different cell population. A. *DNA content for normal proliferating cells without 1NM-PP1.* ***B.*** *DNA content for cells arrested 6h with 1NM-PP1.* ***C.*** *DNA content for cells arrested with deprivation of NH4Cl.*

2.1.4. Counting the cells showing a septum

To complete the characterization of the *cdc2-asM17* strain, I needed to verify that it indeed arrested after addition of the ATP analogue 1NM-PP1. One simple way to achieve that aim is to count and thus estimate the percentage of cells having a septum in the population. Indeed, cells about to divide position a septum at the center of the cell. For that, I imaged cells at each time-point after addition of 1NM-PP1 that were previously stained with calcofluor. The percentage of the septating cells on more than 100 cells can be seen on figure 2d (Fig 2.3). After 1h of blocking with 1NM-PP1, no cells with a septum can be counted. Notably, a smaller percentage of septating cells than at 0h are visible at 4h and 6h after addition of 1NM-PP1. However, those septa were detected on cells of large size as illustrated on figure 2.1c (Fig 2.1c). This behavior of large cells assembling a septum was already witnessed previously during Amalia Segura Martínez's PhD (data not shown), but no division were associated to this observation since the same number of cells can be counted after 6h (Martínez Segura 2017).

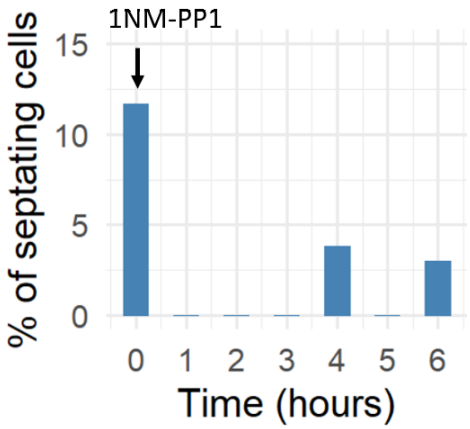


Figure 2.3: Percentage of cells having a septum *(~100 cells were counted for each time-point).*

2.1.5. Protein extraction

Cells were resuspended in commercial lysis buffer (Cell Signalling, 9803S) containing 50X of protein inhibitor (Roche EDTA-free tablets 04693132001). The volume of the lysis buffer is adapted depending on the size of the pellet but more importantly, the volume of lysis buffer was the same for each condition (0h to 6h samples after addition of 1NM-PP1). Glass beads were added to the tubes and the cells wall was destroyed with 3 bead-beating cycles at 6.0 m/s in the Fast-Prep machine (from MP biomedicals), letting the samples cool down on ice between cycles for 5 min. The lysate was extracted by centrifugation at 3000 rpm, 3 minutes at 4^{o}C. The samples were incubated on ice for 20 minutes and centrifuged once more at 3000g, 5 minutes at 4^{o}C.

2.1.6. Protein measurement

The measurement of the protein concentration was done by Bradford assay. To this aim, the Pierce® Microplate BCA Protein Assay Kit (23252) was used. Briefly, the protocol consists in measuring the absorbance with a plate reader (SPECTROstar Omega) of standard samples containing a known concentration of BSA, concomitantly with the absorbance of the samples of interest. The concentration of the samples of interest can thus be deduced thanks to the standard curve of the BSA samples.

2.1.7. Western Blot

After measurement of the protein lysate by a Bradford assay, the volume corresponding to 20 µg of proteins at 0h was loaded in each well of a 13% acrylamide gel and subject to SDS-PAGE. After transferring the proteins from the gel to a nitrocellulose membrane, the loading of the proteins was revealed by staining with red Ponceau (0.1% (x/v) Ponceau S in 1% (v/v) acetic acid). After washing out the stain, the membrane was blocked overnight at 4^{o}C using 10% skim milk (BD Biosciences 232100). The membrane was then incubated 50 minutes at room temperature with Histone 3 antibodies (ab1791, polyclonal and ChIP Grade) to a final concentration of 0.5 µg/ml and with TBS-T (10% TBS #1706435 and 1% Tween-20 #1610781). After stripping the membrane for 10 minutes using a stripping buffer (Thermo Scientific™ 21059), the membrane was incubated 45 minutes with H3K9ac (abcam 4441, polyclonal and ChIP Grade) to a final concentration of 1 µg/ml. The secondary antibody incubation was performed with goat anti-rabbit polyclonal HRP (P0448), diluted 3000 times. The membranes were imaged by chemiluminescence with ImageQuantLAS4000. The quantification was done using imageJ.

2.1.8. Microscopy

A small volume (around 1 ml) from a 50 ml C3 culture (OD600 = 0.8) is transferred into a microcentrifuge tube.

To measure the cell size, the cells were pelleted 3 min at 3000 rpm and then resuspended in PBS. 10 µl is then is deposited on a glass slide and covered by a coverslip. As mentioned above, calcofluor white is excited between 340 and 360 nm wavelength and its emission peak is around 400–440 nm. Therefore, the same filter as for detection of DAPI staining can be used.

All images were acquired with the Leica DM4B microscope with a X40 objective, retransmitted thanks to the Leica DFC365 FX camera and with the Leica Application Suite X (LAS X). GFP channel was captured by 'L5' filter (LED:470, Excitation: 480/40, Emission: 527/30), mcherry channel by the 'N21' filter (LED: 530, Excitation: 538/46, Emission: LP 590) and calcofluor was captured by 'DAP' (LED:365), Exc: 350/50, Em: 460/50). The illumination of the channels was provided by the EL6000 with 120W metal halide lamp (also Leica).

2.1.9. MNase-seq

The following protocol is an adaptation of the MNase digestion from Mason and Mellor (Mason and Mellor, 1997). In this study, the MNase-seq experiment was first tested on normal proliferating cells harvested at OD600 ≈ 0.8 and cells from the same initial population arrested 6h with 1NM-PP1. Then, the MNase-seq was conducted three times on normal proliferating cells (OD600 ≈ 0.8) and on cells from the same initial population collected each hour after addition of 1NM-PP1, and on cells arrested 6h with 1NM-PP1 and glucose removal. Each condition contains around $1x10^9$ cells. All the centrifugations for this protocol were done at 4^oC and the samples were kept on ice until extraction of the DNA.

When needed for the MNase digestion, the pellets of cells were transferred on ice, and around $1 x10^9$ cells were resuspended in 5 ml of 1.2 M sorbitol, 14 mM β-mercaptoethanol, 2.5 mM EDTA. Zymolase (125 mg) was added to the cells and incubation for 7 min at 30^oC and shaking at 200 rpm was done in order to digest the cell wall. The spheroplasts were observed under the microscope: 80 % of cells should appear black. Cells were pelleted for 3 minutes at 3000 rpm and washed twice in cold 1.2 M sorbitol. To enrich the nuclei, 250 µl of 18% Ficoll was added to the pellet and mix a yellow tip or a loop. When the solution is homogenous, another 250 µl of 18% Ficoll was added. The nuclei were isolated by spinning full speed, 30 seconds, washed in cold 1 M sorbitol and re-pelleted at full speed, 8 seconds. After removal of the supernatant, the enriched nuclei were re-suspended in 2.5 ml cold MNDB buffer (1.2 M sorbitol, 10 mM CaCl 2, 100 mM NaCl, 1 mM EDTA, 14 mM 2-mercaptoethanol, 50 mM Tris-HCl pH 8.0). Five aliquots of 0.5 ml of the MNDB preparation were made and different MNase amount were added to each of them: 0, 4, 8, 16 and 32 U/ml of MNase (Sigma-Aldrich, N5386). The MNase stock solution is 850 µl of 10 mM HEPES-KOH, pH 7.6, 1.5 mM MgCl2, 0.5 mM EGTA, 10% (v/v) glycerol, 50 mM KCl, 1 mM DTT, 0.2 mM PMSF added to 1 vial of 500 U Micrococcal Nuclease. This stock solution is considered to be 50 units per µl. The tubes were incubated 20 minutes at 37 ^oC. The reaction was stopped by adding 100 µl of 250 mM EDTA and 50 µl of 10 % SDS. Then, DNA was extracted by one phenol-chloroform and supernatant was incubated 30 minutes at 37 ^oC with 100 µL of 10 mg/ml RNase A (Merck, 9001-99-4). Again, DNA was extracted with phenol:chloroform isoamylalcohol and precipitated with isopropanol (0.7 times the volume of the supernatant) and NaCl (1/25 times the volume of the

supernatant). The precipitate was collected by spinning full speed and washed with ice cold 80% ethanol.

The DNA of each sample was run in a 1.5 % agarose gel (Fig 2.4). The mononucleosome band was taken out and purified with the QIAquick Gel Extraction Kit (Qiagen, 28704). After measurement of the purified DNA with the qubit with the high sensitivity (HS) assay (Q33230), the DNA was run into a bioanalyzer (Agilent) to check out the size of the fragments.

One possible limitation of cutting out the mononucleosomal band in these gels is to inconsistently select the mononucleosomal fragments for each condition. To limit this bias, a large portion of the gel containing the mononucleosomal band was selected for DNA extraction (data not shown).

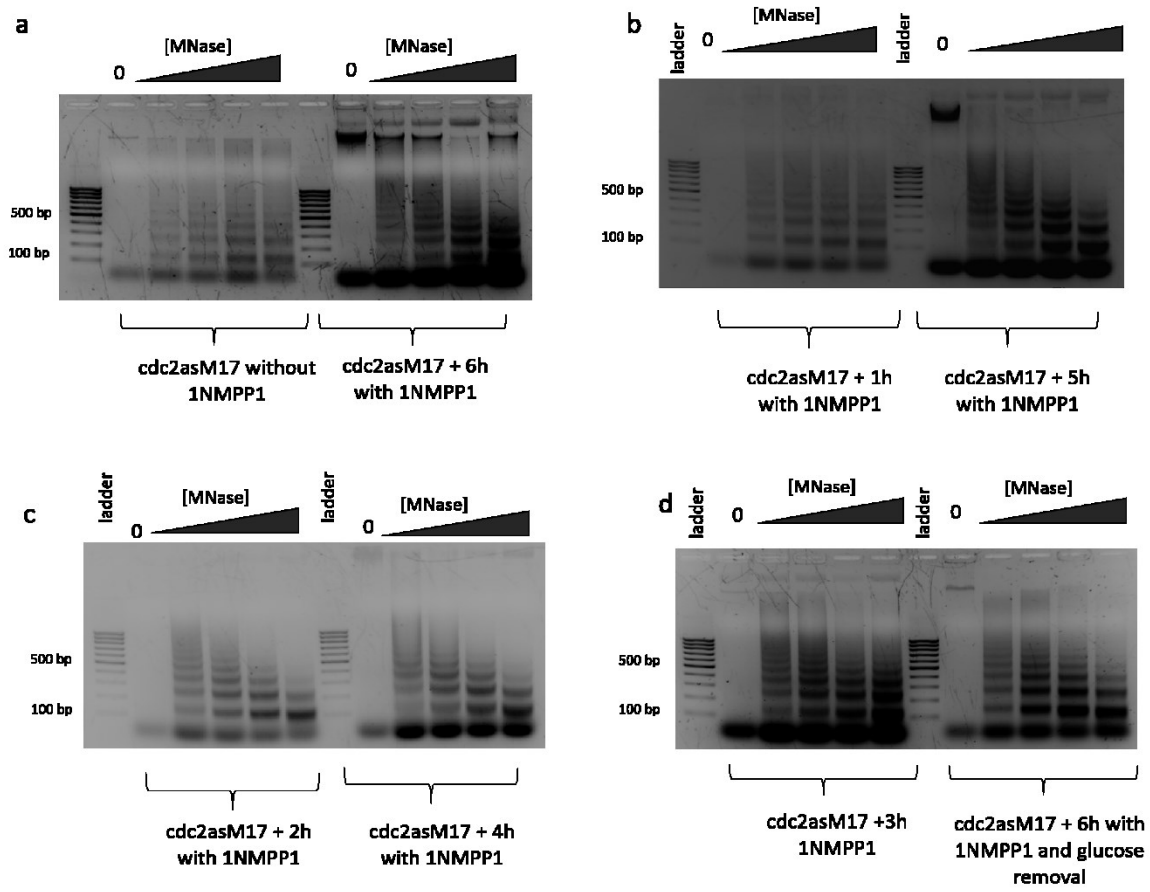


Figure 2.4: DNA fragments after MNase digestion *of increasing concentration (4, 8, 16, 32 U/ml) at each time-point after addition of 1NM-PP1 and for cells arrested 6h with 1NM-PP1 and glucose starvation.*

2.1.10. ChIP-seq with a H3 antibody

The ChIP-seq procedure for this work in based on published protocols (Rodríguez-López et al., 2020). Two conditions are performed for this experiment: normal proliferating cells not cultivated with 1NM-PP1, and large cells arrested 6h with 1NM-PP1. Each condition is harvested at around $2x10^9$ cells in 200 ml of culture. The cells are cross-linked with 5.4 ml of 37% formaldehyde (1% final), then quenched with 10 ml of 2.5M glycine. After washing the cells with 40 ml ice-cold PBS, the pelleted and frozen at -80°C. The following day, the cell pellet is thaw on ice for 5 minutes and then resuspended in 800 µl of ice-cold Lysis Buffer (50mM Hepes pH 7.6, 1mM EDTA pH8, 150mM NaCl, 1% Triton X-1000, 1% Na-Doc) containing 2x protease inhibitors (Roche EDTA-free tablets 04693132001) and 1 mM PMSF. Around 500 µl of acid-washed glass beads were added to the tubes. The cells were then broken in the Fast Prep machine (from MP biomedicals) with 12 cycles of 20 seconds at 5.5. Each cycle was separated by 5 min of incubation on ice. After visualizing the breaking efficiency (>90%) under the microscope, the lysate from the broken cells is extracted by centrifugation at 4 ºC, 1 min at 1000 rpm. The glass beads are washed again with a lysis buffer containing 2x protease inhibitors (Roche EDTA-free tablets 04693132001) and 1 mM PMSF to increase the yield. After that, the beads-free lysate is spined down for 10 min at full speed (14000 rpm) and 4 ºC. The supernatant is discarded, and the pellet is washed with 800 µl Lysis buffer (with protease inhibitors and PMSF) to increase yield. Then, the pellet is resuspended in 750 µl of cold Lysis Buffer (50mM Hepes pH 7.6, 1mM EDTA pH8, 150mM NaCl, 1% Triton X-1000, 1% Na-Doc). The two tubes corresponding to one condition are pooled together so we are back to 1 microcentrifuge tube per condition). PMSF is added to 1 mM final. Split the volume into several 300 µl aliquots in Axygen 1.5 ml to proceed to the sonication of the cells. The sonication of the cells is done using the Diagenode Bioruptor® 300, 30 sec ON/ 30 sec OFF at HIGH setting for two 2 rounds of 20 min and 25 min. The two rounds are separated by 15 min to allow the bioruptor to rest. The sonicated material is spined 10 min at full speed (14000 rpm) at 4°C. This time, the supernatant, which is the chromatin extract, is collected. From this chromatin extract, 5 µl is used to measure the protein concentration (see the protocol mentioned above), and 50 µl is used as input and to check the sonication efficiency. To extract the DNA from the 50 µl chromatin extract that will be used as input, the chromatin extract is incubated with TES

(10mM Tris-HCL pH8, 1mM EDTA, 1% SDS) overnight at 65°C. Then, 200 µl of TE (10mM Tris-HCL pH8, 1mM EDTA) and 5 µl of RNase A (0.5 mg/ml) (Merck, 9001-99-4) are added and the resulting samples are incubated 30 min at 37°C. The proteins are degraded using 7 µl of proteinase K (20 mg/ml) (Kisker 348569) and incubation 2h at 55°C. Finally, the DNA is extracted using the MinElute Reaction Cleanup Kit (Qiagen, 28204). The DNA showing the sonication efficiency is visualized on a 1% agarose gel (Fig 2.5 a, b and c). The DNA fragments are between 100 and 500 bp, which satisfies the conditions of a ChIP-seq experiment. Therefore, the immuno-precipitation step can take place.

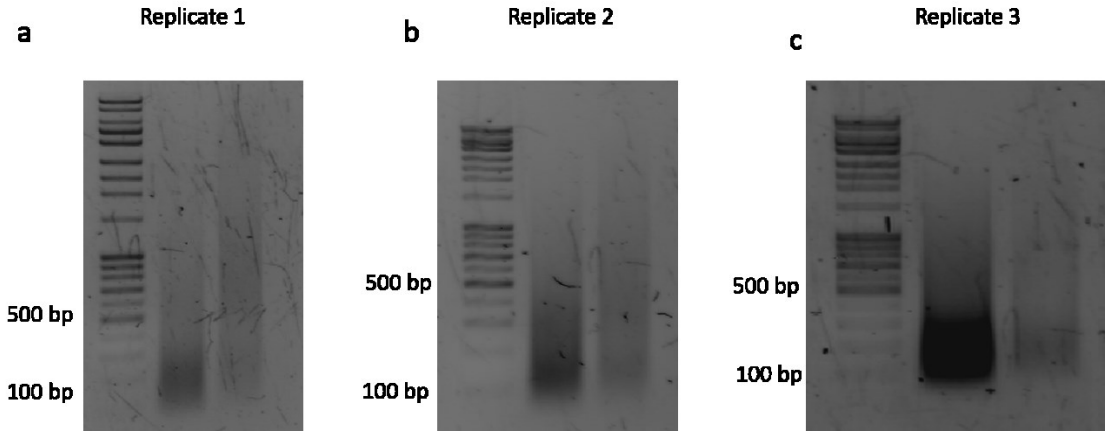


Figure 2.5: Sonicated DNA before H3 antibody incubation *during the ChIP-seq protocol and for replicate 1 (**a**), 2 (**b**) and 3 (**c**).*

Before immune-precipitating the DNA bound to H3 proteins, the H3-antibody (ab1791, polyclonal and ChIP Grade) must be first bound to magnetic beads. To this end, 100 µl of magnetic beads (Invitrogen Dynabeads, 50 µl of protein G and 50 µl of protein A, 10003D and 10001D respectively) are added to a microcentrifuge tube and washed with 1 ml of block solution (0.5% BSA (w/v) in Lysis Buffer). The beads are then collected using a magnetic stand and washed twice more with the block solution. The beads are resuspended in final volume of 250 µl of block solution containing 5 µg of H3 antibody (ab1791). The beads containing the H3 antibody are incubated all day on a rotating wheel at 4°C. The beads are then washed 3 times in 1 ml block solution and resuspended in 50 µl of block solution. Based on the protein measurement performed previously, 5 mg of protein from the chromatin extract are added to the beads-antibody solution. In my condition, this corresponds to approximately 600 µl of chromatin extract. The samples are incubated overnight on a rotating wheel at 4°C.

Then, the beads are collected using the magnetic stand and several crucial washes are performed: two washes with 0.8 ml Lysis Buffer, 2 washes with 0.8 ml Lysis 500 buffer (50mM Hepes pH 7.6, 1mM EDTA pH8, 500mM NaCl, 1% Triton X-100, 0.1% Na-Doc), two washes with 0.8 ml LiCl/NP-40 buffer (10mM Tris-HCl pH8, 1mM EDTA pH8, 250mM LiCl, 1% NP-40, 1% Na-Doc) and one wash in TE buffer (10mM Tris-HCL pH8, 1mM EDTA).

The beads are then resuspended in 200 µl TE and incubated 65°C for 18h with few vortexing at the beginning. This step is to elute and reverse crosslink. The RNA and proteins are digested by adding 200 µl TE having 5 µl of RNase at 0.5 mg/ml (Merck, 9001-99-4) and incubating the samples 1h at 37°C and then by adding 7 µl of proteinase K (20 mg/ml) and incubating the samples 2h at 55°C.

Finally, the DNA is extracted as previously using the MinElute Reaction Cleanup Kit (Qiagen, 28204). After checking the DNA concentration by qubit, the samples are ready for library preparation!

2.1.11. Library preparation

The fragments were then transformed to make sequencing libraries with an Illumina kit (NEBNext® Ultra™ II, NEB #E7645S). The starting material was 10 ng of DNA, the adaptor dilution was 1:10 and the amplification of the adaptor-linked fragments was performed with 6 PCR (Polymerase Chain Reaction) cycles. After the library preparation, the size was checked again on a bioanalyzer (Fig 2.6 for MNase-seq and Fig 2.7 for ChIP-seq). Finally, the libraries were measured with the qubit HS assay (Q33230) and all the samples were pooled together at the same concentration to proceed to the sequencing.

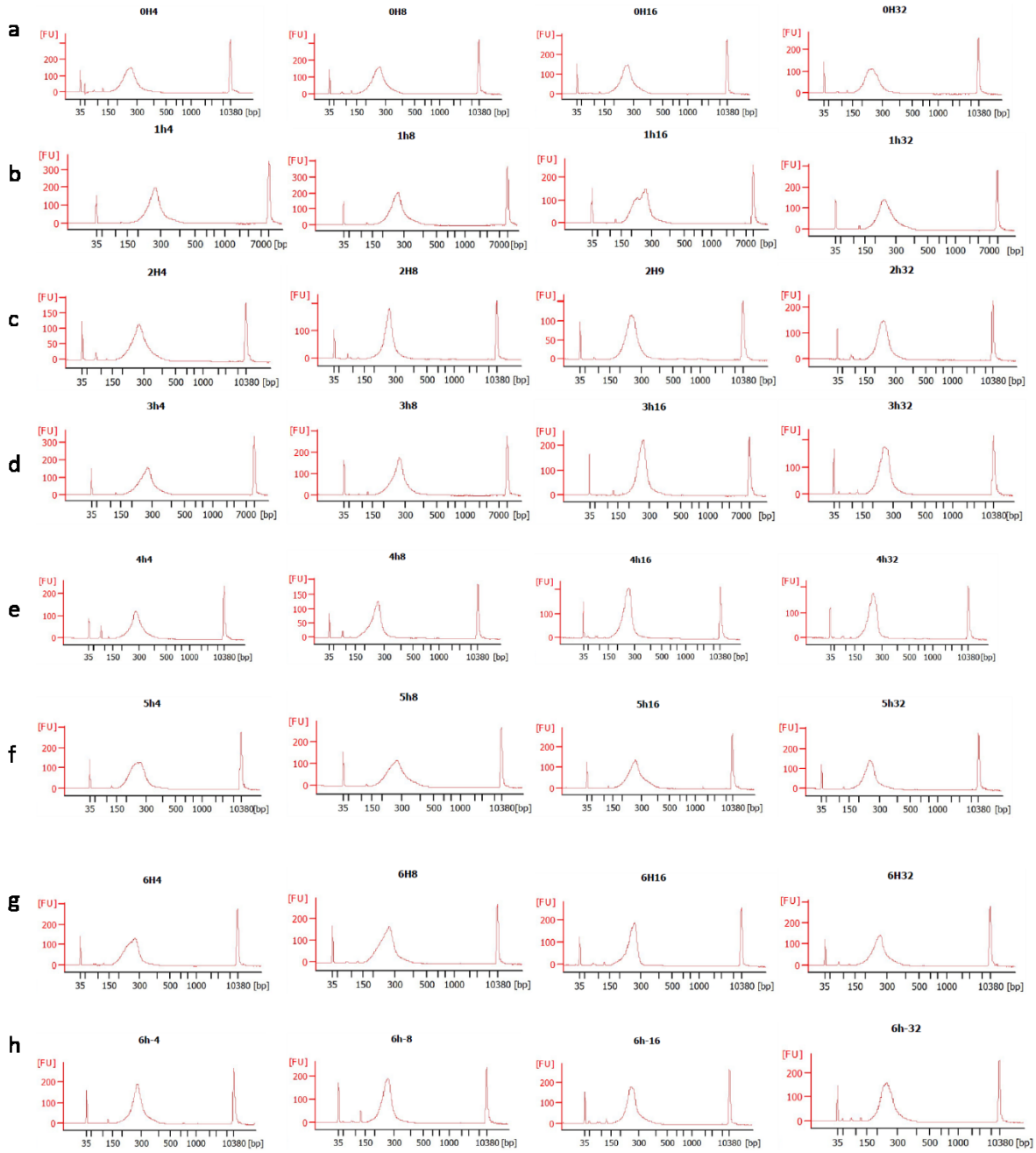


Figure 2.6: DNA fragments after library preparation of the MNase-seq samples *for time-point 0h (**a**), 1h (**b**), 2h (**c**), 3h (**d**), 4h (**e**), 5h (**f**), 6h (**g**), 6h without glucose (**h**) for increasing MNase concentration (from left to right).*

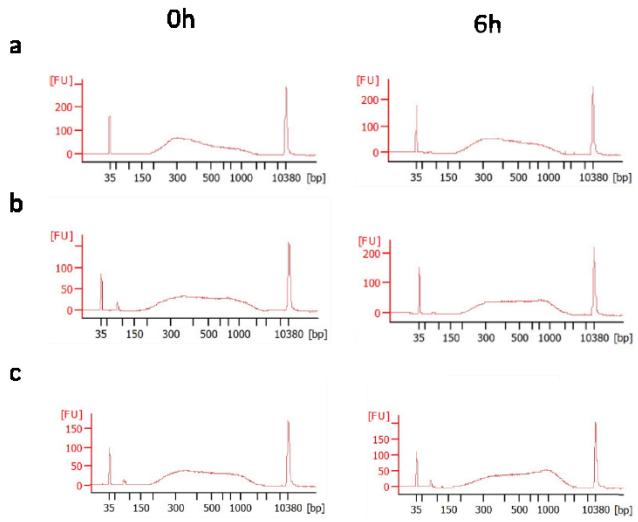


Figure 2.7: Library preparation of the ChIP-seq samples *for replicate 1 (**a**), 2 (**b**) and 3 (**c**).*

2.2. Computational analysis

2.2.1. Computational analysis of the MNase-seq data

The paired-end sequencing of the MNase-seq samples using the HiSeq instrument, the paired reads were aligned to the fission yeast genome as available in PomBase in July 2019 (Lock et al., 2019) using Bowtie with the following command lines:

```
bowtie --verbose --suppress 1,6,7,8 -t --fr -p 1 -m 1 -v 2
$HOME/SEQ_PIPELINE/GENOMES/ANNOT_2019/SP -1 ${toDo}${suf}_R1.fastq -
2 ${toDo}${suf}_R2.fastq ${toDo}${suf}.bowtie 2>
${toDo}${suf}.bowtieout.
```

Using the software samtools, I was able to list, for each sample and for the two MNase repeats used in this study, the total number of reads, the mean of coverage and the number of covered bases in mega base-pair (Mbp). Here is the line for such operation:

```
samtools coverage bowtie_output.sorted.bam -m > output.txt
```

Table 2.1 and 2.2 recapitulates these number for the two replicate 1 and 2 respectively (**Table 2.1 and 2.2**).

Table 2.1: Three properties of read coverage on genomic DNA for the first replicate.

Time-points	MNase concentration	Number of reads	Mean coverage	Covered bases (Mbp)
0h	4U	6,407,946	47.5	11.14
	8U	5557726	39	8.57
	16U	3993306	28.475	9.5823
	32U	6376418	46.425	11.5
1h	4U	7186354	54.125	11.5
	8U	7073318	51.55	11.3
	16U	7633764	56.45	12
	32U	4688660	33.55	11.5
2h	4U	5308470	41	12
	8U	12693864	100.17	12.16
	16U	5346604	39.675	11.8
	32U	6314188	45.375	12.26
3h	4U	3543680	27.3	11.78
	8U	10349790	82.5	12.16
	16U	13551000	107.5	12.16
	32U	13248016	102.65	12.16
4h	4U	7444274	58.5	12.14
	8U	4688660	33.55	11.5
	16U	5308470	41.05	12
	32U	13760829	105.5	12
5h	4U	6854550	53.67	12.14
	8U	3595454	26.15	11
	16U	4366778	32.6	12
	32U	8754642	64	12
6h	4U	6321094	48.15	12
	8U	5854799	44.57	12
	16U	13329420	105	12.15
	32U	5452562	39.625	11.617
6h without glucose	4U	8009156	36.2	12
	8U	10577944	81.55	12.16
	16U	8782936	65.1	12
	32U	9349020	68.475	12

Table 2.2: Three properties of read coverage on genomic DNA for the second replicate.

Time-points	MNase concentration	Number of reads	Mean coverage	Covered bases
0h	4U	2,161,773	15.1	10.3639
	8U	14941652	114	12.14
	16U	6013302	43.55	12.2
	32U	7989858	57.475	15.1
1h	4U	8209940	61	11.83
	8U	7077074	52	11.66
	16U	8755282	64	11.5
	32U	7438958	53.75	11.5
2h	4U	9215460	71.6	12
	8U	6637518	50	11.8
	16U	5056602	37	11.3
	32U	9577802	72.6	12
3h	4U	12589492	97.6	12.17
	8U	7442380	57.35	12.14
	16U	14035574	106	12.14
	32U	6250038	45.4	11.8
4h	4U	5116872	38.33	11.53
	8U	6160089	45.8	11.6
	16U	8857900	65.42	11.8
	32U	8822396	65.8	12
5h	4U	6476179	49.7	12.14
	8U	72920	0.5	3.9
	16U	6584086	49	12
	32U	7571702	59.5	12
6h	4U	8390037	64.2	12
	8U	12254734	94.35	12
	16U	12517312	93.78	12
	32U	16351686	122	12
6h without glucose	4U	5478332	40.3	12
	8U	8054758	59.3	12
	16U	4975626	36	11.6
	32U	4309394	30.8	11.6

To generate accurate genome-wide maps of nucleosomes, the NUCwave tool, created by Luis Quintales and colleagues was used (Quintales et al., 2015) with the following command line:

```
pythonNUCwave_pe.py -g Schizosaccharomyces_pombe_all_chromosomes.fa
-o output_directory -a input_aligned_file.bowtie -w -p
output_file_prefix_name
```

After running this program through ubuntu on each of my sample, several wig files are generated (Quintales, et al., 2015). From all those wig files which each correspond to a time-point after 1NM-PP1 and a specific MNase concentration, only one file will be used. This wig file is the last file generated by NUCwave and it contains a nucleosome map in which the signal

from a nucleosome is deduced from the two paired reads, then smoothened and normalized (Quintales et al., 2015). This file can be visualized in a genome browser such as IGV.

Next, the wig files generated by NUCwave are transformed into bigwig files. These bigwig files are then compressed into a matrix that gathers for all the genes (each row is a gene) the signal in 20 bp bin within 1.5 kb upstream and 1.5 kb downstream from the TSS (each column is a 20 bp bin within this 3 kb centered on the TSS). This was done using the computeMatrix command from deeptools (Ramírez et al., 2016) with the following command line:

```
computeMatrix reference-point --referencePoint TSS  -S
NUCwave_output_file_depth_wl_trimmed.bw -R TSS.bed --outFileName
compressed_matrix.gz --binSize 20 --beforeRegionStartLength 1500 --
afterRegionStartLength 1500
```

This implies inputting the precise coordinate of the TSS of each gene (-R in the above command line). Luckily, such information has been previously published (Thodberg et al., 2019). In this article, the authors map the TSS of 5887 genes in fission yeast thanks to a CAGE-seq experiment (Cap Analysis of Gene Expression).

The generated matrix is then imported on R and will constitute the scaffold for all the figures seen on chapter 3 and 4 that are elaborated with distinct functions and packages (Table 2.3). Firstly, the matrix is normalized as such: for each gene (or row), the signal contained in each bin (or column) is divided by the sum of the signal within that same row. This results in the average signal of all the genes having the same y-intercept between different conditions (Chapter 3, Fig 1a, Fig 2a and Fig 4a, c, e). This additional normalization allows the proper comparison between all conditions. The statistical analysis was made by the function compare_means from ggpubr in R using paired comparison when comparing the different time-points and unpaired comparison when comparing different conditions within a time-point.

Table 2.3: R packages used in this study.

Package name	Version	Reference
ggplot2	3.3.5	Hadley Wickham, 2016
pheatmap	1.0.12	Raivo Kolde, 2019
tidyverse	1.3.1	Wickham et al., 2019
ggpubr	0.4.0	Alboukadel Kassambara, 2017
ggVennDiagramm	1.2.0	Chun-Hui Gao, 2021
reshape2	1.4.4	Hadley Wickham, 2007
ComplexHeatmap	2.8.0	Zuguang Gu et al., 2016
dplyr	1.0.7	Wickham et al., 2021

The normalized matrices were analyzed by Theodoulos Rodosthenous, a PhD student from Vahid Shahrezaei's group at Imperial College London. He performed a t-SNE (t-distributed Stochastic Neighbour Embedding) to compare the chromatin structure of the three replicates at each time-point after addition of 1NM-PP1. This analysis results in the 2-D visualization of the data (Fig 2.8). On the plots of figure 2.8, each dot represents a gene from the matrix. Those dots are organized so that genes having similar chromatin structure are spatially close to one another. Strikingly, the third replicate (pointed with the red arrow) shows a different pattern compared to the other replicates: from 0h, the dots (or genes) form clusters while the dots are uniformly distributed for the two other repeats. This specific pattern of the third replicate can be seen at 2h, 4h, 5h, 6h and control. This proves that, for unknown reasons, the cells from the third replicate were not in the same state as the ones from other replicates. That is why I am focusing my analysis only on the two first replicates.

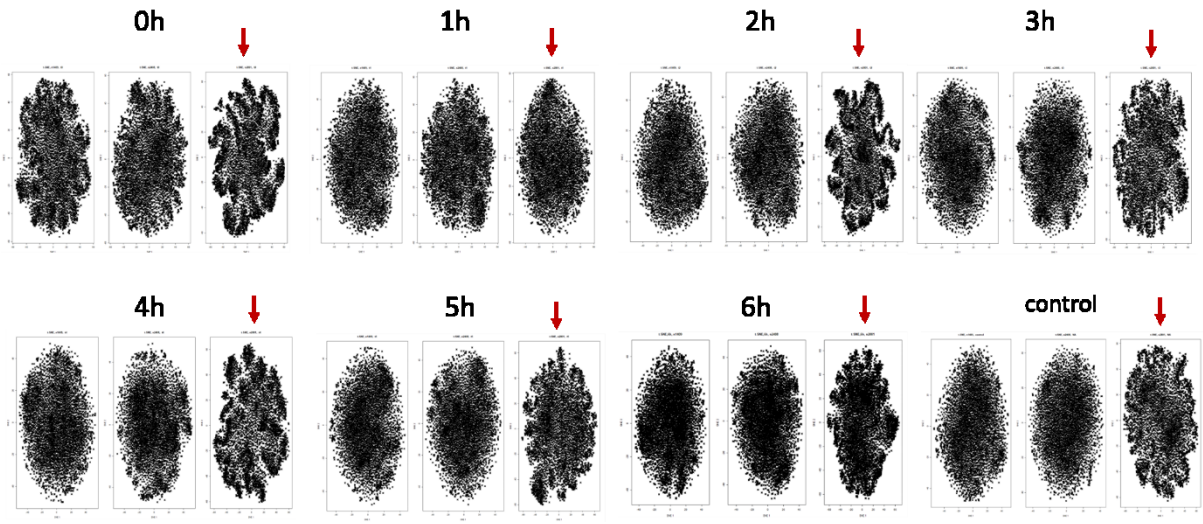


Figure 2.8: tSNE analysis made by Theodoulos Rodosthenous from the mathematic Department (Imperial College London) on three MNase-seq replicates at each time-point after 1NM-PP1and for the four Mnase concentrations simultaneously. *The two first replicates (from left to right) were the ones used for the further analyses (see Chapter 3 and 4) and the last replicate pointed with a red arrow (on the right) was not excluded since it presents different patterns of chromatin organization from 0h and throughout the time-course compared to the two other repeats.*

To support the decision to exclude the third replicates, we observed in this batch an unusual aspect of the cells at 6h after 1NM-PP1 by microscopy (Fig 2.9). Indeed, a large proportion of the population exhibit one or several septa.

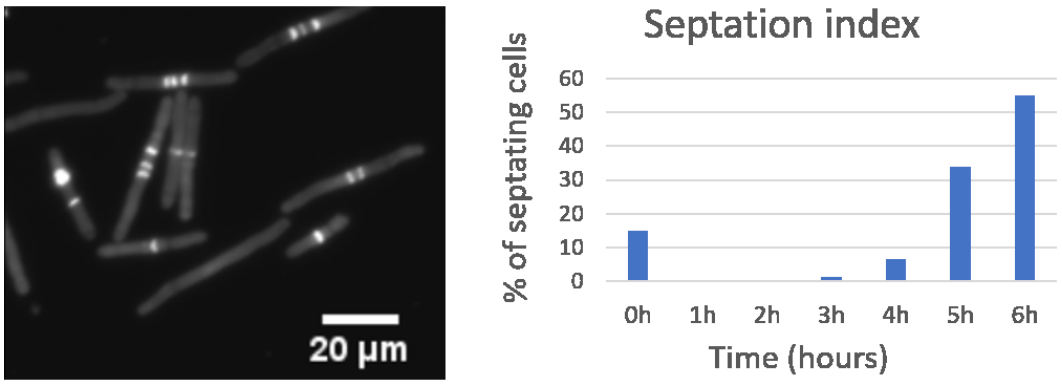


Figure 2.9: Cells at 6h after 1NM-PP1 in the third replicate shows unusual septa*. The septation index is reported for each hour after addition of 1NM-PP1 on the right panel.*

2.2.2. ChIP-seq computational analysis

The alignment of the reads on the fission yeast genome was done as in (Sun et al. 2020a). After sorting and indexing the files, we can list the number of reads, the mean coverage and the covered bases, like previously for the MNase:

```
samtools coverage ChIP_seq_sorted.bam -m > output.txt
```

Table 2.4 illustrates these indications on the read coverage after sequencing (Table 2.4).

Table 2.4: Three properties of read coverage on genomic DNA for the ChIP_seq.

Replicate	Time-point and experience	Number of reads	Mean coverage	Covered bases (Mbp)
1	0h INPUT	15,073,527	59	12.1
	0h IP	18,260,051	70.45	12.1
	6h INPUT	8,119,973	33	12.1
	6h IP	14923602	59.6	12.1
2	0h INPUT	18475672	72.6	12.1
	0h IP	17918859	68.78	12.1
	6h INPUT	11844818	47	12.1
	6h IP	14951183	58.6	12.1
3	0h INPUT	16890757	67	12.1
	0h IP	17246667	66.4	12.1
	6h INPUT	10712060	43.2	12.1
	6h IP	14675718	57.7	12.1

The command bamCompare from deeptools was used (Ramírez et al. 2016) to calculate, at each genomic position, the log2 ratio of signal in large cells over the signal in normal cells. This calculation was performed for both the IP and the input with the following command line:

```
bamCompare -b1 treatment_6h.sorted.mapped.bam -b2
control_0h.sorted.mapped.bam -o H3.bw
```

The default method for normalization is chosen: a normalization by read count (Ramírez et al. 2016). The output is a bigwig file that can be input in the computeMatrix command (Ramírez et al. 2016) in the exact same way as for the MNase-seq:

```
computeMatrix reference-point --referencePoint TSS  -S H3.bw -R
TSS.bed --outFileName H3.gz --binSize 20 --beforeRegionStartLength
1500 --afterRegionStartLength 1500
```

Finally, this matrix is imported in R and the matrices from the three replicates are averaged out. The figures are made with different packages and functions in R (Table 2.1).

Chapter 3: The global changes in chromatin architecture as a function of cell size

This chapter aims to define the global organization of chromatin structure in a fission yeast mutant growing in size. In *S. pombe*, the genome is highly gene-compact (Allshire and Ekwall, 2015) and most genes are expressed, though at various levels (Marguerat and Bähler, 2012). The TSS of the gene is surrounded by important regulatory players in the transcriptional process. Upstream of the TSS lies the promoter that recruits TFs, the PIC and the RNAPII. On the downstream side of the TSS, the +1 nucleosome forms the first barrier encountered by the RNAPII before its progression on the gene body. In the introduction, I reviewed the principle of transcription scaling to cell size, namely the increase in cell size is accompanied by an increase in transcription rate to keep the biomolecule homeostasis. Since chromatin is tightly linked to transcription regulations, how does this translate into chromatin structure, especially around the TSS?

3.1. Chromatin structure of the proliferating and normal-sized *cdc2-asM17*

To test whether size and chromatin structure are related, the chromatin of normal and large cells was compared with an MNase-seq approach (see Introduction and chapter 2, 2.1.9 for further details about the MNase-seq technique). In this protocol, the nucleosomes were not uniquely selected. Instead, all the proteins that were protecting up to ~200 bp of DNA (DNA fragments corresponding to the mononucleosomal fragments) were recovered so the result could offer a complete view of the chromatin structure in those cells. Despite this effort, it is admitted that mostly nucleosomes are detected on the DNA, as they are the main constituent of the chromatin (Mieczkowski et al., 2016). Moreover, the cells were treated with several increasing MNase concentrations. What is the advantage of using multiple MNase concentrations? Traditionally, genome-wide nucleosome maps were established with mononucleosome fragments from a unique MNase digestion. However, the pattern of nucleosomes is dependent on the level of MNase digestion (Chereji et al., 2019; Mieczkowski et al., 2016; Vera et al., 2014; Henikoff et al., 2009). This leads to discrepancies when interpreting the results from a single MNase condition. These dependencies to the MNase

concentration stem from distinctive states of chromatin in which DNA binding proteins are differently exposed to MNase or differently stable on the DNA.

After treating the chromatin from different samples with multiple MNase conditions, the samples were sequenced (MNase-seq) and nucleosome maps are generated with NUCwave from the sequencing reads (Quintales et al., 2015). Then, the signal for each MNase concentration and for each gene is calculated in consecutive bins of 20 bp. Here, I report this signal for a region spanning 1.5 kb on each side of the TSS. The information about the TSS was extracted from the study from Thodberg and colleagues who performed a CAGE-seq experiment to map precisely the TSS for almost all the genes (5887 genes) (Thodberg et al., 2019). After calculating the average signal of all the genes, the signal value in each bin is normalized by the total signal contained in this 3 kb region. The resulting average profile reflects the average signal distribution of all the genes on a TSS-centered region. Hence, the signal in each bin represents a fraction of the total signal from this 3 kb region (Fig 3.1). The promoter is defined here as the 200 bp region right upstream of the TSS, as shown by the red rectangle on each average profile (Fig 3.1).

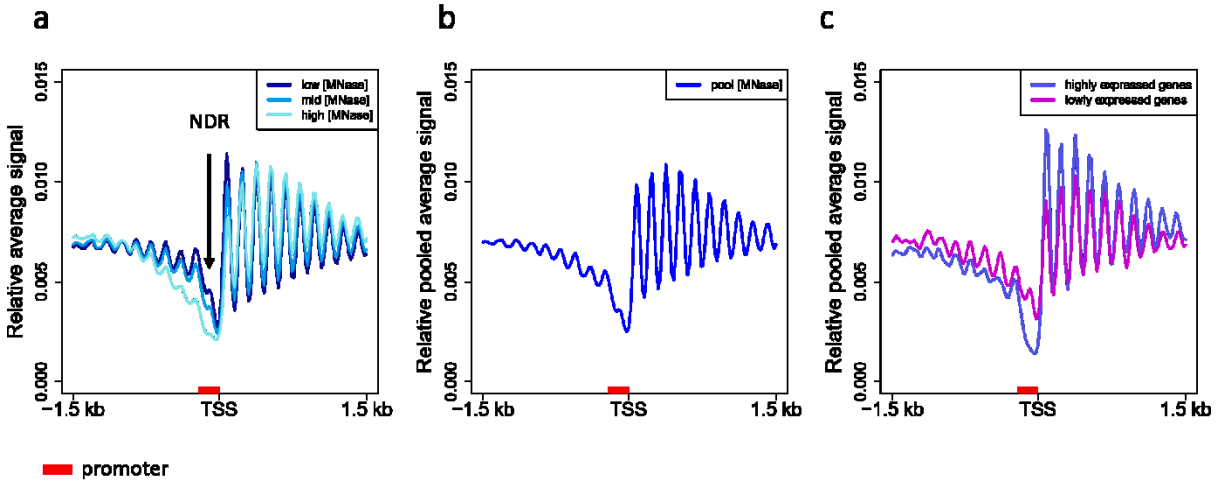


Figure 3.1: TSS-centred average chromatin profile for normal proliferating cells. ***A.*** *Average profile of nucleosomes released at 3 different MNase concentrations.* ***B.*** *Average profile for pooled MNase concentrations.* ***C.*** *Average profile for pooled MNase concentrations and for highly and lowly expressed genes.*

For proliferating cells of normal size, the average profile seen on figure 3 is characteristic of the nucleosome organization around the TSS for fission yeast: i) a fuzzy nucleosome organization upstream of the promoter, ii) a nucleosome-depleted region on the promoter and

iii) a phased array of nucleosomes in the gene body (Fig 3.1a) (Murawska et al., 2020; Yague-Sanz et al., 2017; Lieleg et al., 2015; Soriano et al., 2013; Pointner et al., 2012; Givens et al., 2012; Lantermann et al., 2010). The presence of these three features confirms the good quality of the data. The nucleosome-depleted region (NDR) is reflected by a drop of signal at the TSS that can span around 150 bp depending on the MNase concentration (Fig 3.1a). This clear NDR is widely conserved across eukaryotes, from fungi, to flies, to humans, to plants (Murawska et al., 2020; Chereji et al., 2017; Mieczkowski et al., 2016; Liu et al., 2015; Lantermann et al., 2010; Ozsolak et al., 2007). The weak ordering of the nucleosomes upstream of the TSS is not observed in budding yeast chromatin structure, which, instead, exhibits a regular array of nucleosomes upstream of the promoter followed by a positioned -1 nucleosome (Chereji et al., 2017; Lantermann et al., 2010; Shivaswamy et al., 2008). In some multicellular organisms, such as flies or plants, the chromatin structure upstream of the TSS shares similarities with the one in *S. pombe* with a rather fuzzy nucleosome organization upstream of the TSS (Chereji et al., 2019; Mieczkowski et al., 2016; Li et al., 2014; Mavrich et al., 2008a). Finally, ten nucleosomes can be counted on the gene body up to 1.5 kb downstream of the TSS (Fig 3.1a). This corresponds to what was observed before in *S. pombe* (Soriano et al., 2013; Givens et al., 2012; Pointner et al., 2012; Lantermann et al., 2010) but also in other eukaryote species such as flies (Chereji et al., 2016; Mieczkowski et al., 2016), plants (Zhang et al., 2018, 2015; Li et al., 2014) and mammals (Hu et al., 2017; Jung et al., 2017). However, the comparison of the overlaid nucleosome average profile of *S. pombe* and *S. cerevisiae* shows a progressive de-phasing of their respective nucleosomes, from the TSS to the 3' end of the gene (Yague-Sanz et al., 2017; Bai and Morozov, 2010; Lantermann et al., 2010). This tendency reveals the shorter nucleosome repeat length (calculated between two nucleosome dyad or nucleosome center) in *S. pombe*, possibly due to the absence of the linker histone H1 in this organism (Moyle-Heyrman et al., 2013; Godde and Widom, 1992).

In addition, the nucleosome profile for normal-sized cells shows slight differences depending on the MNase concentration. For example, in the proximity of the TSS, namely on the +1 nucleosome and at NDR, the signal decreases with higher MNase concentrations, showing their respective sensitivity to MNase digestion (Fig 3.1a). Again, this characteristic has been previously noticed in *S. pombe* (Murawska et al., 2020). However, no peak that would reflect the presence of fragile nucleosomes (Jeffers and Lieb, 2017; Kubik et al., 2017, 2015; Vera et

al., 2014; Xi et al., 2011; Weiner et al., 2010; Henikoff et al., 2009; Lee et al., 2007) appears on the promoter after the lowest digestion. This suggests that the MNase concentrations used here, which were carefully chosen to assess a difference between cells of increasing sizes, might be too high to reveal such a peak on the promoter in normal conditions (without 1NM-PP1).

To combine the patterns from different MNase conditions into a single profile, nucleosome occupancies can be calculated as the average signal between different MNase conditions (**Fig 3.1b**) (Mieczkowski et al., 2016); this leads to one profile of pooled nucleosome occupancies. This profile shows the conservation of the nucleosome features described above with a nucleosome-free region upstream of the TSS followed by a regular array of nucleosomes.
Since it has been established that the nucleosome organization was perturbed on genes of different transcriptional state (Mieczkowski et al., 2016; Lantermann et al., 2010), it would be relevant to confirm this statement with the present data in a WT situation. In *S. pombe*, absolute quantities of RNA were previously reported (Marguerat et al., 2012), allowing, in this study, the grouping of genes depending on their expression level. After filtering out the non-coding RNAs, the LTR (long terminal repeat) and the non-detected RNAs in the cell, the genes were sorted in descending order according to their number of copies per cell. Highly expressed genes were defined as the top 1000 genes and lowly expressed genes as the 1000 bottom genes. Because the CAGE experiment mentioned above does not find the TSS of all the genes (5887 vs ~7000 in total), the average profile is drawn here for 826 genes for the highly expressed genes and 826 for lowly expressed genes (Thodberg et al. 2019). As noticed before (Mueller et al., 2017; Mieczkowski et al., 2016; Tolstorukov et al., 2013; Lantermann et al., 2010), highly expressed genes show a high occupancy of the nucleosomes downstream of the TSS as well as a profound drop of signal on the promoter (**Fig 3.1c**). On the contrary, nucleosome occupancy on the gene body of lowly expressed genes is lower, although their position is preserved. In addition, the promoter, instead of being unprotected, displays a small peak that could be explained by the presence of a nucleosome. This is in accordance with the idea that the formation of a nucleosome on the NDR inhibits transcription (Kornberg and Lorch, 2020; Korber and Barbaric, 2014; Rizzo et al., 2011; Hirschhorn et al., 1992; Han and Grunstein, 1988; Lorch et al., 1987) and, consistently, that the depth of the NDR correlates with expression levels (Nabilsi et al., 2014; Kelly et al., 2012; Lantermann et al., 2010; Lee et al., 2004). It is

noteworthy that this peak is yet less represented at this locus compared to the gene body. This could be explained by either a sub-population of cells not exhibiting this peak or by a sub-population of genes not exhibiting the peak. The latter will be assessed in the next chapter in which I will explore the chromatin structure of different sets of genes (chapter 4), although it will not rule out the implication of cell-to-cell variation that could only be addressed by single-cell MNase-seq.

3.2. Chromatin structure of the large *cdc2-asM17* strain

Is the chromatin structure described above perturbed on large cells, arrested for 6h with 1NM-PP1? The answer is yes: cells of larger size exhibit a very peculiar chromatin structure. First, a wide footprint on the promoter is visible after a low MNase treatment and progressively disappears with a more extensive MNase digestion (Fig 3.2a).

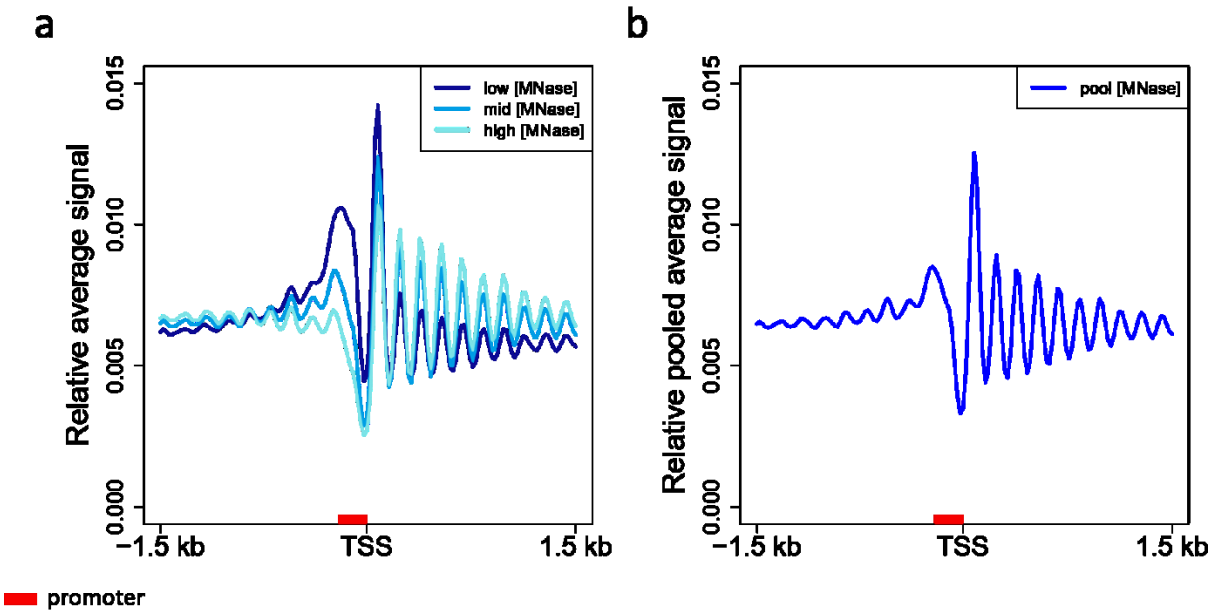


Figure 3.2: TSS-centred average chromatin profile of large 1NM-PP1 arrested cells. A. *Average profile of nucleosomes released at 3 different MNase concentrations,* ***B.*** *Average profile for pooled MNase concentrations.*

In addition, the +1 nucleosome occupancy is also higher than in normal cells; this is the case for the pooled occupancies as well as with the different MNase digestions. These two observations will be quantified later but it is important to first explain why the comparison of nucleosome occupancy within the gene body between large and small cells will not be

addressed. On the one hand, NUCwave generates nucleosome maps that are normalized by the average genome-wide depth coverage, meaning that the signal seen across the genome cannot account for absolute nucleosome occupancy (Quintales et al., 2015). However, comparison between groups of genes is possible to assess their relative nucleosome occupancy, as was done in Figure 3.1c for lowly and highly expressed genes (Fig 3.1c). On the other hand, we are interested here in the global differences of chromatin structure between two cell populations that can be visualized through a nucleosome average profile. As mentioned above, this nucleosome average profile represents a normalized distribution of the signal across a 3 kb region. Therefore, we can apply the same principle as before when comparing only groups of genes, namely focus only on small regions within this 3 kb region and compare their relative signal. The signal observed in those small regions are thus relative to each other and does not reflect the absolute occupancy of a DNA-binding protein at these regions. What are the regions presenting the most interest in this 3kb region spanning either side of the TSS? As reported before, the promoter has been extensively characterized as a main player in transcription regulation as well as a locus with intense chromatin remodeling connected to transcription. Moreover, the +1 nucleosome is the first barrier delimiting the gene, located right upstream from the TSS and has been linked to efficient transcription (Challal et al., 2018; Weber et al., 2014; Rhee and Pugh, 2012; Martinez-Campa et al., 2004). That is why only the promoter and the locus occupied by the +1 nucleosome, as depicted in red and green respectively on the average profiles (Fig 3.3a), will be considered from now on.

Indeed, an orthogonal way of understanding the differences of pooled occupancy between 0h and 6h would be to calculate for each gene the mean signal of these two regions: 200 bp upstream and 200 bp downstream of the TSS (Fig 3.3a).

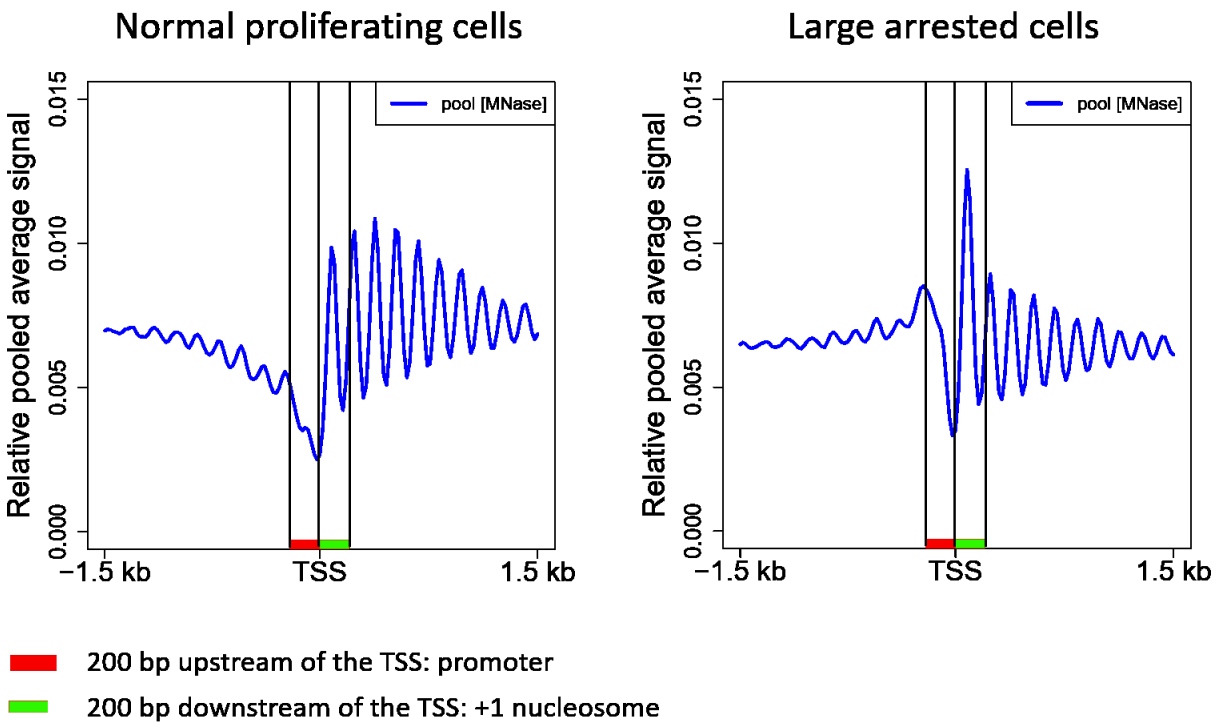


Figure 3.3: Average profile showing the regions considered for the next analysis: the promoter and the +1 nucleosome locus defined herein as the 200 bp upstream and downstream from the TSS respectively.

These new quantitative metrics are meant to reflect the occupancy distribution across the genes of the +1 nucleosome and of the proteins on the promoter. Pooled occupancies on all the promoters clearly show a relative doubling of the median occupancy between normal and big cells with a fraction of 0.003 at 0h and 0.006 at 6h (Fig 3.4a), although the distribution of the signal on the promoter shows some overlap of the interquartile range between the two conditions (ranging from 0.004 to 0.005). Consistent with the average profile, the median of the +1 nucleosome occupancy is already around 0.0055 at 0h and reaches 0.007 at 6h (Fig 3.4b). In large cells, such a significant increase in the average occupancy of the +1 nucleosome might be caused by the increased occupancy at the promoter that would protect the surrounding area from MNase digestion. Furthermore, the overlap of the interquartile range and the tail of the boxplots between 0h and 6h for the +1 nucleosome is larger than the overlap seen on the promoter between the same conditions, 0h and 6h. Therefore, the difference of +1 nucleosome occupancy between large and normal-sized cells less obvious than on the promoter, even though it is significantly higher in large cells

Moreover, these types of measure can be used to compare the signal from low and high MNase concentrations (Vera et al., 2014) because those conditions often tell distinct stories regarding nucleosome organization (Mueller et al., 2017; Mieczkowski et al., 2016; Vera et al., 2014; Weiner et al., 2010). Firstly, the median occupancy on both loci is always higher at a low MNase condition compared to a high MNase condition, with corresponding significant differences in the mean (Fig 3.4c and 3.4d). This illustrates well how those two loci are sensitive to an increasing MNase digestion. More importantly, at a low MNase concentration, the promoter of large cells presents a relative doubling of the median occupancy, similarly to what was observed for pooled occupancy, this time ranging from 0.004 at 0h to 0.008 at 6h (Fig 3.4c). Here, however, there is no overlap of the interquartile range, highlighting the global upshift of occupancy for this condition. At a high MNase concentration, the same relative doubling in occupancy occurs between 0h and 6h but the separation of the distribution is not clear with a part of the genes still having a low occupancy on the promoter at 6h (Fig 3.4c). The same tendency appears on the +1 nucleosome but somehow less remarkable because of the overlap between the MNase conditions (Fig 3.4d). A way to better visualize these results is to calculate the difference of occupancy between a high and a low MNase concentration for both normal and large cells (Fig 3.4e and 3.4f). Indeed, this indicates how DNA protection reacts to different digestion treatments which can then reveal the differential nature of chromatin in those loci. On the promoter, the median difference between low and high MNase concentrations is more than double (Fig 3.4e). This new observation corroborates what was seen for big cells, namely that the promoter of large cells presents a very singular chromatin structure. However, the same conclusion cannot be drawn for the +1 nucleosome for which the median signal between 0h and 6h is less than double (Fig 3.4f), although the means of both distributions, 0h and 6h, are significantly different. Taken together, the average profile combined with the signal mean of restricted loci shows that the promoters in large cells are occupied by proteins that are sensitive to a high level of MNase but resistant to the lowest MNase concentration used in this study. In addition, the +1 nucleosome is also sensitive to the extent of the MNase digestion, and, like the promoter, the DNA underlying the +1 nucleosome is more protected in large cells compared to normal-sized cells, although the range of variation is less clear than for the promoter.

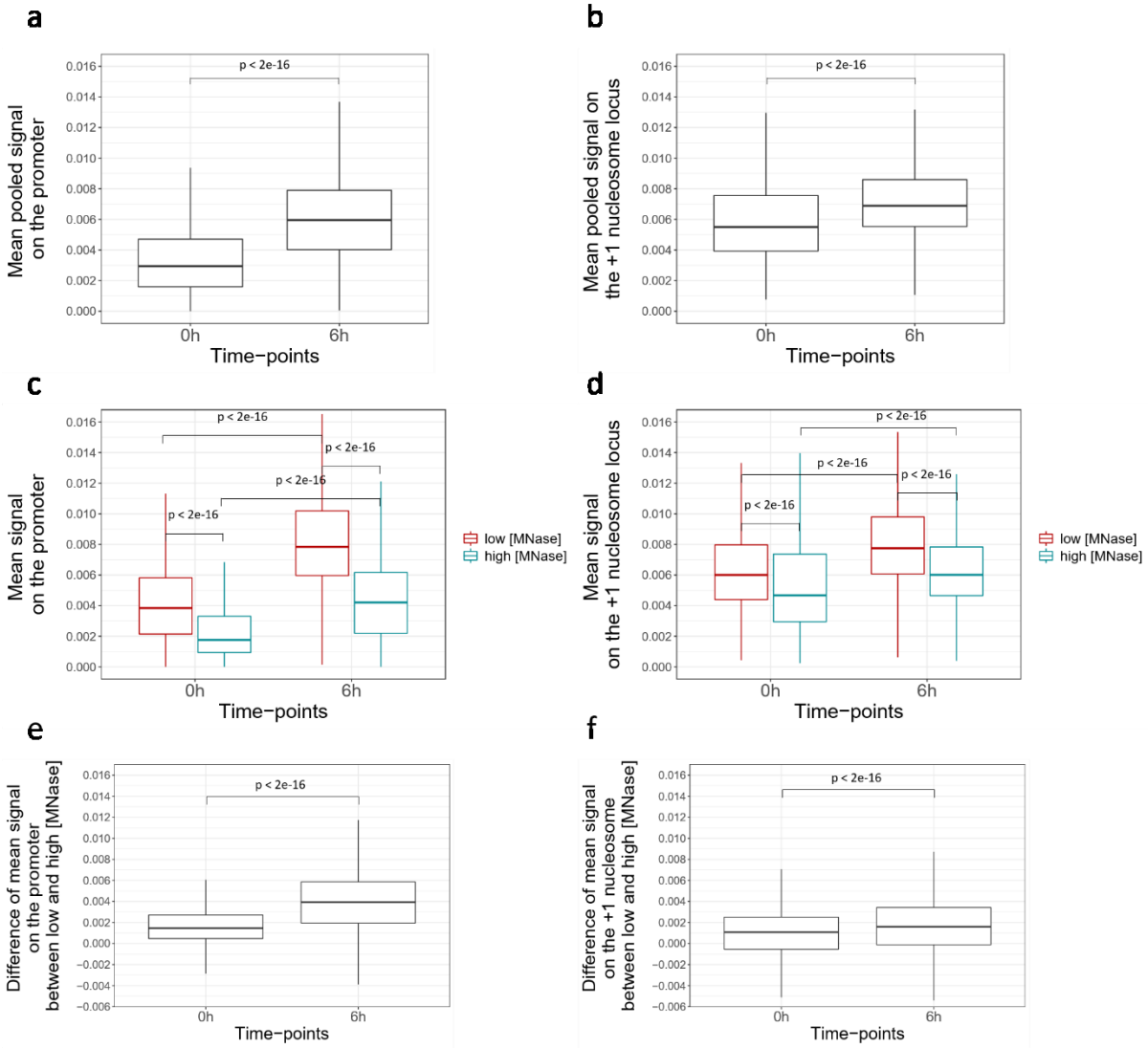


Figure 3.4: Mean signal on the promoter and +1 nucleosome locus for normal and large cells. A-b. *Mean pooled occupancy on the promoter (200 bp upstream of the TSS) (paired t-test, n = 5887 genes)* ***(a)*** *and. on the +1 nucleosome locus (200 bp downstream of the TSS) (paired t-test, n = 5887 genes)* ***(b)****. **C-d.** Mean signal for the low and high MNase concentration on the promoter* ***(c)*** *and. on the +1 nucleosome locus* ***(d)*** *(for both loci, paired t-test is performed to compare 0h and 6h and unpaired t-test to compare low [MNase] and high [MNase], n= 5887 genes). **E-f**. Subtraction of high MNase signal to low MNase signal on the promoter and* ***(e).*** *on the +1 nucleosome locus* ***(f)*** *(paired t-test, n = 5887 genes).*

3.3. Chromatin structure of cells of increasing size

As shown previously, the size of the cell seems to be linked with chromatin structure. One question arises from this observation. Does chromatin of intermediate cell sizes show distinct nucleosome arrangements? Or does the growth in size cause fission yeasts to adapt their chromatin state in a continuous fashion? Are these changes in chromatin states proportional

to cell size? Besides, are the changes witnessed above, between large and normal proliferating cells truly size-dependent?

To obtain the high-resolution sequence of changes in chromatin structure while cell size increases, the population of increasing cell size is harvested every hour after 1NM-PP1 addition (Fig 3.5 and annex 3.1). A control condition consists in cultivating the cells with the drug but without glucose so that the cells are arrested in the cell cycle but not growing in size (for further details about the control, refer to chapter 2).

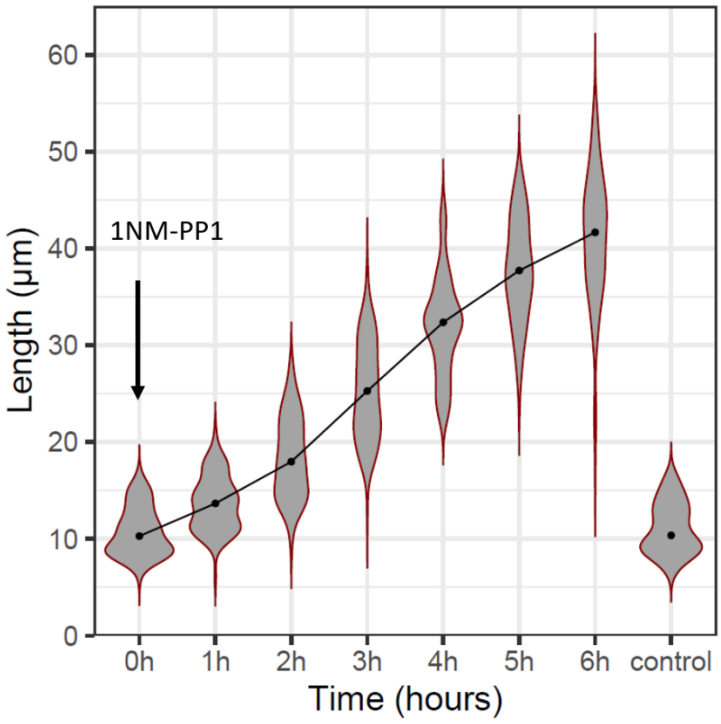


Figure 3.5: Cell size increase after addition of 1NM-PP1 (replicate 1). *The connected points are showing the median length of the population at each time-point.*

An additional lower MNase treatment is conducted to enrich the recovery of the protected DNA on the promoter. For this experiment, two replicates were conducted. As previously, the comparison between normal proliferating cells, large, arrested cells and normal arrested cells indicates that the nucleosome configuration around the TSS is dependent on the MNase condition and this for both replicates (Fig 3.6a, b, c, and annex 3.2a, b and c, replicate 2 is on the annexes).

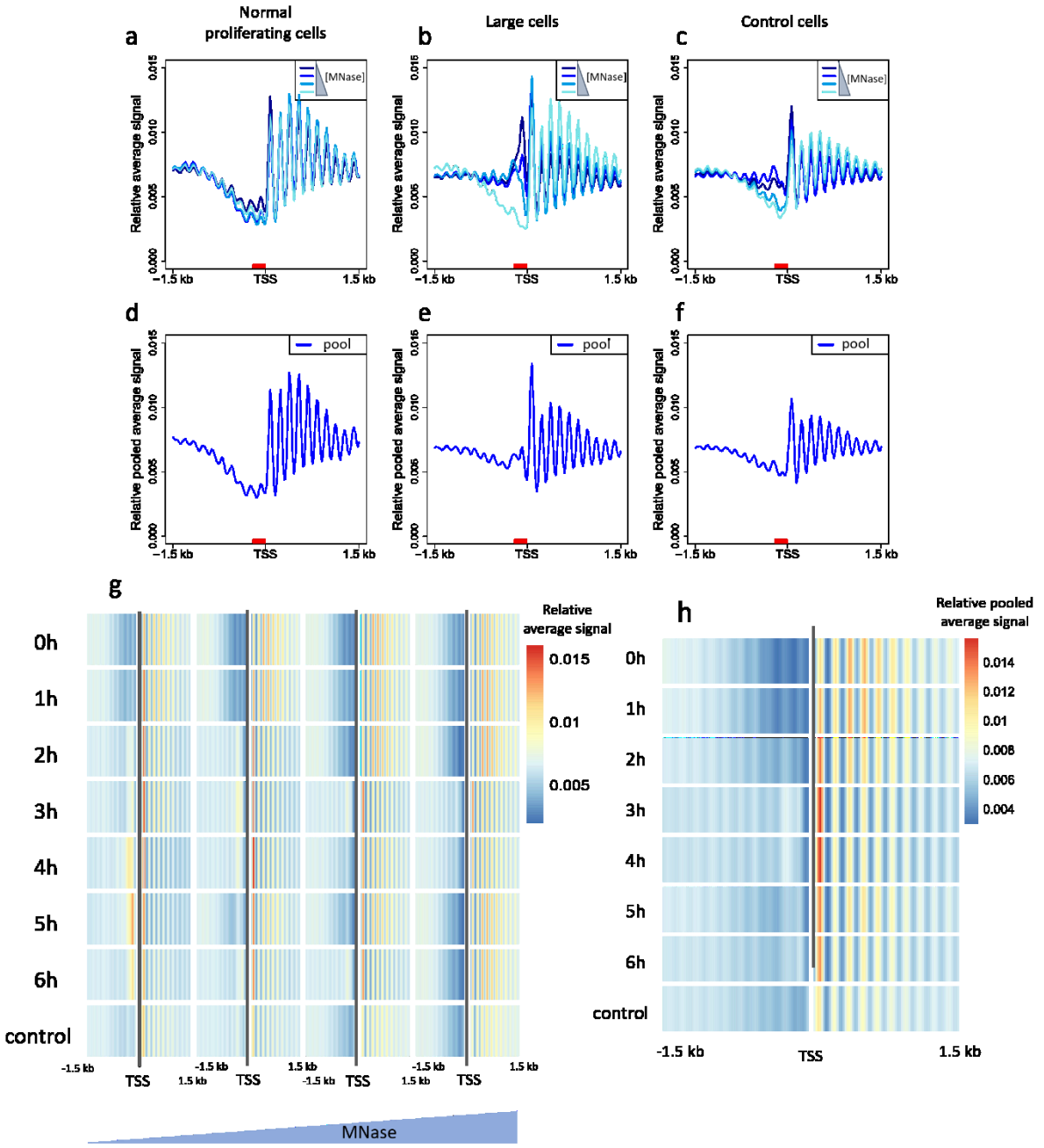


Figure 3.6: TSS-centred average chromatin profile of cells of growing size arrested with 1NM-PP1 for 6h (replicate 1). A-b*. Average profile of nucleosomes released at 4 different MNase concentrations for normal proliferating cells* ***(a)****, large arrested cells* ***(b)*** *and. normal arrested cells* ***(c)****. **D-f**. Average profile for pooled MNase concentrations for normal proliferating cells* ***(d)****, large arrested cells* ***(e)*** *and normal arrested cells* ***(f)****. **G-h.** Average profile for cells of growing size arrested with 1NM-PP1 and cells of normal arrested cells (control condition at the bottom of the heatmap) with four increasing MNase concentration (**g**), and with pooled MNase concentrations* ***(h).***

For normal cells, the promoter is still almost denuded of nucleosomes compared to the gene body in which the nucleosomes are well-organized (Fig 3.6a and d and annex 3.2a and d). For large cells, a mild MNase digestion allows the visualization of the same high and wide footprint

on the promoter which is also visible on the pooled average profile (Fig 3.6b and e and annex 3.2b and e). The +1 nucleosome reaches a relatively higher signal for the large cells when studying the pooled average profile which is also consistent with the first data set (Fig 3.6e and annex 3.2e) and might indicate the protection of the underlying DNA sequence by the factors leaving the high footprint on the promoter. The control sample is at an intermediate state, showing in some cases high signal on the promoter (Fig 3.6c and annex 3.2c, low MNase treatment) and other cases where it behaves more like wild type (Fig 3.6c and annex 3.2c, mid MNase treatment). The transcriptional state of these control cells is perturbed with different sets of stress genes being simultaneously repressed and induced upon glucose starvation (Oda et al., 2015). In addition, this change in transcriptional state is associated with a decrease of histone density on certain groups of genes (Oda et al., 2015). Hence, it is not surprising that the chromatin state for those glucose-starved cells deviates from the chromatin state in proliferating normal-sized cells. However, the difference of the chromatin structure is even more noticeable in large cells at a low MNase concentration. Taken together, this new result suggests that the unique chromatin organization on the promoter of large cells is indeed size-dependent. Moreover, the signal encountered at the promoter of large cells is much greater compared to the control condition, although glucose starvation also alters the transcriptional state (Oda et al., 2015). This observation introduces the possibility of distinct regulations of chromatin structure depending on the cellular state (size increase or glucose starvation).

The intermediate stages of the time-course, which correspond to intermediate cell size, add a layer of information. Visually, the pooled occupancies, together with the occupancies from different MNase digestions (Fig 3.6g and h and annex 3.2g and h), indicate a progressive change of signal on the promoter and the +1 nucleosome as the cell size increases. This supports the model of a continuous response of chromatin structure to cell size. Here, it is conceivable to observe the nucleosome positioning on the gene body for all the time-points since this metric is not affected by normalization as is the case for nucleosome occupancy. Interestingly, the oscillation of the signal in the gene body is aligned across all the conditions, showing the conservation of nucleosome positioning while cell size increases despite a progressively higher transcription rate.

The mean signal of the 200 bp promoter region and the 200 bp region corresponding to the +1 nucleosome is now analyzed (Fig 3.7 and annex 3.3). First, we will focus our attention on the promoter.

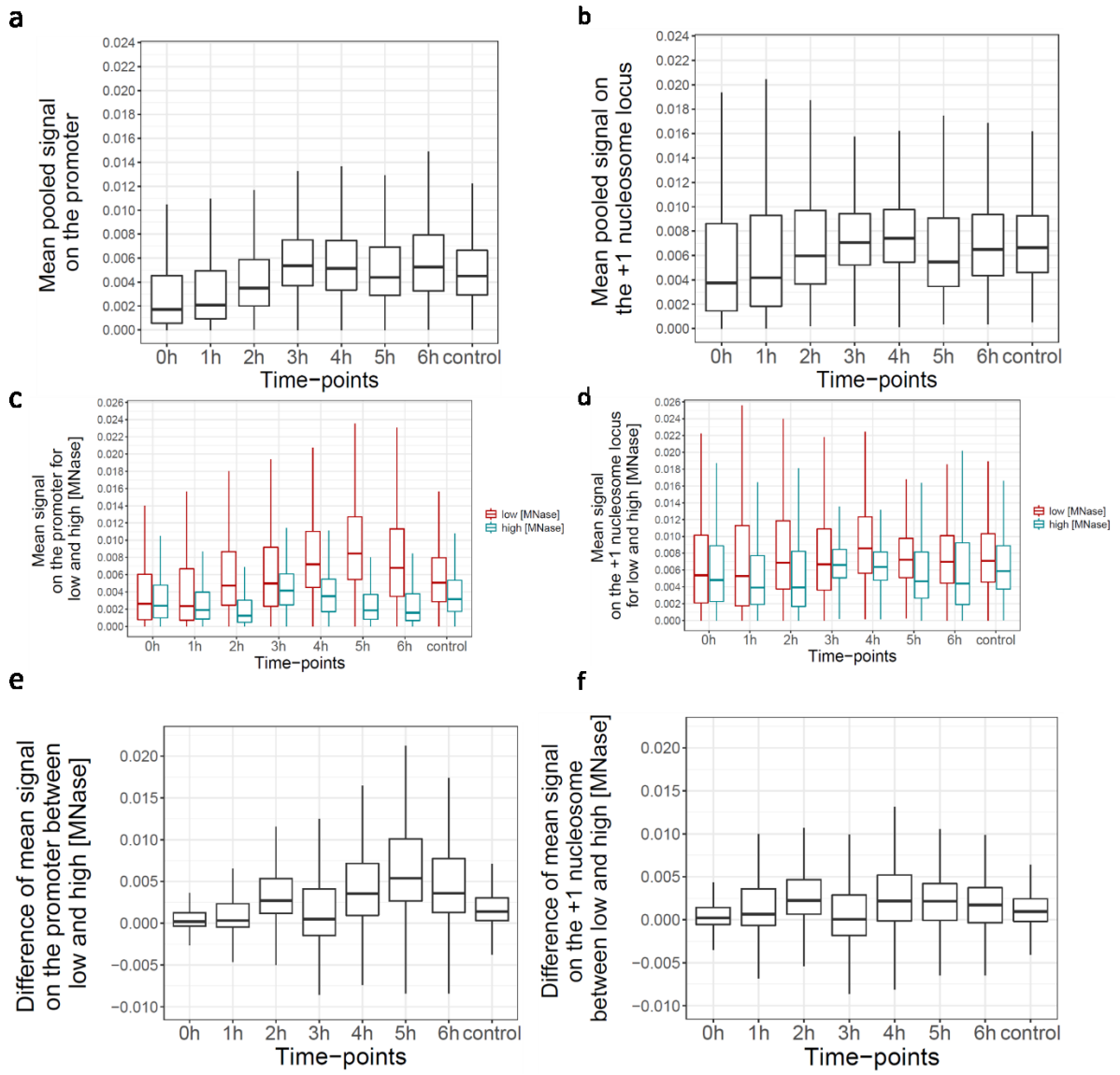


Figure 3.7: Mean signal on the promoter and +1 nucleosome for cells of growing size arrested with 1NM-PP1 for 6h (replicate 1). A.b. *Mean pooled occupancy on the promoter, 200 bp upstream of the TSS* ***(a)*** *and on the +1 nucleosome locus, 200 bp downstream of the TSS* ***(b)*** *(differences tested by paired t-test with n = 5887).* ***C-d.*** *Mean signal occupancy for the low and high MNase concentration on the promoter* ***(c)****, and on the +1 nucleosome locus (differences between time-points tested by paired t-test and between MNase concentrations by unpaired t-test, n = 5887)* ***(d)****.* ***E-f****. Differences between low and high MNase signal on the promoter* ***(e)*** *and on the +1 nucleosome locus* ***(f)*** *(differences tested by paired t-test with n = 5887).*

- Signal on the promoter locus

The distribution of pooled occupancies on the promoter does not show dramatic changes but an apparent stepwise trend, starting from 0.0018 at 0h to 0.0052 at 6h for replicate 1 (Fig 3.7a) and from 0.0025 at 0h to 0.0052 at 6h for replicate 2 (Annex 3.3a). The relative median occupancy is more than doubling between 0h and 6h, consistently with the first dataset (Fig 3.4a). Because the median cell length is four times more at 6h than at 0h (Fig 3.5), we can now exclude that the relative pooled occupancy at the promoter is strictly proportional to cell size. While the signal occupancy of the normal glucose-deprived cells is lower than at 6h, it is still significantly higher than at 0h (Fig 3.7a and Table 3.1a), recapitulating the intermediate state of the chromatin structure in these cells in terms of pooled occupancy. As mentioned above, glucose starvation leads to the expression of specific genetic programs (Oda et al., 2015) that are associated to a remodeling of the chromatin organization that might, for a specific set of genes, resemble the remodeling happening on the promoter of larger cells.

Table 3.1: Comparisons of pooled occupancies on the promoter between time-points tested with paired t-test for both replicates.

Mean pooled signal on the promoter				
			p-value	
Time-points	Group 1	Group 2	Rep 1	Rep 2
0h	0h	1h	< 2e-16	< 2e-16
	0h	2h	< 2e-16	< 2e-16
	0h	3h	< 2e-16	< 2e-16
	0h	4h	< 2e-16	< 2e-16
	0h	5h	< 2e-16	< 2e-16
	0h	6h	< 2e-16	< 2e-16
	0h	control	< 2e-16	< 2e-16
1h	1h	2h	< 2e-16	< 2e-16
	1h	3h	< 2e-16	< 2e-16
	1h	4h	< 2e-16	< 2e-16
	1h	5h	< 2e-16	< 2e-16
	1h	6h	< 2e-16	< 2e-16
	1h	control	< 2e-16	< 2e-16
2h	2h	3h	< 2e-16	< 2e-16
	2h	4h	< 2e-16	2.30E-12
	2h	5h	< 2e-16	< 2e-16
	2h	6h	< 2e-16	< 2e-16
	2h	control	< 2e-16	0.00058
3h	3h	4h	< 2e-16	< 2e-16
	3h	5h	< 2e-16	< 2e-16
	3h	6h	0.082	< 2e-16
	3h	control	< 2e-16	< 2e-16
4h	4h	5h	1.00E-09	< 2e-16
	4h	6h	< 2e-16	< 2e-16
	4h	control	< 2e-16	2.90E-12
5h	5h	6h	< 2e-16	3.00E-10
	5h	control	< 2e-16	< 2e-16
6h	6h	control	< 2e-16	< 2e-16

The pattern of pooled occupancies might be hiding other interesting patterns between low and high MNase concentrations that are averaged out when combined. That is why we now consider the lowest and the highest MNase concentrations separately. First, the lowest MNase digestion reports a progressive but clear increase in signal on the promoter. (Fig 3.7c and annex 3.3c). The occupancy on the promoter reaches the highest values from 5h with a median occupancy around 0.0085 vs 0.0025 at 0h for replicate 1 and 0.0075 vs 0.0045 at 0h for replicate 2 (Fig 3.7c and annex 3.3c). Again, this observation shows that the changes in occupancy at the promoter are not exactly proportional to cell size since cells at 6h are larger than those at 5h (Fig 3.5). It is also worth noting that the control sample shows measurements closer to 0h than 6h. This means that DNA-binding proteins, which are sensitive to high levels of MNase, are appearing progressively on the promoter as a function of cell size and not because of the blocking with 1NM-PP1. In addition, at a low MNase concentration and at late time-points, the distribution of the occupancy at this locus across the genome is wider compared to a high MNase concentration. This reflects a heterogeneous protection of the promoter among the genes that might be due to the different nature of the MNase-sensitive complexes occupying the promoter. While most promoters should be increasingly protected by the transcription machinery upon cell size increase (Swaffer et al., 2021b; Sun et al., 2020; Zhurinsky et al., 2010), this complex might be associated with non-canonical nucleosomes (unstable nucleosome or non-histone particle) (Kubik et al., 2015; Knight et al., 2014; Xi et al., 2011) in some cases and with non-histone proteins in other cases (Chereji et al., 2017).

Table 3.2: Comparisons of occupancies on the promoter between time-points tested with paired t-test for both replicates and at a low and a high MNase concentration.

Mean signal on the promoter at a low [MNase]					
[MNase]	Time-points		p-value		
			Rep 1	Rep 2	
low [MNase]	0h	0h 1h	< 2e-16	< 2e-16	
		0h 2h	< 2e-16	< 2e-16	
		0h 3h	< 2e-16	< 2e-16	
		0h 4h	< 2e-16	< 2e-16	
		0h 5h	< 2e-16	< 2e-16	
		0h 6h	< 2e-16	< 2e-16	
		0h control	< 2e-16	< 2e-16	
	1h	1h 2h	< 2e-16	< 2e-16	
		1h 3h	< 2e-16	< 2e-16	
		1h 4h	< 2e-16	< 2e-16	
		1h 5h	< 2e-16	< 2e-16	
		1h 6h	< 2e-16	< 2e-16	
		1h control	< 2e-16	< 2e-16	
	2h	2h 3h	< 2e-16	< 2e-16	
		2h 4h	< 2e-16	9.20E-06	
		2h 5h	< 2e-16	< 2e-16	
		2h 6h	< 2e-16	< 2e-16	
		2h control	< 2e-16	< 2e-16	
	3h	3h 4h	< 2e-16	< 2e-16	
		3h 5h	< 2e-16	< 2e-16	
		3h 6h	< 2e-16	< 2e-16	
		3h control	< 2e-16	< 2e-16	
	4h	4h 5h	< 2e-16	< 2e-16	
		4h 6h	0.00031	< 2e-16	
		4h control	< 2e-16	< 2e-16	
	5h	5h 6h	< 2e-16	< 2e-16	
		5h control	< 2e-16	< 2e-16	
	6h	6h control	< 2e-16	< 2e-16	

Mean signal on the promoter at a high [MNase]					
[MNase]	Time-points		p-value		
			Rep 1	Rep 2	
high [MNase]	0h	0h 1h	< 2e-16	< 2e-16	
		0h 2h	< 2e-16	< 2e-16	
		0h 3h	< 2e-16	0.0108	
		0h 4h	1.70E-12	< 2e-16	
		0h 5h	< 2e-16	< 2e-16	
		0h 6h	< 2e-16	< 2e-16	
		0h control	< 2e-16	< 2e-16	
	1h	1h 2h	< 2e-16	0.0815	
		1h 3h	< 2e-16	< 2e-16	
		1h 4h	< 2e-16	< 2e-16	
		1h 5h	< 2e-16	0.53654	
		1h 6h	1.20E-12	0.06827	
		1h control	< 2e-16	0.57576	
	2h	2h 3h	< 2e-16	< 2e-16	
		2h 4h	< 2e-16	< 2e-16	
		2h 5h	< 2e-16	5.70E-09	
		2h 6h	< 2e-16	< 2e-16	
		2h control	< 2e-16	0.01698	
	3h	3h 4h	< 2e-16	< 2e-16	
		3h 5h	< 2e-16	< 2e-16	
		3h 6h	< 2e-16	< 2e-16	
		3h control	< 2e-16	< 2e-16	
	4h	4h 5h	< 2e-16	< 2e-16	
		4h 6h	< 2e-16	< 2e-16	
		4h control	6.90E-06	< 2e-16	
	5h	5h 6h	7.50E-08	0.00017	
		5h control	< 2e-16	0.84733	
	6h	6h control	< 2e-16	0.10795	

To better understand the link between high and low MNase concentrations, I calculated the difference between those two MNase treatments (Fig 3.7e and annex 3.3e), as explained earlier (Fig 3.4e, f). Again, these measurements support the fact that cell size increase is associated with a progressive increase in the difference between high and low MNase conditions and with an increasing variability of the occupancy depending on the genes (Fig 3.7e and annex 3.3e). Nevertheless, a sudden and significant drop of signal at 3h for replicate 1 and 1h and 4h for replicate 2 is clearly visible (Fig 3.7e, annex 3.3e and Table 3.3). Hence, it seems that only a low MNase treatment is essential to reveal the net and gradual occupancy of the promoter as cell size increases. Besides, the increase in the difference is not seen for the control population for which the median occupancy is around 0.0025, similarly to 0h for replicate 2 (annex 3.3e) and a slightly more than 0h for replicate 1 (Fig 3.7e). Again, this suggests that control cells are undergoing a chromatin remodeling that is different to the one associated to cell size increase. However, there is a reproducible and significant increase in MNase-resistant proteins or complexes occupying the promoter at intermediate stages of the time-course (at 3h in replicate 1 and at 2h in replicate 2) (Fig 3.7c, annex 3.3c and Table 3.2) from 0.0025 at 0h to 0.0035 at 4h in replicate 1. This result is even more visible when analyzing the difference

between high and low MNase concentrations in replicate 1 where a significant decrease occurs at 3h (Fig 3.7e and Table 3.3). One category of MNase-resistant proteins contains histones, especially when they form the nucleosome that tightly binds the DNA. Thus, it would be interesting to visualize the pattern of histone expression. This idea will be explored later in the thesis (chapter 5).

Table 3.3: Comparisons of the occupancies on the promoter between time-points tested with paired t-test for both replicates and for the difference between a low and a high MNase concentration.

Difference of mean signal on the promoter between low and high [MNase]				
			p-value	
Time-points	Group 1	Group 2	Rep 1	Rep 2
0h	0h	1h	< 2e-16	< 2e-16
	0h	2h	< 2e-16	< 2e-16
	0h	3h	< 2e-16	< 2e-16
	0h	4h	< 2e-16	< 2e-16
	0h	5h	< 2e-16	< 2e-16
	0h	6h	< 2e-16	< 2e-16
	0h	control	< 2e-16	< 2e-16
1h	1h	2h	< 2e-16	< 2e-16
	1h	3h	< 2e-16	< 2e-16
	1h	4h	< 2e-16	< 2e-16
	1h	5h	< 2e-16	< 2e-16
	1h	6h	< 2e-16	< 2e-16
	1h	control	< 2e-16	< 2e-16
2h	2h	3h	< 2e-16	< 2e-16
	2h	4h	< 2e-16	2.70E-09
	2h	5h	< 2e-16	< 2e-16
	2h	6h	< 2e-16	< 2e-16
	2h	control	< 2e-16	< 2e-16
3h	3h	4h	< 2e-16	< 2e-16
	3h	5h	< 2e-16	< 2e-16
	3h	6h	< 2e-16	< 2e-16
	3h	control	5.10E-05	< 2e-16
4h	4h	5h	< 2e-16	< 2e-16
	4h	6h	< 2e-16	< 2e-16
	4h	control	< 2e-16	< 2e-16
5h	5h	6h	< 2e-16	< 2e-16
	5h	control	< 2e-16	< 2e-16
6h	6h	control	< 2e-16	< 2e-16

- Signal on the +1 nucleosome locus

Next, despite the comparatively smaller range of variation between time-points, the pooled occupancy in the +1 nucleosome window follows the same direction with a global relative increase in the median occupancy (Fig 3.7b and annex 3.3b). At 0h, the distribution of the signal distribution is wider for the +1 nucleosome compared to the promoter (Fig 3.7a, b and annex 3.3a, b). This must reflect the existence of groups of genes that exhibit distinct occupancy of the +1 nucleosome. The clustering analysis in the next chapter will explore this statement (chapter 4). Furthermore, it seems that the +1 nucleosome signal reaches at 3h (replicate 2) or

4h (replicate 1) the maximal signal (0.007) that is not exceeded in large cells (Fig 3.7b and annex 3.3b). This might suggest that after 3h (replicate 2) or 4h (replicate 1), the +1 nucleosome locus is protected in most cells and/or most genes so that the occupancy of the +1 nucleosome cannot increase more. Since the +1 nucleosome locus is in direct contact with the promoter, such protection might emerge from the occupancy of proteins on the promoter. Then, while the control median occupancy is lower than the one at 6h for replicate 2 (0.006 vs 0.0075), the other replicate (rep 1) reports similar distributions that are not statistically different between these two conditions (Table 3.4). This shows that the +1 nucleosome, when analyzed through pooled occupancy, is not strictly correlated to cell size but again, to the protection of the promoter which also increases in the control condition (Fig 3.7a, b and annex 3.3a, b). To support this, one can observe that replicate 1, showing similar distribution of +1 nucleosome occupancy between 6h and control, also shows a similar occupancy distribution on the promoter between 0h and 6h (Fig 3.7a, b).

Table 3.4: Comparisons of pooled occupancies of the +1 nucleosome between time-points tested with paired t-test for both replicates.

Mean pooled signal on the +1 nucleosome				
			p-value	
Time-points	Group 1	Group 2	Rep 1	Rep 2
0h	0h	1h	< 2e-16	0.372
	0h	2h	< 2e-16	< 2e-16
	0h	3h	< 2e-16	< 2e-16
	0h	4h	< 2e-16	< 2e-16
	0h	5h	< 2e-16	< 2e-16
	0h	6h	< 2e-16	< 2e-16
	0h	control	< 2e-16	< 2e-16
1h	1h	2h	< 2e-16	< 2e-16
	1h	3h	< 2e-16	< 2e-16
	1h	4h	< 2e-16	< 2e-16
	1h	5h	< 2e-16	< 2e-16
	1h	6h	< 2e-16	< 2e-16
	1h	control	< 2e-16	< 2e-16
2h	2h	3h	< 2e-16	< 2e-16
	2h	4h	< 2e-16	< 2e-16
	2h	5h	0.0118	4.20E-06
	2h	6h	0.00022	< 2e-16
	2h	control	0.01397	< 2e-16
3h	3h	4h	< 2e-16	< 2e-16
	3h	5h	< 2e-16	< 2e-16
	3h	6h	0.082	0.017
	3h	control	< 2e-16	< 2e-16
4h	4h	5h	1.00E-09	0.979
	4h	6h	< 2e-16	< 2e-16
	4h	control	< 2e-16	< 2e-16
5h	5h	6h	4.20E-06	< 2e-16
	5h	control	0.00076	< 2e-16
6h	6h	control	0.69565	< 2e-16

Nevertheless, it is worth considering the MNase treatments separately as well (Fig 3.7d and annex 3.3d). At 0h and for both high and low MNase treatments, the median signal is relatively higher in the +1 nucleosome compared to the promoter (Fig 3.7c, d and annex 3.3c, d). This reflects the relative nucleosome-depleted state of the promoter, followed by the positioned +1 nucleosome. While the low MNase concentration gives slightly higher occupancy than the highest MNase concentration (Fig 3.7d and annex 3.3d), the difference between the two conditions is very mild and does not compare to the difference seen on the promoter (Fig 3.7f and annex 3.3f). Besides, at the low MNase concentration the difference between 3h, 5h and control is not statistically different in replicate 1 and between 4h and 6h in replicate 2 (Table 3.5). This might be because the +1 nucleosome, formed by a histone octamer, is resistant to high levels of MNase. Therefore, according to these measurements, the strict dependence between the +1 nucleosome occupancy and cell size can be ruled out (Fig 3.7f and annex 3.3f).

Instead, the change in +1 nucleosome occupancy at the beginning of the time-course might be due to the increased protection of the promoter.

Table 3.5: Comparisons of +1 nucleosome occupancy between time-points tested with paired t-test for both replicates and at a low and a high MNase concentration.

Mean signal on the +1 nucleosome at a low [MNase]					
[MNase]	Time-points			p-value	
				Rep 1	Rep 2
low [MNase]	0h	0h	1h	< 2e-16	< 2e-16
		0h	2h	< 2e-16	< 2e-16
		0h	3h	< 2e-16	< 2e-16
		0h	4h	< 2e-16	< 2e-16
		0h	5h	< 2e-16	< 2e-16
		0h	6h	< 2e-16	< 2e-16
		0h	control	< 2e-16	< 2e-16
	1h	1h	2h	< 2e-16	< 2e-16
		1h	3h	< 2e-16	< 2e-16
		1h	4h	< 2e-16	< 2e-16
		1h	5h	5.40E-08	< 2e-16
		1h	6h	1.60E-05	< 2e-16
		1h	control	1.10E-13	< 2e-16
	2h	2h	3h	< 2e-16	< 2e-16
		2h	4h	< 2e-16	9.70E-14
		2h	5h	< 2e-16	0.0023
		2h	6h	< 2e-16	< 2e-16
		2h	control	< 2e-16	< 2e-16
	3h	3h	4h	< 2e-16	< 2e-16
		3h	5h	0.042	< 2e-16
		3h	6h	5.20E-11	< 2e-16
		3h	control	0.239	< 2e-16
	4h	4h	5h	< 2e-16	< 2e-16
		4h	6h	< 2e-16	0.5505
		4h	control	< 2e-16	< 2e-16
	5h	5h	6h	1.40E-06	< 2e-16
		5h	control	0.313	< 2e-16
	6h	6h	control	1.70E-06	< 2e-16

Mean signal on the +1 nucleosome at a high [MNase]					
[MNase]	Time-points			p-value	
				Rep 1	Rep 2
high [MNase]	0h	0h	1h	< 2e-16	0.0079
		0h	2h	< 2e-16	< 2e-16
		0h	3h	< 2e-16	< 2e-16
		0h	4h	1.30E-13	< 2e-16
		0h	5h	0.00013	< 2e-16
		0h	6h	0.00018	< 2e-16
		0h	control	< 2e-16	< 2e-16
	1h	1h	2h	1.10E-07	< 2e-16
		1h	3h	< 2e-16	< 2e-16
		1h	4h	< 2e-16	< 2e-16
		1h	5h	2.80E-08	< 2e-16
		1h	6h	< 2e-16	< 2e-16
		1h	control	< 2e-16	< 2e-16
	2h	2h	3h	< 2e-16	< 2e-16
		2h	4h	< 2e-16	< 2e-16
		2h	5h	< 2e-16	2.10E-06
		2h	6h	< 2e-16	< 2e-16
		2h	control	< 2e-16	3.70E-06
	3h	3h	4h	< 2e-16	2.60E-07
		3h	5h	< 2e-16	< 2e-16
		3h	6h	< 2e-16	< 2e-16
		3h	control	1.00E-13	8.20E-16
	4h	4h	5h	< 2e-16	< 2e-16
		4h	6h	1.90E-05	< 2e-16
		4h	control	0.01807	0.2299
	5h	5h	6h	2.80E-14	< 2e-16
		5h	control	< 2e-16	0.0056
	6h	6h	control	5.20E-14	2.50E-16

Table 3.6: Comparisons of the +1 nucleosome occupancy on the promoter between time-points tested with paired t-test for both replicates and for the difference between a low and a high MNase concentration.

Difference of mean signal on the +1 nucleosome locus between low and high [MNase]				
			p-value	
Time-points	Group 1	Group 2	Rep 1	Rep 2
0h	0h	1h	< 2e-16	< 2e-16
	0h	2h	< 2e-16	0.439
	0h	3h	< 2e-16	< 2e-16
	0h	4h	< 2e-16	5.50E-15
	0h	5h	< 2e-16	0.351
	0h	6h	< 2e-16	0.001
	0h	control	< 2e-16	3.60E-07
1h	1h	2h	< 2e-16	< 2e-16
	1h	3h	< 2e-16	< 2e-16
	1h	4h	< 2e-16	< 2e-16
	1h	5h	0.0167	< 2e-16
	1h	6h	0.0372	< 2e-16
	1h	control	1.70E-08	1.10E-08
2h	2h	3h	< 2e-16	< 2e-16
	2h	4h	1.30E-09	< 2e-16
	2h	5h	< 2e-16	7.69E-01
	2h	6h	< 2e-16	1.60E-06
	2h	control	< 2e-16	3.00E-13
3h	3h	4h	< 2e-16	< 2e-16
	3h	5h	< 2e-16	< 2e-16
	3h	6h	7.00E-06	< 2e-16
	3h	control	7.90E-05	< 2e-16
4h	4h	5h	< 2e-16	< 2e-16
	4h	6h	< 2e-16	< 2e-16
	4h	control	< 2e-16	< 2e-16
5h	5h	6h	< 2e-16	1.30E-07
	5h	control	< 2e-16	2.60E-09
6h	6h	control	0.0036	< 2e-16

Altogether, these data reinforce the idea that the promoter, initially almost denuded of nucleosomes, is occupied by proteins in a cell size dependent manner. Importantly, it seems that the progressive change of the chromatin structure is correlated to the gradual size increase in the cell. Since it was found previously that RNAPII occupancy and transcription initiation rates correlate with cell size, one likely component responsible of the change is chromatin architecture is the RNAPII itself. However, its efficient recruitment at the promoter of large cells probably requires other actors that would also bind the DNA at this locus. The change of chromatin composition at the promoter seems to affect the +1 nucleosome structure, although those changes remain minor compared to those on the promoter. Additionally, no main change in nucleosome positioning occurs in the gene body. This suggests that chromatin structure at the promoter might be the main regulatory player to coordinate transcription rate to cell size.

3.4. MNase accessibility: MACC

In the year 2016, a new protocol relying on MNase digestions was described in fly cells (Mieczkowski et al., 2016). This innovative approach permits integration of the data from several MNase concentrations into one new measurement called MACC for MNase accessibility. Briefly, the read frequencies for each genomic bin are plotted against the four MNase concentrations. Then, the slope of the linear fit of the resulting curve is calculated and defined as the MACC score (Fig 3.8).

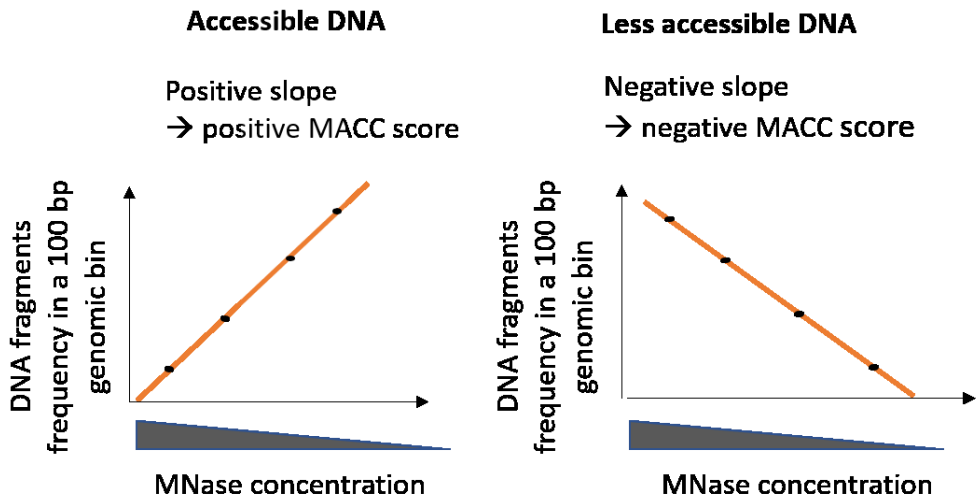


Figure 3.8: MACC score method can be used on MNase-seq to measure DNA accessibility at each locus. *Adapted from Miezckowski et al., 2016.*

In a genomic environment where DNA is accessible, a mild MNase digestion is be able to recover the mononucleosomal fragments; on the contrary, an extensive MNase digestion might degrade all DNA in this region so that the signal after the digestion is low. The situation is reverse in a less accessible genomic region. The MACC score is used as a proxy for DNA accessibility to MNase. Consequently, all loci can be described according to their accessibility to MNase which reflects the ability of nuclear factors to bind the chromatinized DNA. In their study, Mieczkowski and colleagues profile the MACC score of all the genome using an MNase digestion followed or not by a ChIP-seq against Histone 3 (Mieczkowski et al., 2016). These two variants of the technique allow the distinction between histone binding proteins that form the nucleosome and the non-histone binding proteins such as TFs or chromatin remodelers.

However, the outcome of these two versions is very similar since most of the proteins contacting the DNA are nucleosomes (Mieczkowski et al., 2016). Nevertheless, when overlapping MACC results to previous ChIP-seq data, they saw that a local increase in MACC values on the promoter was often associated with the binding of non-histone proteins, even though those factors also occlude DNA. The authors saw that regions with high pooled nucleosome occupancy could have either high MACC values or low MACC values. This study, in combination with others (eg, Mueller et al. 2017), brought the idea that changes in chromatin accessibility are not necessarily accompanied with changes in nucleosome occupancy and consequently, that MACC is a better predictor of chromatin dynamics. However, cell size increase leads to an increase in occupancy in the promoter, meaning that the effect on the chromatin is sufficient to be observed with the occupancy metric. Therefore, what could the MACC score calculated for cells of increasing size reveal? Could we witness a size-dependent change in the +1 nucleosome structure that was not visible with the occupancy metrics?

In the present work, because *S. pombe*'s genome is smaller than fly's, the MACC method is adapted so that MACC scores are calculated in 20 bp bins (instead of 50 bp in fly) in 1.5 kb, either side of the TSS. Strikingly, the MACC average profile for normal proliferating cells recapitulates the phased pattern seen on the gene body with the nucleosome average profile (Fig 3.9a).

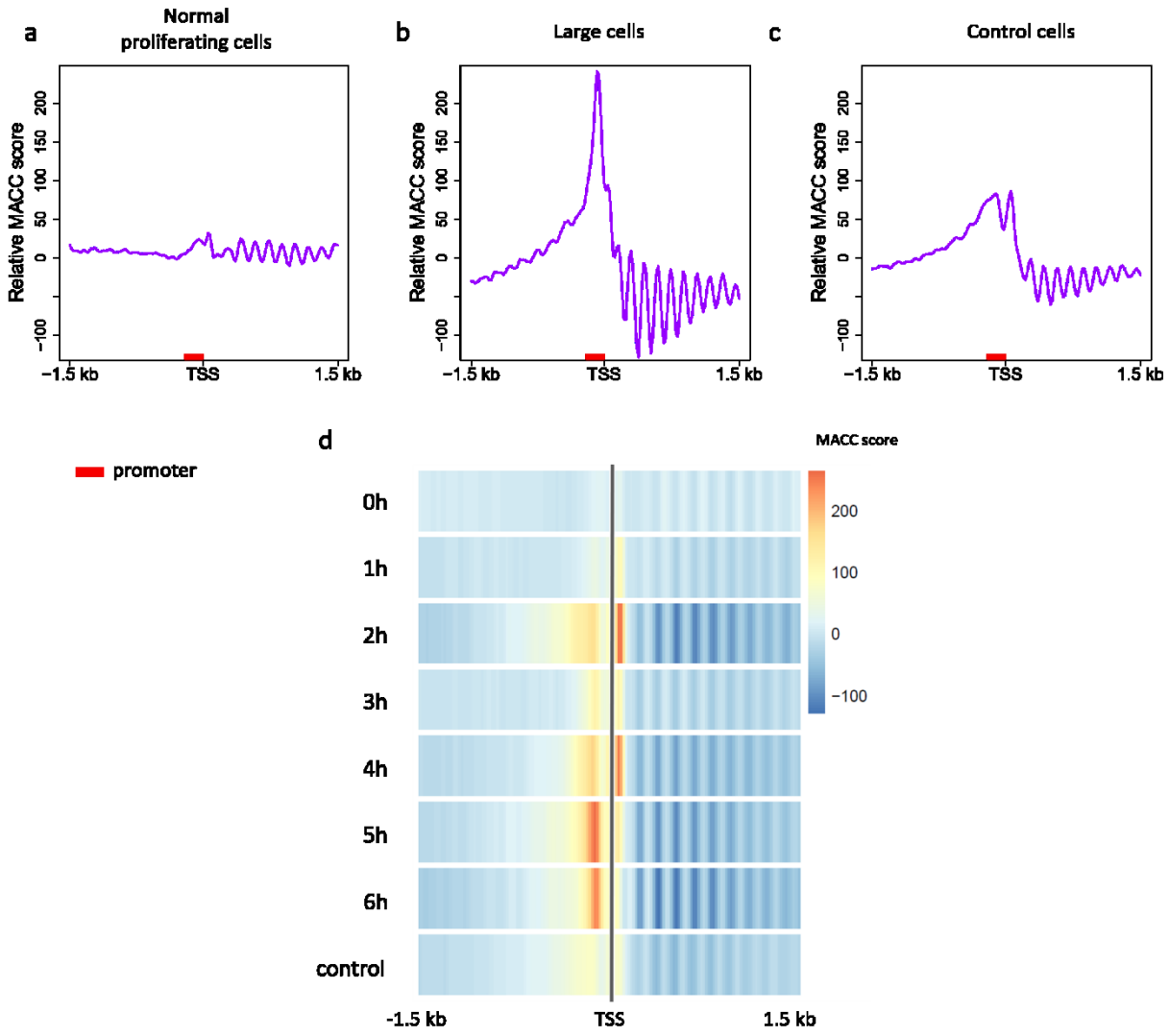


Figure 3.9: TSS-centred average MACC profile of cells of growing size arrested with 1NM-PP1 for 6h (replicate 1). A-c*. MACC average profile for normal proliferating cells* ***(a)****, large cells arrested 6h with 1NM-PP1* ***(b)*** *and normal arrested cells* ***(c).***

However, while the signal on the promoter is low on the nucleosome average profile, it is the opposite for the MACC score: the signal on the promoter and the +1 nucleosome is relatively higher than in the gene body. This highlights the accessibility of the NDR around the TSS (Mieczkowski et al. 2016), potentially due to a faster nucleosome turnover or a non-histone binding protein having a lower residence time at this locus. Both could be intertwined if the non-nucleosome protein is a remodeling factor that favors nucleosome turnover.

For large cells, the amplitude of the signal is greater on the whole region (Fig 3.9b). Notably, the promoter and the +1 nucleosome, while presenting the same tendency as normal cells,

exhibit a much higher MACC signal. Finally, the MACC average profile for the control condition is very similar to normal cells in terms of signal amplitude, showing that the arrest with 1NM-PP1 is not responsible for the pattern seen in large cells (Fig 3.9c and annex 3.4c). In summary, the trend for normal proliferating cells and large arrested cells is the same, with higher accessibility close to the TSS. However, the accessibility of the promoter is much higher in large cells compared to normal cells, as if the events occurring in normal circumstances were exacerbated in large cells. As mentioned above, high MACC peaks are often associated with non-histone factors such as RNAPII, consistently with the higher transcription rate occurring in large cells. In addition, a high MACC peak might indicate the binding of chromatin remodeling complex that could induce more accessibility at this locus. Cells of increasing size present an interesting sequential change in chromatin accessibility around the TSS (Fig 3.9d and annex 3.4d). Indeed, the promoter seems to be more accessible as the cells are growing in size and in this occurrence, it appears to also be the case for the +1 nucleosome.

The same method as before, consisting in calculating the total MACC values in different loci (200 bp on the promoter or 200 bp upstream of the TSS) can be applied. The distributions of the MACC values are displayed for each time-point (Fig 3.10 and annex 3.5). In both replicates, cells harvested at 5h show the maximal median accessibility, between 0.2 and 0.25 compared to 0.12 for replicate 1 (Fig 3.10a) or 0 for replicate 2 at 0h. For replicate 2 (annex 3.5a), time point 1h has the lowest median accessibility (0.03) with many genes having very close MACC values to the median. When comparing the lowest MACC value (either 0h or 1h) with the highest one (5h), no overlap between the interquartile range of the 2 conditions can be observed, showing that for most genes, promoter of cells at late time-points are more accessible than cells at earlier time-points. In this case, the MACC values can be negative, describing regions of the genome having a close chromatin structure. Because the chromatin of *S. pombe* is known to be generally open, those inaccessible loci are in a strong minority.

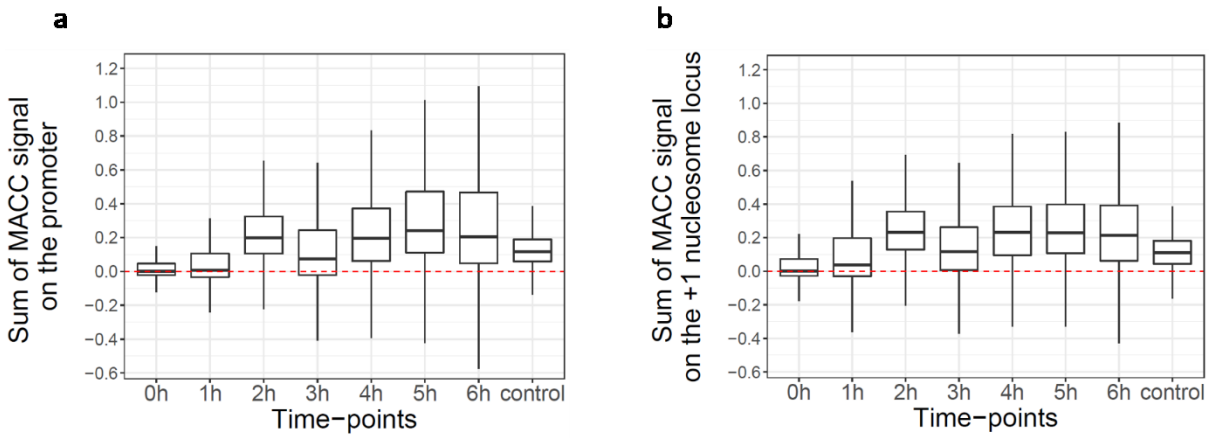


Figure 3.10: Total MACC profile for cells of growing size arrested for 6h with 1NM-PP1 (replicate 1). *Total MACC signal on the promoter, 200 bp upstream of the TSS* ***(a)****, and on the +1 nucleosome locus, 200 bp downstream of the TSS* ***(b)*** *(paired t-test with n = 5887)*

Considering the whole time-course rather than the two extreme time-points, the accessibility on the promoter is suddenly dropping at 1h and 3h for replicate 1 at 0 and 0.05 respectively and at 1h and 4h in replicate 2 (around 0) (Fig 3.10a and annex 3.5a). This is in contrast with what was observed with the pooled occupancy and the occupancy at low MNase conditions (Fig 3.7a, c and annex 3.3a, c). However, this trend is similar to the trend of the difference of occupancies between high and low MNase concentrations (Figure 3.7e and annex 3.3e). In other words, in this scenario wherein cell size increases following addition of a drug, the trend of promoter accessibility recapitulates the trend observed when subtracting the occupancy signal from a low MNase concentration to the one obtained after a high MNase concentration. This means that both metrics can be used to measure the change of chromatin structure on the promoter. Again, it is possible that a difference in histone expression levels as well as their localization on the DNA could provoke a decrease of accessibility at intermediate time-points, but this will be further explored later in the thesis. Then, the median and distribution of chromatin accessibility is similar between normal proliferating cells and normal arrested cells, showing that cell size might be intimately linked to primary chromatin structure on the promoter. Besides, the distribution of the MACC values becomes more spread towards the end of the time-course thus revealing the variability of promoter accessibility in the genome of large cells. This suggests that cell size increase is associated with a genome-wide variable chromatin context around the TSS, which was initially similar for all the genes at 0h.

The locus in which sits the +1 nucleosome follows the same trend as for the promoter with a wavy pattern of accessibility throughout the time-course and with higher accessibility and variability at later time-points (Fig 3.10 and annex 3.5b). At 0h, the median MACC value of the +1 nucleosome is the same as in the promoter (~0 for replicate 1 and ~0.1 for replicate 2). Conversely, pooled occupancy at this locus is higher than in the promoter. This situation is the reciprocal of what was described before where the same values of MACC correspond to different values of pooled occupancy (Mieczkowski et al., 2016): here, different nucleosome occupancy can result in the same MACC score, demonstrating that occupancy is still a valid metric to study changes in chromatin structure. This time again, the median accessibility of the +1 nucleosome is higher at late time-points, more precisely at 5h (0.2 vs 0 at 0h for replicate 1 and 0.25 vs 0.1 at 0h for replicate 2).

In comparison to the promoter, accessibility around the +1 nucleosome is diminishing slightly at 6h. Furthermore, its increase in accessibility with cell size is not continuous because time-point 3h for replicate 1 and 4h for replicate 2 present a sudden and significant drop with a median accessibility of 0.1 for replicate 1 at 3h and for replicate 2 at 4h (Fig 3.10b and Table 3.7). Finally, median accessibility of the +1 nucleosome locus decreases in the control condition, reaching ~0.12 in replicate 1 and 0.11 in replicate 2 (Fig 3.10b and annex 3.5b). These observations suggest that the +1 nucleosome undergoes different states of accessibility while cell size increases.

Table 3.7: Comparison of the sum of MACC signal for the promoter and the +1 nucleosome between time-points tested by paired t-test.

Sum of MAC signal on the promoter					Sum of MAC signal on the +1 nucleosome				
			p-value					p-value	
Time-points	Group 1	Group 2	Rep 1	Rep 2	Time-points	Group 1	Group 2	Rep 1	Rep 2
0h	0h	1h	< 2e-16	< 2e-16	0h	0h	1h	< 2e-16	< 2e-16
	0h	2h	< 2e-16	< 2e-16		0h	2h	< 2e-16	< 2e-16
	0h	3h	< 2e-16	< 2e-16		0h	3h	< 2e-16	< 2e-16
	0h	4h	< 2e-16	< 2e-16		0h	4h	< 2e-16	< 2e-16
	0h	5h	< 2e-16	< 2e-16		0h	5h	< 2e-16	< 2e-16
	0h	6h	< 2e-16	< 2e-16		0h	6h	< 2e-16	< 2e-16
	0h	control	< 2e-16	< 2e-16		0h	control	< 2e-16	0.0087
1h	1h	2h	< 2e-16	< 2e-16	1h	1h	2h	< 2e-16	< 2e-16
	1h	3h	< 2e-16	< 2e-16		1h	3h	< 2e-16	< 2e-16
	1h	4h	< 2e-16	< 2e-16		1h	4h	< 2e-16	< 2e-16
	1h	5h	< 2e-16	< 2e-16		1h	5h	< 2e-16	< 2e-16
	1h	6h	< 2e-16	< 2e-16		1h	6h	< 2e-16	< 2e-16
	1h	control	< 2e-16	< 2e-16		1h	control	6.60E-09	< 2e-16
2h	2h	3h	< 2e-16	< 2e-16	2h	2h	3h	< 2e-16	< 2e-16
	2h	4h	6.40E-09	9.40E-05		2h	4h	4.50E-10	< 2e-16
	2h	5h	< 2e-16	< 2e-16		2h	5h	< 2e-16	< 2e-16
	2h	6h	< 2e-16	< 2e-16		2h	6h	0.2952	< 2e-16
	2h	control	< 2e-16	0.002		2h	control	< 2e-16	< 2e-16
3h	3h	4h	< 2e-16	< 2e-16	3h	3h	4h	< 2e-16	< 2e-16
	3h	5h	< 2e-16	< 2e-16		3h	5h	< 2e-16	< 2e-16
	3h	6h	< 2e-16	0.084		3h	6h	< 2e-16	2.30E-15
	3h	control	1.50E-13	< 2e-16		3h	control	< 2e-16	< 2e-16
4h	4h	5h	< 2e-16	< 2e-16	4h	4h	5h	< 2e-16	< 2e-16
	4h	6h	< 2e-16	< 2e-16		4h	6h	0.0022	< 2e-16
	4h	control	< 2e-16	4.82E-01		4h	control	< 2e-16	< 2e-16
5h	5h	6h	< 2e-16	< 2e-16	5h	5h	6h	< 2e-16	< 2e-16
	5h	control	< 2e-16	< 2e-16		5h	control	< 2e-16	< 2e-16
6h	6h	control	< 2e-16	< 2e-16	6h	6h	control	< 2e-16	< 2e-16

The MACC score analysis is a new way to describe chromatin state. Initially applied in WT conditions to track the changes of accessibility genome-wide, it is used here to explore cell size-dependent dynamics of DNA accessibility in the vicinity of the TSS. Occupancy being a component of accessibility, it was expected that a size-dependent changes of occupancy would provoke a change in accessibility. Indeed, when the size of cells increases, the accessibility in the promoter and the +1 nucleosome locus, measured through MACC scores, varies. However, the trend is somehow different and despite a higher MACC scores at large size, the intermediate time-points show that accessibility is not directly linked to cell size. This corroborates the idea that MACC values and occupancy assess distinctive features of the chromatin. Because occupancy can capture differences invisible with the MACC score (for example when comparing the +1 nucleosome with the promoter in normal cells), occupancy measures should still be used but in combination with the MACC scores.

3.5. Discussion

Altogether, these data demonstrate the close relationship between cell size and the 2D chromatin organization through nucleosomes. In this chapter, I focused on an important regulatory component of genes, the promoter, and on the first barrier encountered by Pol II during transcription, the +1 nucleosome.

In a normal situation, the promoter is generally depleted of nucleosomes and the gene body presents a regular array of nucleosomes queuing downstream of a well-positioned +1 nucleosome (Moyle-Heyrman et al., 2013; Soriano et al., 2013; Lantermann et al., 2010). This definite chromatin structure found in normal proliferating cells is known to favor a controlled transcription with the recruitment of TF, PIC and RNAPII on the nucleosome-depleted promoter and with the nucleosomal barrier that avoids aberrant and pervasive transcription (DeGennaro et al., 2013; Kornberg and Lorch, 2020). Besides, this conformation around the TSS is well reported in this chapter (**Fig 3.1a, b**) with occupancy measurements, for cells cultivated without 1NM-PP1. Specifically, these measurements show that +1 nucleosome occupancy is higher than occupancy on the promoter for both pooled occupancies and occupancy at extreme MNase concentrations (**Fig 3.7,** 0h). However, when calculating MACC score for both loci, this difference is lost. This is surprising knowing that accessibility by MACC score has been described as a better predictor of DNA accessibility, as it could vary where occupancy remains stable. For example, this has been observed on genes of distinct transcriptional status where highly expressed genes report a higher MACC score compared to lowly expressed genes (Mueller et al. 2017). However, it is now acknowledged that accessibility values can change in different chromatin structures that present various occupancy levels (Klemm et al., 2019). Finally, it was observed here that the MACC score trend was highly similar to the trend of occupancy at a low MNase treatment minus occupancy at a high MNase treatment at the promoter, showing that the occupancy metrics at this locus could be used to achieve the same purpose.

Elongating *cdc2-asM17* cells exhibit a striking change in chromatin state on the promoter (**Fig 3.2, 3.6b, e and 3.7**). A progressive but unmistakable footprint is left on the promoter and is

visible for the lowest MNase concentration but not apparent for the highest one. The height and width of this footprint is puzzling (Fig 3.2, 3.6b) as it conflicts with the conventional concept of a low nucleosome occupancy or even an absence of nucleosome at this location, as is the case for normal-sized cells (Fig 3.1 and 3.6a) (Lantermann et al., 2010, 2009). The sensitivity to high MNase concentrations suggests that the proteins responsible for the footprint are not canonical nucleosomes, as opposed to those found on the gene body regardless of the MNase concentration (Weiner et al., 2010; Henikoff et al., 2009).

More recent studies indeed revealed that the promoters of cells were in fact occupied by proteins that are not nucleosome, therefore protecting the DNA only upon a mild MNase treatment. The nature of these binding factors has been highly debated. Some studies identified non-canonical nucleosomes (Brahma and Henikoff, 2019; Kubik et al., 2015; Knight et al., 2014; Xi et al., 2011) whereas others described non-nucleosomal proteins (Chereji et al., 2017).

The conditions of MNase used here do not recover those MNase-sensitive proteins in normal proliferating cells, as the footprint of such proteins is not clearly visible even for the lowest MNase concentration. This can arise from different scenarios: the binding factors are not identical to what has been described before; fission yeasts possess them only in large cells; the average profile stems from different patterns of promoter structure across the genome. Indeed, the variability (wide distribution) of occupancy signal in larger cells shows that all genes might respond differently to the addition of 1NM-PP1; this will be explored in the next chapter (chapter 4).

The accessibility is also higher on the promoter at some stages where the cells are bigger but, strikingly, the increase is not gradual. As the size increase occurs in the same cells, one can imagine a model in which sequential events happen in the vicinity of the promoter. The residence time of the proteins occupying the promoter could increase while the cells are growing; this would lead to a gradual increase in occupancy. Moreover, most genes are scaling and consequently, the ratio DNA-to-RNAPII, TFs, chromatin remodelers and any other proteins able to contact the DNA is getting lower as cells are elongating. At large cells (~40 µm), the histone concentration goes down (Claude et al., 2021; Jiang and Zhang, 2021; Lanz et al., 2021; Swaffer et al., 2021a; Martínez Segura, 2017). These two observations might indicate that the

competition for DNA at certain loci is in favor of the non-histone proteins that are able to perturb the local chromatin structure in large cells (Fig 3.11). This group of non-histone proteins are likely to contain the RNAPII (Swaffer et al., 2021b; Sun et al., 2020; Zhurinsky et al., 2010). In addition, they can be GRFs which were described to compete with the formation of a nucleosome at the promoter and, with the cooperation with RSC, to shift the +1 nucleosome aside (Mivelaz et al., 2020; Kubik et al., 2015). Moreover, the pool of proteins contacting the promoter is heterogeneous and the composition of the proteins might differ depending on the stages of the elongation. For instance, although the concentration of histones negatively correlates with cell size (Claude et al., 2021; Lanz et al., 2021; Martínez Segura, 2017), histone abundance could be different depending on the stages of elongation. Is it possible that, at intermediate time-points, histones are rising transiently, which consequently decreases the promoter accessibility? Are other proteins expressed differently during the time-course having higher affinity with the DNA than the rest? In any case, the diversity in the pool of proteins binding the promoter during the time-course might provoke a wavy pattern in median accessibility.

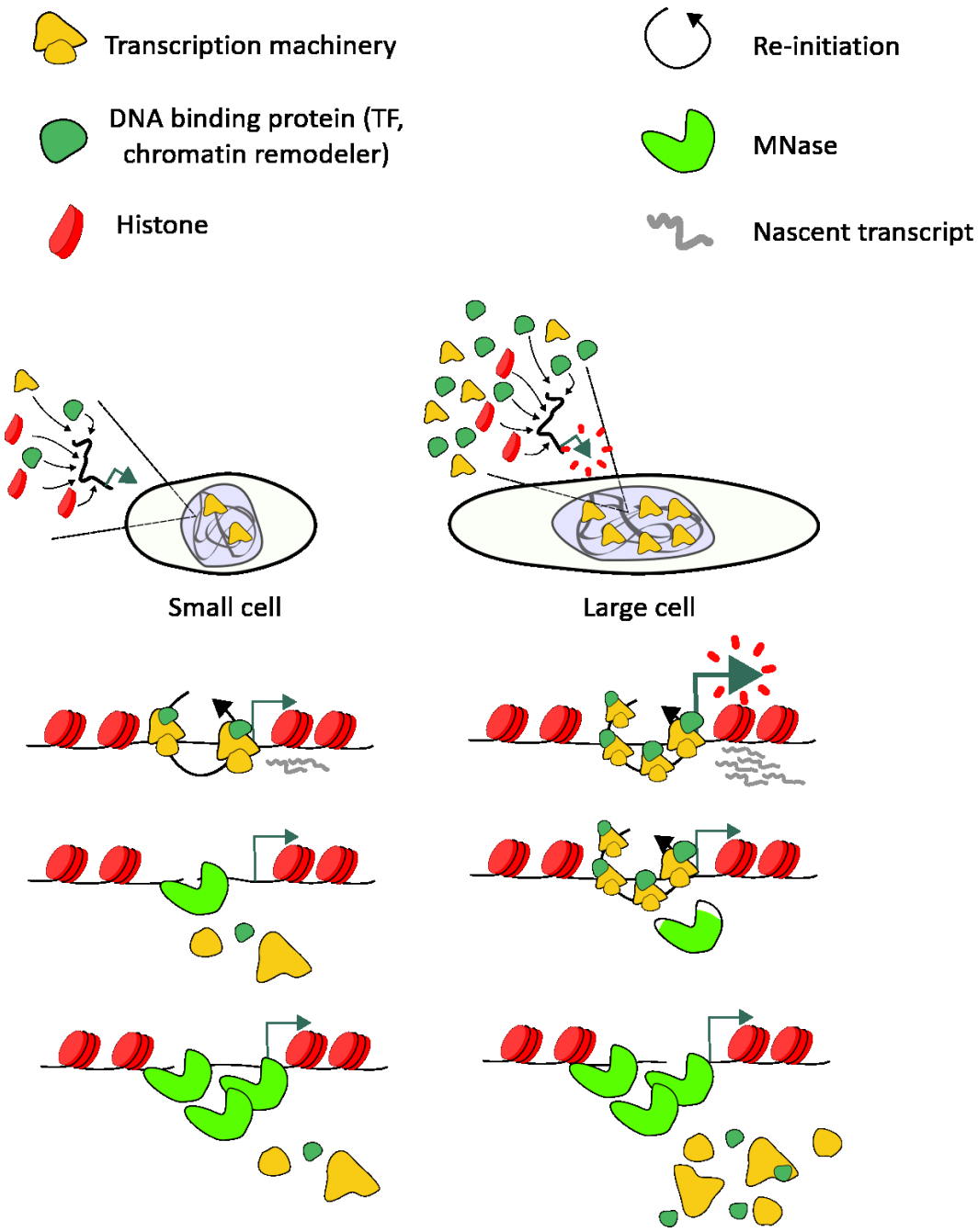


Figure 3.11: Potential model explaining the increase in MNase-sensitive factors at the promoter of larger cells. *In large cells the accumulation of RNAPII and TFs in the nucleus competes with histones to bind promoter. Those factors are resistant to a low MNase concentration in large cells only. At the population level, the frequency of promoter binding by RNAPII and TFs is higher in large cells compared to small cells. This would lead to a higher efficiency of transcription initiation in large cells.*

Why would the promoter of elongating cells be more and more occupied by MNase-sensitive factors? It was previously established that RNAPII occupancy on the DNA is proportional to cell size (Sun et al., 2020; Zhurinsky et al., 2010). As the volume of these cells increases, they need to adjust their transcription rate so that RNA molecules scale with cell size (Sun et al., 2020;

Martínez Segura, 2017; Zhurinsky et al., 2010). Hence, one can imagine that cell elongation is accompanied by a reshuffling of the chromatin structure in regulatory regions such as the promoter and a subsequent increase in RNAPII occupancy at this locus that would promote transcription scaling to cell size. However, the median occupancy by MNase-sensitive factors at the promoter is not exactly proportional to cell size. This observation might indicate that the whole pool of scaling proteins binding to the DNA (e.g. TF) are binding the promoter at each stage of the cell elongation. Other studies show that the mutants for Rsc1 and Snf21, which are both part of the RSC chromatin remodeler complex, present a cellular elongation phenotype (Yague-Sanz et al., 2017). The RSC chromatin complex is known to evict nucleosomes on the promoter of genes to create the NDR and, consequently, to activate transcription (Kornberg and Lorch, 2020; Ganguli et al., 2014; Floer et al., 2010). One possibility is that the RSC complex positively controls cdc25 which would induce a cell cycle delay and consequently a division at large size. However, another role for RSC complex was discovered in 2016: by creating a NDR, this chromatin remodeler is facilitating the condensing loading at nucleosome-depleted site during mitosis (Toselli-Mollereau et al., 2016). The latter mechanism, which was confirmed later (Muñoz et al., 2019) elegantly shows how nucleosome regulation by RSC could be linked to DNA condensation, which is mainly attributed to condensins (Kornberg and Lorch 2020). It seems that the incapacity of loading condensin on the DNA at NDR, as well as the inactivation of cdc2, both lead to the same phenotype which is cell elongation. Furthermore, a recent study in the *cdc2-asM17* model depicts condensin as an important regulator of the volume taken by the DNA during interphase (Kakui et al., 2020). Thus, all this work might open the path to a model where condensin, nucleosomes and size control are connected, although the role of condensin is still very mysterious.

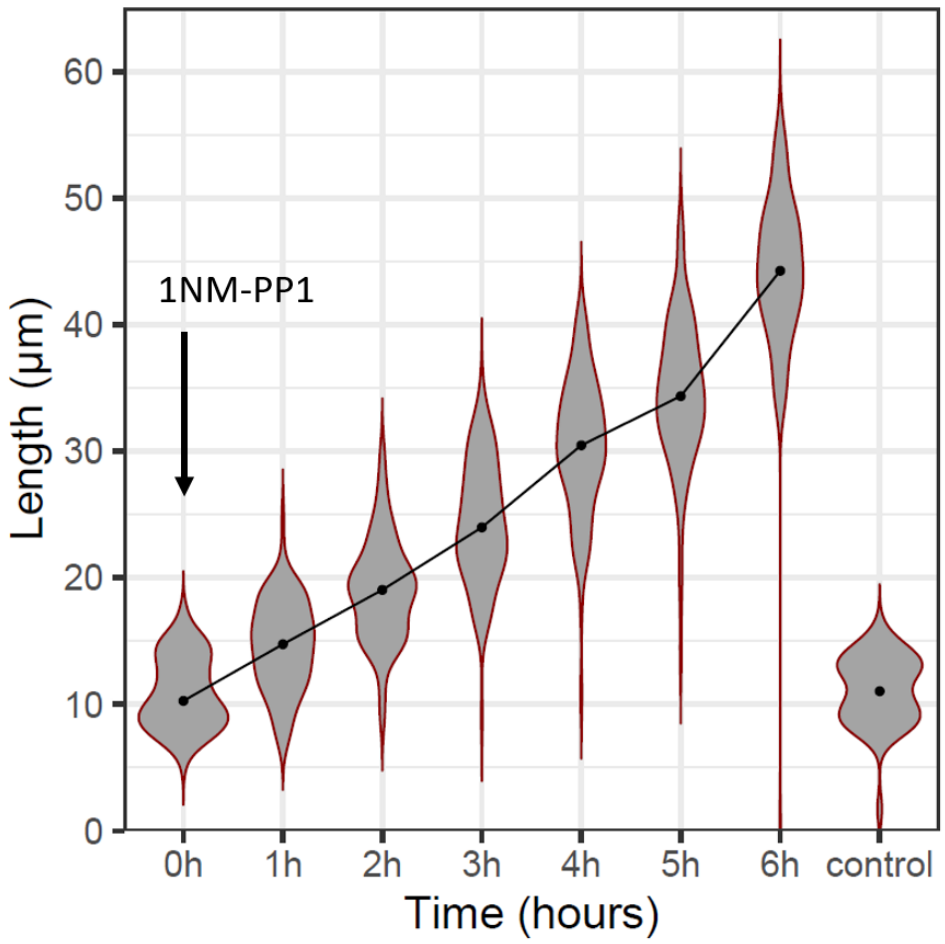
Finally, the +1 nucleosome, whose position depends on remodeling complex such as RSC or SWI/SNF, has been portrayed as an important regulatory player of the transcription (Kubik et al., 2018; Nagai et al., 2017; Rhee and Pugh, 2012; Martinez-Campa et al., 2004) (introduction). In the present study, occupancy of the +1 nucleosome is not directly associated with cell size since an increase in occupancy was also witnessed on the control condition. Moreover, the +1 nucleosome accessibility followed in both replicates the trend observed at the promoter. Therefore, since the +1 nucleosome sits in direct proximity of the promoter, its changes of occupancy and accessibility might be connected to the changes on the promoter while cell size

increases. The presence of proteins on the promoter would promote the downstream protection of the +1 nucleosome locus. On the one hand, the influence of the promoter structure on the +1 nucleosome structure could be a passive mechanism resulting from the limited access of the MNase to the sequence underlying the +1 nucleosome at a low MNase concentration. On the other hand, the proteins on the promoter could actively induced structural changes of the +1 nucleosome, involving its stabilization, incorporation of the histone variant H2A.Z for rapid transcription activation or else its accurate positioning downstream of the TSS (Farnung et al., 2021; Clapier et al., 2017; Venkatesh and Workman, 2015; Billon and Côté, 2012) to facilitate transcription in large cells. This would enhance the MNase digestion at both edges of the +1 nucleosome, thus increasing its occupancy and its accessibility. The last scenario is appealing as it would explain both the results from occupancy and accessibility.

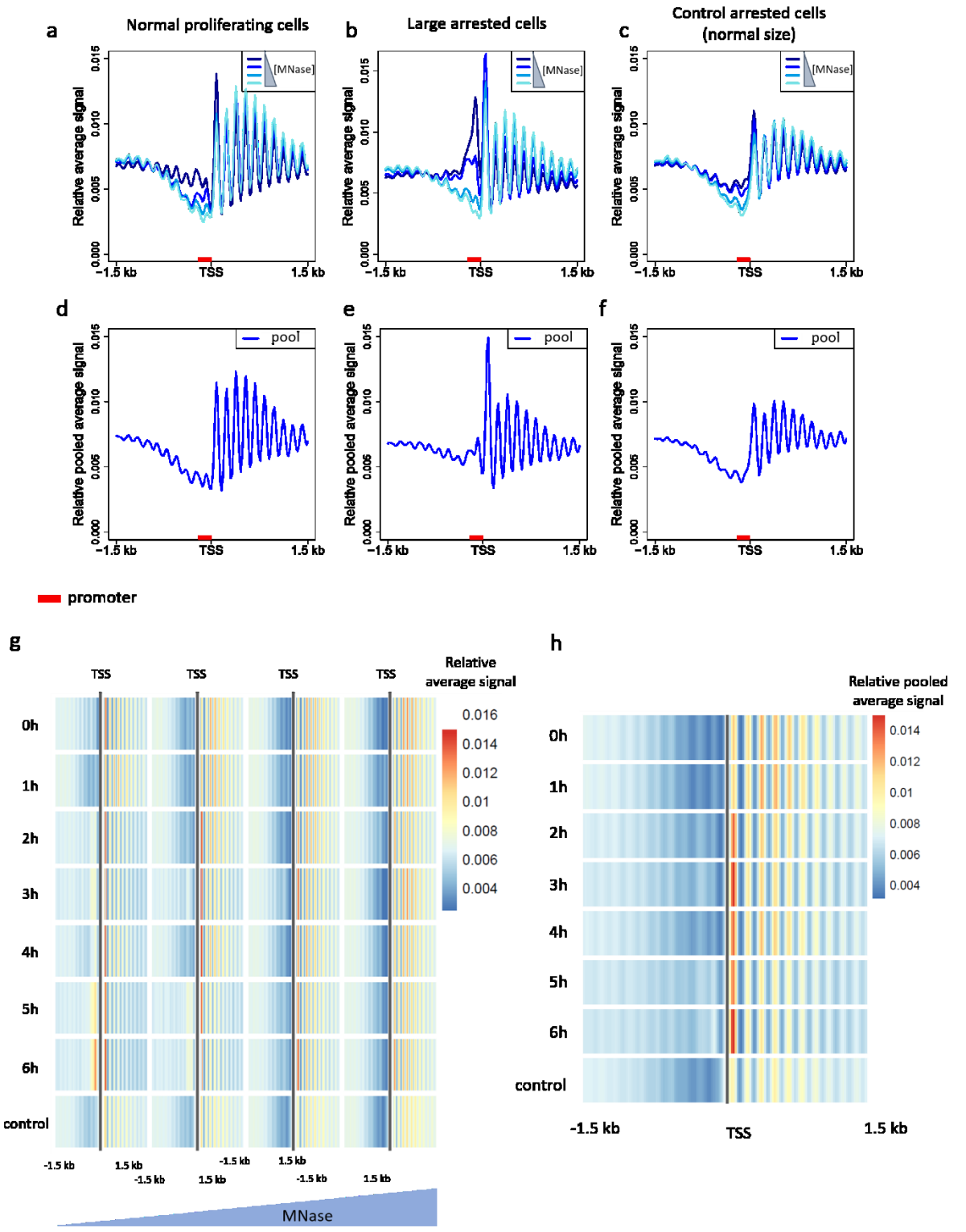
Strikingly, the nucleosome positioning on the gene body is not affected by cell size increase, although the transcription rate is higher in large cells. As reviewed in the introduction, cells possess various mechanisms to maintain the nucleosome organization on the gene body during transcription elongation (Evrin et al., 2022; Farnung et al., 2021; Venkatesh and Workman, 2015; Kulaeva and Studitsky, 2010; Workman, 2006). Such mechanisms could explain why no change in nucleosome positioning is observed by MNase-seq on the gene body of large cells.

Finally, an interesting aspect of these results is the increasing range of signal distribution around the TSS with cell size increase. This reflects the differential changes of chromatin structure depending on the genes. Therefore, it would now be interesting to illuminate such differential changes and to understand whether they are associated with specific functions or transcriptional states.

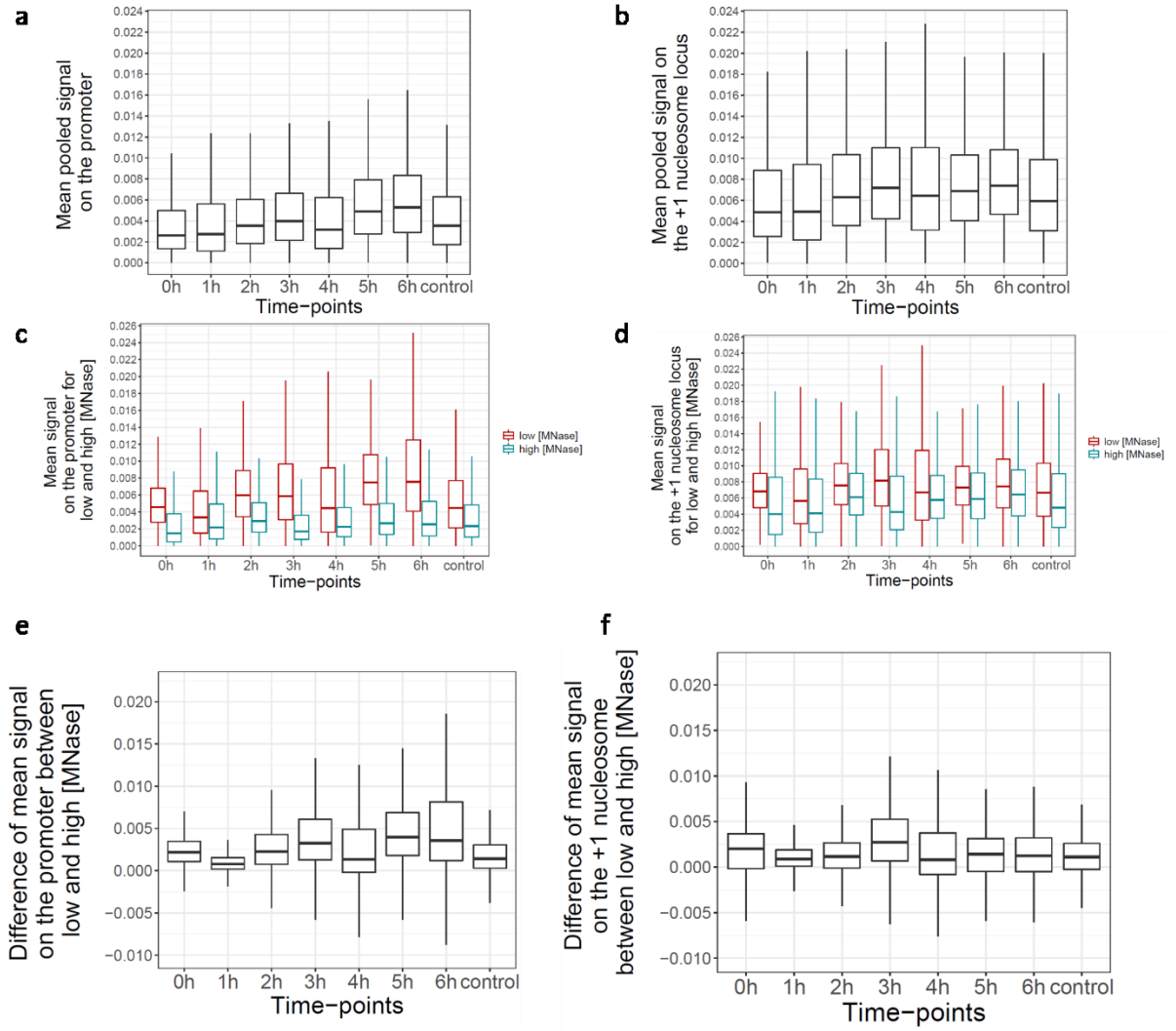
ANNEXES



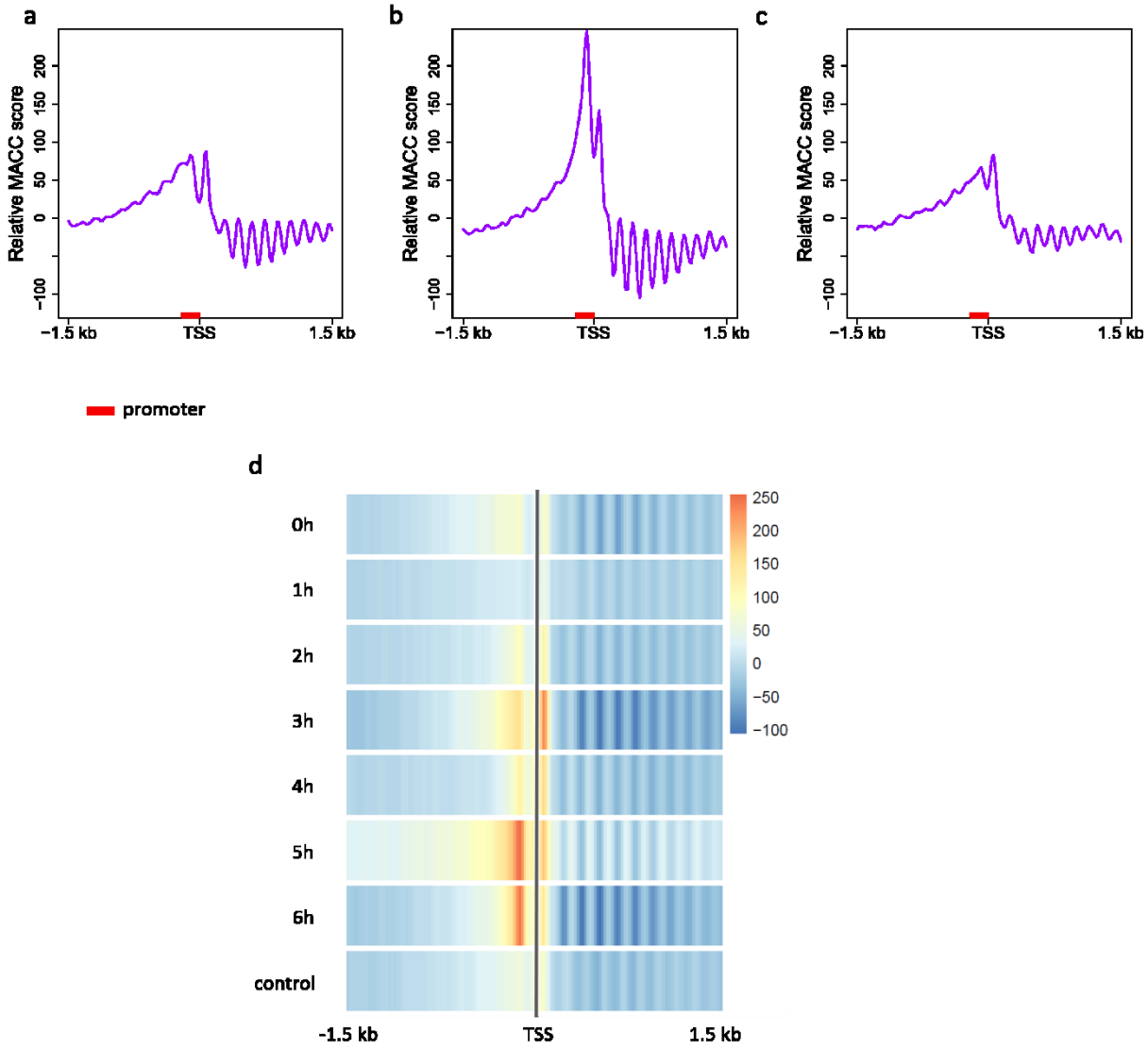
Annex 3.1: Cell size increase after addition of 1NM-PP1 (replicate 2).



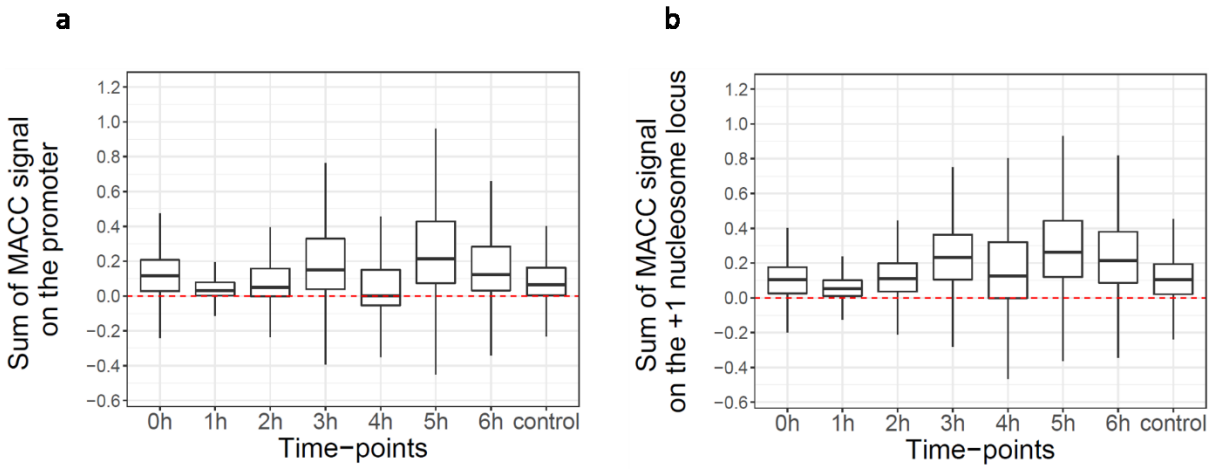
Annex 3.2: TSS-centred average chromatin profile of cells of growing size arrested with 1NM-PP1 for 6h (replicate 2). A-b*. Average profile of nucleosomes released at 4 different MNase concentrations for normal proliferating cells* ***(a)****, large arrested cells* ***(b)*** *and. normal arrested cells* ***(c)****. **D-f**. Average profile for pooled MNase concentrations for normal proliferating cells* ***(d)****, large arrested cells* ***(e)*** *and normal arrested cells* ***(f)****. **G-h.** Average profile for cells of growing size arrested with 1NM-PP1 and cells of normal arrested cells (control condition at the bottom of the heatmap) with four increasing MNase concentration (**g**), and with pooled MNase concentrations* ***(h).***



Annex 3.3: Mean signal on the promoter and +1 nucleosome for cells of growing size arrested with 1NM-PP1 for 6h (replicate 2). ***A.b.*** *Mean pooled occupancy on the promoter, 200 bp upstream of the TSS* ***(a)*** *and on the +1 nucleosome locus, 200 bp downstream of the TSS* ***(b)*** *(differences tested by paired t-test with n = 5887).* ***C-d.*** *Mean signal occupancy for the low and high MNase concentration on the promoter* ***(c)****, and on the +1 nucleosome locus (differences between time-points tested by paired t-test and between MNase concentrations by unpaired t-test, n = 5887)* ***(d)***. ***E-f****. Differences between low and high MNase signal on the promoter* ***(e)*** *and on the +1 nucleosome locus* ***(f)*** *(differences tested by paired t-test with n = 5887).*



Annex 3.4: TSS-centred average MACC profile of cells of growing size arrested with 1NM-PP1 for 6h (replicate 2). A-c*. MACC average profile for normal proliferating cells* ***(a)****, large cells arrested 6h with 1NM-PP1* ***(b)*** *and normal arrested cells* ***(c).***



Annex 3.5: Total MACC profile for cells of growing size arrested for 6h with 1NM-PP1 (replicate 2). *Total MACC signal on the promoter, 200 bp upstream of the TSS* ***(a)****, and on the +1 nucleosome locus, 200 bp downstream of the TSS* ***(b)*** *(paired t-test with n = 5887)*

Chapter 4: Genes and genomic regions showing differential chromatin organization as a function of cell size

In the previous chapter I analysed the global chromatin structure of the cdc2-asM17 up to 6h with 1NM-PP1. This chromatin structure, more precisely around the TSS, indeed changes with cell size. However, the previous chapter does not portray precisely the changes in chromatin structure for different sets of genes. In other words, how many different underlying patterns in chromatin structure lead to the average pattern reported in chapter 3? To answer this, I will explore here the chromatin structure of different groups of genes. The goal is to understand whether changes in chromatin organization with cell size are global or localized and if such changes correlate with the level of gene expression and/or specific functions of the gene. In a Wild Type situation, the chromatin structure indeed influences the transcription regulation of the underlying genes, as it was reported on figure 3.1c and by decades-long research on the matter (see introduction, (Murawska et al., 2020; Brahma and Henikoff, 2019; Kubik et al., 2018; Mieczkowski et al., 2016; Ganguli et al., 2014; Nabilsi et al., 2014; DeGennaro et al., 2013; Lantermann et al., 2009; Bryant et al., 2008; Workman et al., 1988; Lorch et al., 1987). Therefore, one could reasonably hypothesize that differentially expressed genes during size increase is associated with different chromatin structure

4.1. Detailed chromatin organization around the TSS for all genes for cells of different size

One of the most striking features from the global profile of chromatin architecture around the TSS is the large footprint witnessed on the promoter of large cells (**Chapter 3**). Is this footprint present in all the genes or is it a characteristic of a minority of genes that can affect the average profile? Since this footprint is present only at a low MNase concentration, I will mainly focus here on the profile obtained for this specific condition. Firstly, I would like to remind the reader that the signal presented here represents a distribution of the signal 1.5 kb on both sides of TSS so that each locus presents a fraction of the total signal of this region, exactly like in the previous chapter. Figure 4.1 and annex 4.1 detail the profile of chromatin structure for all the genes ordered by length, from the shortest at the top of the heatmap, to the longest at the bottom. Each heatmap corresponds to a distinct population of cells: normal-sized and proliferating cells (**Fig 4.1a and annex 4.1a**), large cells arrested 6h with 1NM-PP1 (**Fig 4.1b and annex 4.1b**) and normal-sized cells arrested 6h with 1NM-PP1 and cultivated without glucose (**Fig 4.1c and annex 4.1c**).

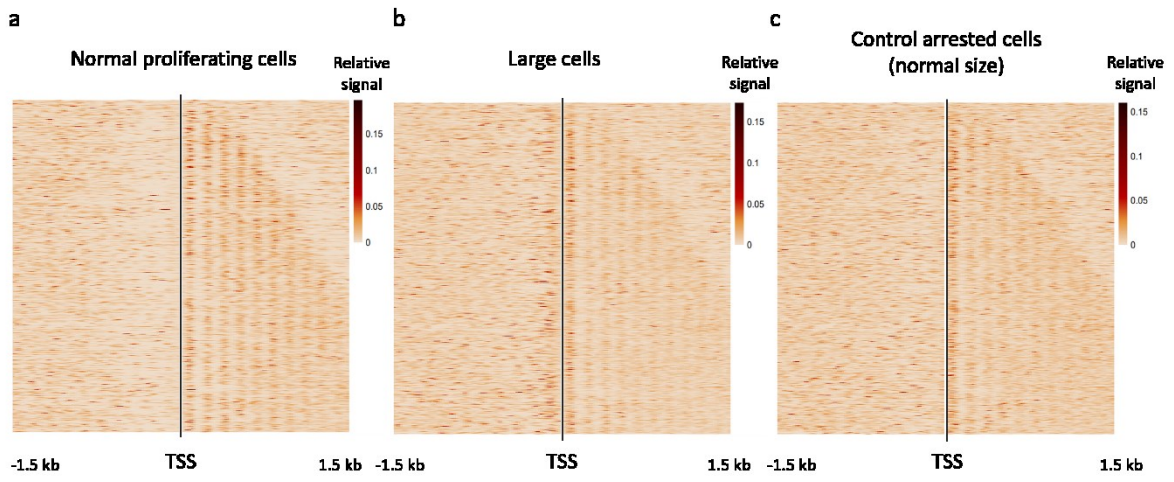


Figure 4.1: Detailed heatmap of the chromatin profile 1.5 kb around the TSS, obtained at low MNase concentration and for cells of different sizes*: normal proliferating cells* ***(a)****, large cells arrested with 1NM-PP1 for 6h* ***(b),*** *normal-sized cells arrested 6h with 1NM-PP1 and deprived of glucose for 6h* ***(c).***

For normal-sized cells, the heatmap recapitulates for all the genes what was previously seen on the average profile: an ordered array of nucleosomes downstream of the TSS with a relatively high signal for the +1 nucleosome, a lower signal on the promoter and a disorganized signal upstream of the promoter (Fig 4.1a and annex 4.1a). For large-sized cells, the situation is different: although the nucleosomes are well-organized on all the genes downstream of the TSS and quite fuzzy upstream of the TSS, one can witness a relatively higher signal on the promoter compared to the gene body and this seems to be the case for all the genes (Fig 4.1b and annex 4.1b). The signal from the +1 nucleosome also appears stronger at this specific localization, an aspect that was also reported in the previous chapter. The enrichment of signal on the promoter appears lower for the control condition, suggesting once again that the footprint on the promoter of large cells is size-specific (Fig 4.1c and annex 4.1c). Together, these heatmaps give us the ability to zoom in on the chromatin structure of all the genes around their TSS. One can propose that the footprint encountered specifically at the promoter of large cells and at a low MNase concentration is indeed present on most genes, rather than on a minority of genes.

4.2. Grouping genes under a common chromatin architecture

To confirm the idea that a higher footprint on the promoter of large cells is present on all genes, clustering analysis was performed using the K-means method. Briefly, the chromatin

profiles were partitioned into different groups, each of them containing close chromatin profiles. I set the numbers of groups to three as this created notably different three chromatin profiles. If I created Forming three distinct groups of genes led to three chromatin profiles looking very different from one another. Creating more than three groups led to two similar chromatin profiles, whereas creating only two groups masked one of the chromatin profile types. That is why I performed the clustering analysis by setting the number of groups to 3. For the two replicates of the MNase experiment (Fig 4.2 and annex 4.2) the clustering creates three groups of genes of comparable size and that possess similar chromatin organization. Drawing Venn diagrams permits us to visualize the overlap between clusters of each replicate (Fig 4.2h). This highlights the significant similarity of the gene cluster composition between the two replicates: cluster 1 presents 90% of common genes for the two replicates, 87% for cluster 2 and 83% for cluster 3 (Fisher exact test generates a p-value inferior to 0.05). These last observations support the fact that the two repeats are very reproducible and can thus be clustered similarly. Although the chromatin structure for the 1.5 kb upstream and downstream of the TSS is presented here, the three clusters distinguish differences of chromatin organization mainly around the TSS.

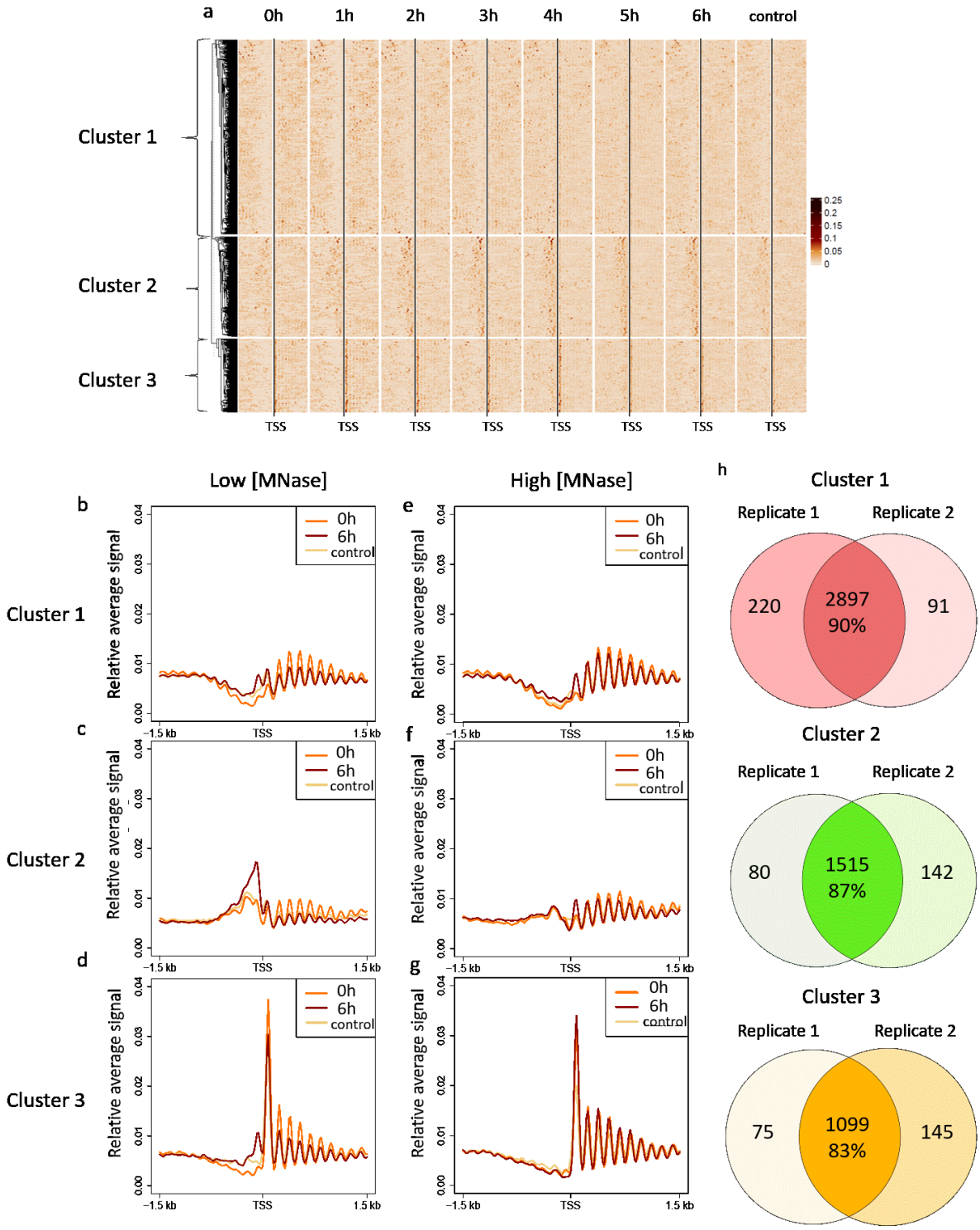


Figure 4.2: Three clusters of genes having specific chromatin organization profile 1.5 kb around the TSS obtained at low MNase concentration and for cells of different sizes. A. *Detailed heatmap for cell population each hour after addition of 1NM-PP1 until 6h and finally 6h with 1NM-PP1 and without glucose (control)* ***(a). B-g.*** *Average profile from a low MNase concentration of cluster 1* ***(b)****, cluster 2* ***(c)*** *and cluster 3* ***(d)*** *and from a high MNase concentration of cluster 1* ***(e)****, cluster 2* ***(f)*** *and cluster 3* ***(g)*** *for normal proliferating cells (0h), large cells (6h with 1NM-PP1) and control (6h with 1NM-PP1 and without glucose).* ***H****. Venn diagram showing the overlap between clusters of different replicates. The overlap, tested by Fisher exact test generates a p-value inferior to 0.05.*

In the first largest cluster (cluster 1) of replicate 1, an absence of signal is largely visible on the promoter for early timepoints (Fig 4.2a and 4.2b). For replicate 2, the promoter at 0h does not seem as much depleted as for replicate 1 (Annex 4.2a and 4.2b) as it exhibits a small peak as seen on the average profile from a low MNase concentration (Annex 4.2b), showing a slight difference between those two replicates. The occupancy of the proteins on the promoter in replicate 2 disappears at higher levels of MNase digestion (Annex 4.2e), showing that these proteins are sensitive to a higher MNase concentration and are thus unlikely to be canonical nucleosomes (Brahma and Henikoff, 2019; Chereji et al., 2017; Kubik et al., 2017).Moreover, in both replicates, after addition of 1NM-PP1, the promoter gets progressively occupied by proteins that are leaving a footprint at this locus (Fig 4.2a and annex 4.2a). This trend can be seen on the average profile for this cluster (Fig 4.2b and annex 4.2b) where the NDR at the promoter is visible at 0h and control but not at 6h which instead exhibits a narrow peak right upstream of the TSS, reaching around 0.0057 units of signal. However, this peak disappears at the highest MNase concentration (Fig 4.2e and annex 4.2e), showing again that the promoter of this subset of genes is occupied by MNase-sensitive proteins in large cells, unlikely to be canonical nucleosomes.

The intermediate cluster (cluster 2) contains genes for which the promoter is already occupied in normal-sized cells (Fig 4.2a, c and annex 4.2a, c). However, the promoter of these genes is not saturated upon normal conditions as the signal at this location keeps increasing. On the average profile (Fig 4.2c), the peak on the promoter for 0h, 6h and control cells appears wider in this cluster 2 than the peak in large cells on the promoter of cluster 1. At 6h, the signal reaches the highest value, around 0.0157, right upstream of the TSS and two to three shoulders of signal are visible going further upstream of the TSS (Fig 4.2c), as if more factors were added and were occupying a region upstream of the promoter when size increases. This specific pattern is not seen in the control condition that looks more like the one at 0h. At a high MNase concentration (Fig 4.2f and annex 4.2f), the signal for both normal-sized and large cells decreases at the promoter as before, showing that MNase-sensitive proteins are occupying the promoter in the two conditions and with higher efficiency in large cells. Nevertheless, a region up to 300 bp upstream of the TSS is occupied by MNase resistant factors. Since this pattern is observed for both 0h and 6h, the presence of those factors mentioned here is not

size-dependent. Finally, the last cluster (cluster 3) gathers genes that display a strongly positioned +1 nucleosome, which is present during the whole time-course and is thus not size-dependant (Fig 4.2a, d and annex 4.2a, d). Once again, the promoter seems progressively more occupied as the cells elongate. On the average profile, the promoter seems to be depleted of proteins at 0h, thus resembling the promoters from cluster 1 rather than cluster 2 (Fig 4.2a, d and annex 4.2a, d). At 6h, the peak on the promoter is narrower and slightly lower, around 0.01, than for genes in cluster 2 but a bit higher than for cluster 1. At a high MNase concentration, no footprint is left at 6h; this demonstrates sensitivity to high MNase digestion, unlike the most upstream factors from cluster 2.

Therefore, in the basal condition (at 0h) and for all time-points, the clustering analysis highlights separate groups of genes which present specific chromatin structure in the vicinity of the TSS. On the one hand, some pattern of chromatin organization, such as a strongly positioned +1 nucleosome (cluster 3) or the high occupancy of proteins on the promoters (cluster 2), is dependent on clusters but not on size. In addition, the different sensitivity of the proteins to the MNase suggests that the nature of the proteins bound upstream of the TSS is different depending on the clusters. It is also possible that the factors are identical but are forming complexes resisting an extensive MNase digestion. On the other hand, beyond these differences, the three clusters show the same trend as the cells are growing in size, namely an increasing occupancy of MNase-sensitive proteins at the promoter.

To verify the latter tendency in a quantitative way, the mean of the signal contained in the 200 bp upstream of the TSS is represented exactly like in the previous chapter (chapter 3), this time for the 3 different clusters. Since we are interested here in the MNase-sensitive proteins occupying the promoter, the mean of the signal on the 200 bp upstream of the TSS is calculated only for the lowest MNase concentration (Fig 4.3 and annex 4.3).

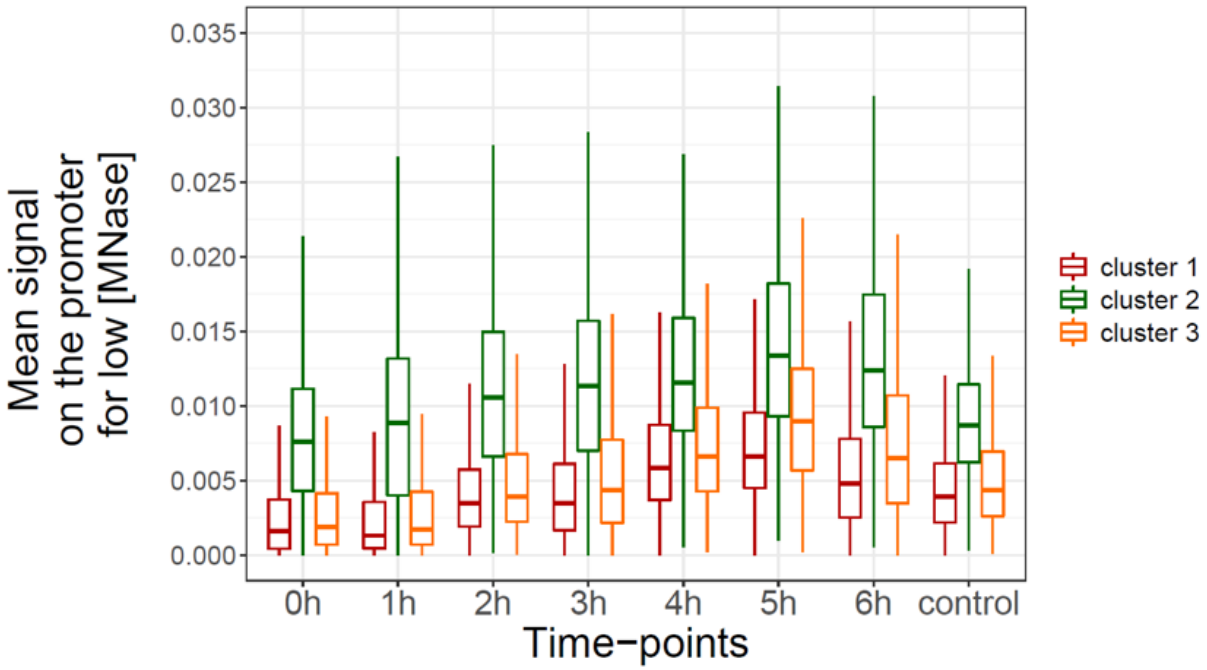


Figure 4.3: Mean signal on the promoter for low MNase concentration and for each time-point after addition of 1NM-PP1 and for the control condition.

From the start of the time-course, promoters from cluster 1 and 3 have a lower median occupancy (~0.002) than cluster 2 promoters (~0.008), as it could be anticipated from the previous heatmaps and average profiles (Fig 4.2b, c, d, and annex 4.2 b, c, and d). There is no meaningful comparison between cluster 1 and cluster 3 in replicate 2; conversely, when these clusters are compared with cluster 2, a p-value inferior to 2e-16 is generated by unpaired t-test (Fig 4.3, annex 4.3 and Table 4.1).

Table 4.1: Comparisons of occupancies tested by unpaired t-test between clusters and paired t-test between time-points for both replicates.

			p-value	
Time-points	Group 1	Group 2	Rep 1	Rep 2
0h	cluster 1	cluster 2	< 2e-16	<2e-16
	cluster 1	cluster 3	2.30E-08	0.65
	cluster 2	cluster 3	< 2e-16	<2e-16
1h	cluster 1	cluster 2	< 2e-16	<2e-16
	cluster 1	cluster 3	1.20E-10	3.00E-06
	cluster 2	cluster 3	< 2e-16	<2e-16
2h	cluster 1	cluster 2	< 2e-16	<2e-16
	cluster 1	cluster 3	< 2e-16	2.20E-08
	cluster 2	cluster 3	< 2e-16	<2e-16
3h	cluster 1	cluster 2	< 2e-16	<2e-16
	cluster 1	cluster 3	< 2e-16	9.40E-10
	cluster 2	cluster 3	< 2e-16	<2e-16
4h	cluster 1	cluster 2	< 2e-16	<2e-16
	cluster 1	cluster 3	< 2e-16	3.40E-14
	cluster 2	cluster 3	< 2e-16	<2e-16
5h	cluster 1	cluster 2	< 2e-16	<2e-16
	cluster 1	cluster 3	< 2e-16	<2e-16
	cluster 2	cluster 3	< 2e-16	<2e-16
6h	cluster 1	cluster 2	< 2e-16	<2e-16
	cluster 1	cluster 3	< 2e-16	<2e-16
	cluster 2	cluster 3	< 2e-16	<2e-16
control	cluster 1	cluster 2	< 2e-16	<2e-16
	cluster 1	cluster 3	< 2e-16	5.00E-10
	cluster 2	cluster 3	< 2e-16	<2e-16

			p-value	
Clusters	Group 1	Group 2	Rep 1	Rep 2
Cluster 1	0h	1h	0.812	< 2e-16
	0h	2h	< 2e-16	< 2e-16
	0h	3h	< 2e-16	< 2e-16
	0h	4h	< 2e-16	< 2e-16
	0h	5h	< 2e-16	< 2e-16
	0h	6h	< 2e-16	< 2e-16
	0h	control	< 2e-16	< 2e-16
Cluster 2	0h	1h	< 2e-16	< 2e-16
	0h	2h	< 2e-16	< 2e-16
	0h	3h	< 2e-16	< 2e-16
	0h	4h	< 2e-16	< 2e-16
	0h	5h	< 2e-16	< 2e-16
	0h	6h	< 2e-16	< 2e-16
	0h	control	< 2e-16	< 2e-16
Cluster 3	0h	1h	9.0e-05	< 2e-16
	0h	2h	< 2e-16	< 2e-16
	0h	3h	< 2e-16	< 2e-16
	0h	4h	< 2e-16	1.2e-11
	0h	5h	< 2e-16	< 2e-16
	0h	6h	< 2e-16	< 2e-16
	0h	control	< 2e-16	1.1e-06

This highlights the specificity of cluster 2 promoters to efficiently recruit proteins in normal conditions which are, for the most upstream ones, resistant to a high MNase treatment. The progressive increase in signal as cell size increase is noticeably clear for all the clusters for replicate 1 (Fig 4.3) but not as much for replicate 2 (Annex 4.3) in which the signal on the promoters decreases at 4h for clusters 1 and 3. However, a progressive increase is striking for cluster 2 for the two replicates (Fig 4.3 and annex 4.3). In fact, for cluster 2, the median signal is always higher compared to the previous time-point, except for the control in which it is close to the one at early time-points, around 0.09. For this cluster (cluster 2), comparisons of all time-points to 0h always show significant differences and this for both replicates (Table 4.1). Finally, cluster 3 follows the same trend as cluster 1 for both replicates (Fig 4.3 and annex 4.3). We can conclude that each cluster presents an increase in MNase-sensitive protein on the promoter at large cells. These proteins in cluster 2 appears particularly connected to cell size because they are gradually increasing with cell size and because the control condition shows the same trend as in normal-sized cells.

These analyses show a global coordination of cell size and chromatin structure on the promoter, rather than a specialized coordination (only on certain genes). In other words, although the clusters present different chromatin architecture close to the TSS in a Wild Type situation, most promoters are more and more occupied by MNase-sensitive proteins, concomitantly with cell size increase. This change in global occupancy is dependent on the initial chromatin state of the genes (e.g., it reaches higher value if the occupancy was already high in basal condition). This is consistent with the idea that most proteins are scaling with cell size thus triggering the increase in their recruitment to their target DNA site at all genes and/or at all genes. In summary, this work describes in detail how the primary chromatin structure can change in the context of cell size.

For the following analyses on the three clusters, the overlap of genes between the two replicates (Fig 4.2h) are considered as cluster 1, 2, and 3 (red, green, and orange respectively).

4.3. Function and expression levels of the three groups of genes having a distinct chromatin structure

Can we find any enrichment of GO-terms in these cluster lists? Are these genes gathered depending on their function in the cell? The AnGeLi tool from the Bähler's lab website is a great resource to answer the question as it was developed specifically for *S. pombe* and systematically performs enrichment analysis through a variety of available datasets (Bitton et al., 2015). I used the default parameters and input each cluster lists for which I decided to consider all genes and not only protein-coding genes. The background for this analysis is all the genes from *S. pombe*. Cluster 1 (2891 genes) and 2 (1509 genes) are highly depleted in translational processes (e.g., p-value = 4.56e-30 for cluster 1 and p-value = 3.71e-13 for cluster 2), macromolecular biosynthetic processes (p-value = 1.11e-05) and metabolic processes (p-value = 0.00012) and are enriched in ion and anion transport (p-value = 0.00058), cell communication and organic acid transport.

Interestingly, cluster 2 contains most core histones, namely hta1 and htb1 (coding for H2A alpha and H2B respectively), hhf1 and hht1 (coding for H3.1 and H4.1 respectively), hht2 and hh2 (coding for H3.2 and H4.2 respectively), and finally hhf3 (coding for H4.3). This set of genes

are not randomly distributed in the genome but, instead, are forming adjacent pairs that are transcribed in opposite direction from the same upstream region (**Fig 4.4**).

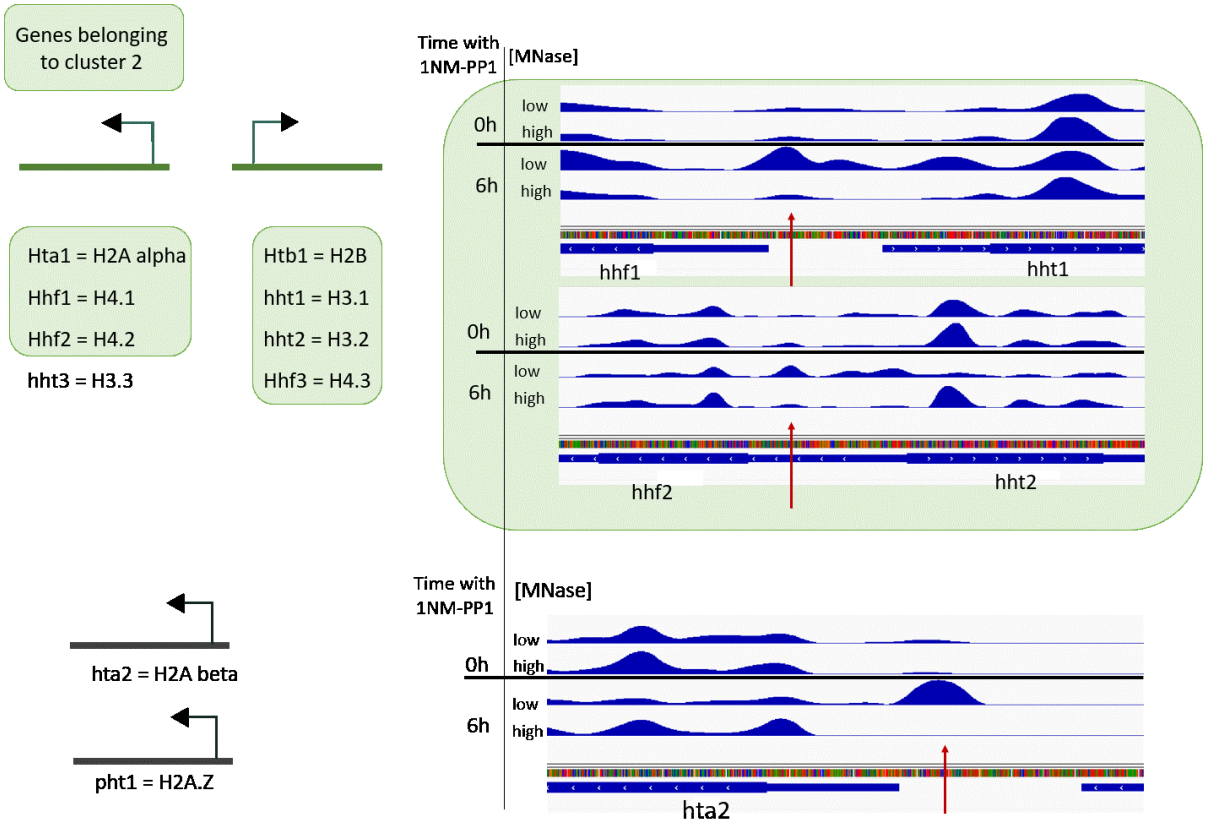


***Figure 4.4: MNase-resisting particles found upstream of histones belonging to cluster 2**. Histones found in cluster 2 are forming adjacent pairs that are transcribed in opposite direction from the same upstream region.*

This is not the case for the other histone genes, pht1 (coding for the variant H2A.Z) and hta2 (coding for H2A beta) which are excluded from this list. Besides, a closer look at histones contained in cluster 2, especially H3 and H4, via a genome browser can confirm the presence of MNase-resisting particle upstream of the promoter in both normal and large cells. Those particles are not detected for hta2 (**Fig 4.4**). One possibility, which is quite likely for hhf2 and hht2 due to their proximity in the genome, is that their upstream MNase-resisting particle corresponds to the +1 nucleosome of their adjacent pairs.

Cluster 3 (1098 genes) is highly enriched in macromolecule metabolic processes (p-value = 3.91e-25), ribosome biogenesis and assembly (p-value = 1.24e-08), translation and

biosynthetic process (p-value = 3.21e-10) (analysis performed on the AnGeLi website). Interestingly, those genes represent one fourth of the R-sectors containing genes which are also enriched in translation, ribosome assembly and biogenesis, biosynthetic processes, etc., and which positively correlate with growth rate (Kleijn et al., 2021; Scott and Hwa, 2011; Scott et al., 2010)(Kleijn et al. 2021; Scott et al. 2010). Therefore, genes having a role in translation and ribosome processes possess a specific pattern of chromatin around their TSS, such as a strongly positioned +1 nucleosome and a depleted promoter even at a low MNase (**Fig 4.2a, d and annex 4.2a, d**). This agrees with previous findings that revealed specific chromatin architecture for genes coding for ribosomal proteins and growth genes (Rawal et al., 2018; Nocetti and Whitehouse, 2016; Kubik et al., 2015; Knight et al., 2014; Mavrich et al., 2008b). In budding yeast, these genes possess a "fragile" nucleosome at their promoter destabilized by GRFs and RSC (Brahma and Henikoff, 2019; Kubik et al., 2015), which are simultaneously establishing a strongly positioned +1 nucleosome behind the TSS (Mivelaz et al., 2020; Kubik et al., 2018; Mavrich et al., 2008b). Such specific chromatin structure around of the TSS is indeed believed to provide the proper transcription regulation for these genes which are often highly expressed in basal conditions, and which respond quickly to changes in the environment (Mivelaz et al., 2020; Metzl-Raz et al., 2017; Kubik et al., 2015).

Nevertheless, when cell size increases, this set of genes, as well as the two others (cluster 1, 2 and 3), all present a higher footprint on their promoters demonstrating the increasing presence of a, or several, factor(s) at this localization.

As stated in the introduction, chromatin structure is intimately linked to transcription. Amalia Martínez Segura, a former PhD student from our group, performed transcriptomics on the *cdc2-asM17* strain for 11h after addition of 1NM-PP1 (Martínez Segura 2017). This dataset is of great interest to test a potential relationship between RNA levels of the elongating *cdc2-asM17* strain in G2 and its chromatin structure. Hence, I re-analyzed the RNA-seq data, starting from the data which Amalia Martínez Segura previously normalized. Firstly, thanks to this RNA-seq experiment and analysis, Amalia Martínez Segura identified genes that are drifting from the major scaling behavior. In addition to the scaling genes, she highlighted two other sets of genes: towards the end of the time-course, the positively non-scaling genes are expressed above the level required for RNA scaling with size whereas the negatively non-scaling genes

are expressed below the scaling level (**Fig 4.5a**). It is noteworthy that the negatively non-scaling genes might still be more expressed at late time-points compared to the basal level at normal size. This is because the reference for non-differentially expressed genes is the scaling genes which expression are supposed to go up with size. On the one hand, the positively scaling genes are enriched for the genes acting in TCA (tricarboxylic acid) cycle, the proteins and carbohydrates involved in the catabolism (Martínez Segura 2017). On the other hand, this group of genes is depleted of genes involved in translation and ribosome biogenesis. The negatively non-scaling genes contain ribosomal genes like rps2202 or rpl22 and few genes encoding tRNA for guiding methionine and serine, genes coding for the four core histones H2A, H2B, H3 and H3, and chromatin remodelers (Martínez Segura 2017). Other scientists also aimed to characterize this kind of genes in budding yeast which they call super-scaling genes for positively non-scaling genes and sub-scaling genes for negatively non-scaling genes (Lanz et al., 2021). Interestingly, they found that lysosome constituents, which are involved in the degradation of proteins, are found in the super-scaling category; this finding is consistent with the enrichment for catabolism found in the positively non-scaling genes in *S. pombe*. Since ribosomes are limiting for growth (Kleijn et al., 2021; Scott and Hwa, 2011), it is possible that large cells need to degrade more proteins to recycle them into the ribosome biogenesis and assembly pathway. Finally, chromatin components such as histones are also found in the sub-scaling category, showing a parallel between both yeast species for histone regulation in the context of cell size. Indeed, histones are now a well-known category of non-scaling proteins since they are coordinated with the genome content (Claude et al., 2021).

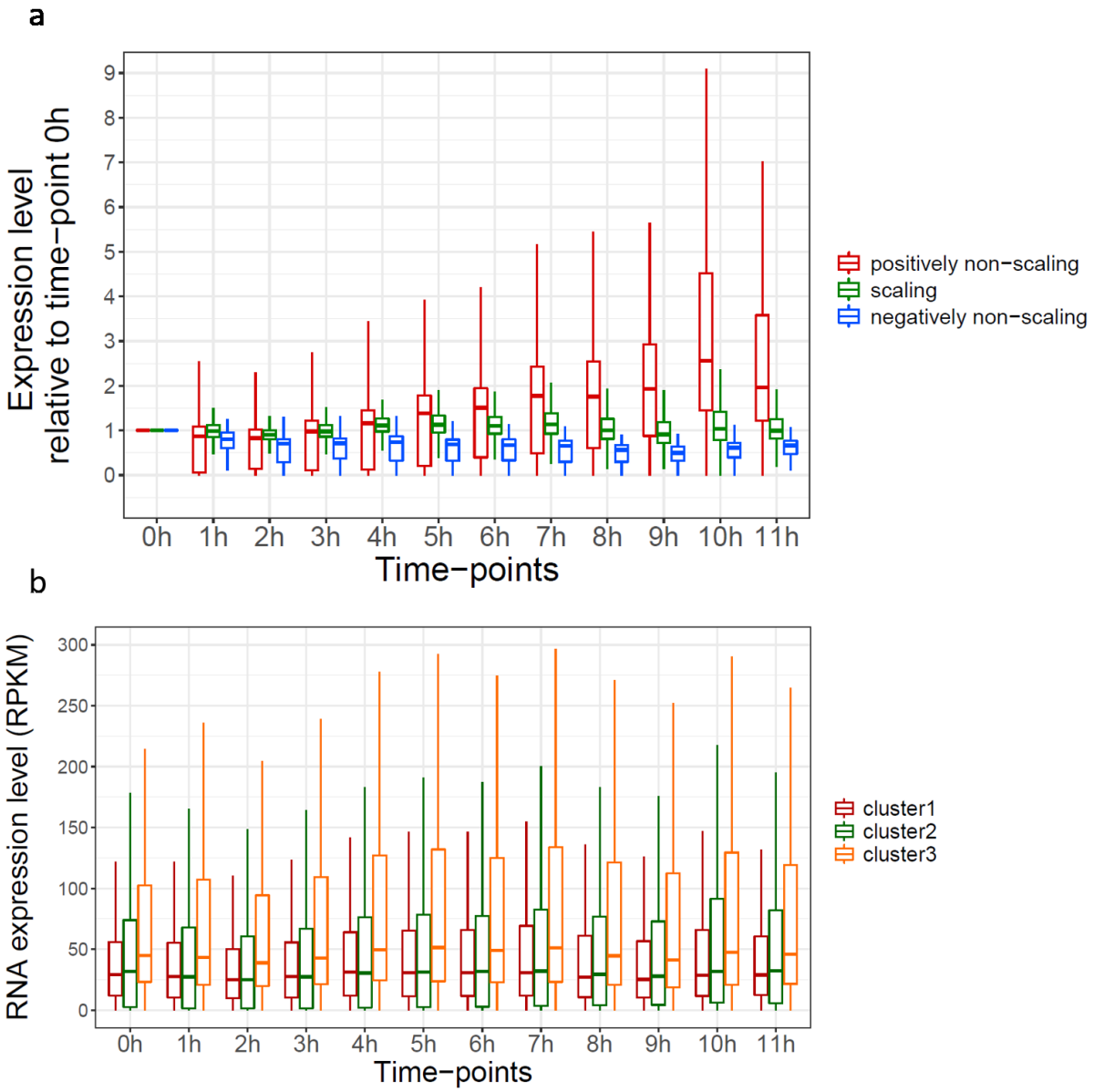


Figure 4.5: Expression levels of groups of genes during cell size increase after addition of 1NM-PP1*. Expression levels normalized by time-point 0h of differentially expressed genes identified by RNA-seq by Amalia Segura Martinez* ***(a).*** *Expression levels in RPKM of the three clusters corresponding to the overlap between the 2 replicates (Fig 4.2h)* ***(b).***

Are some clusters identified from the MNase-seq data having specific expression levels? To answer this, the distribution of RNA expression level for each cluster is represented using the previous RNA-seq data on the *cdc2-asM17* strain cultivated 11h with 1NM-PP1 (Fig 4.5b). It is noteworthy that the level of expression is not normalized to the time-point 0h, so that the RNA levels of each cluster can also be analyzed in normal proliferating conditions (Fig 4.5b). Indeed, the expression of each category of genes is different at time-point 0h. Medians of expression for cluster 1 and 2 are close (slightly more than 20 rpkm) but the variability of signal is larger

for cluster 2. Cluster 3 shows a median close to 50 rpkm which is higher compared to cluster 1 and 2, consistently with the high expression of these genes (Kubik et al., 2015; Marguerat et al., 2012). This trend remains throughout the time-course. Overall, the RNA-seq indicates that the grouping into the three clusters is not strongly linked to the differential expression of these genes with cell size increase. Instead, the grouping of the genes having similar chromatin structure is determined by the pattern of chromatin structure in normal-sized cells (Fig 4.2 and annex 4.2) However, these genes have a different expression pattern in normal conditions suggesting that the specific chromatin organization of each cluster might play a role in the regulation of transcription in basal conditions.

To compare the MNase-seq data with the RNA-seq data as a function of cell size, another strategy would be to analyze the occupancy on the promoter for the scaling and non-scaling genes. For this analysis, the distribution of signal on the 200 bp upstream of the TSS is represented for each category (Fig 4.6a and annex 4.4).

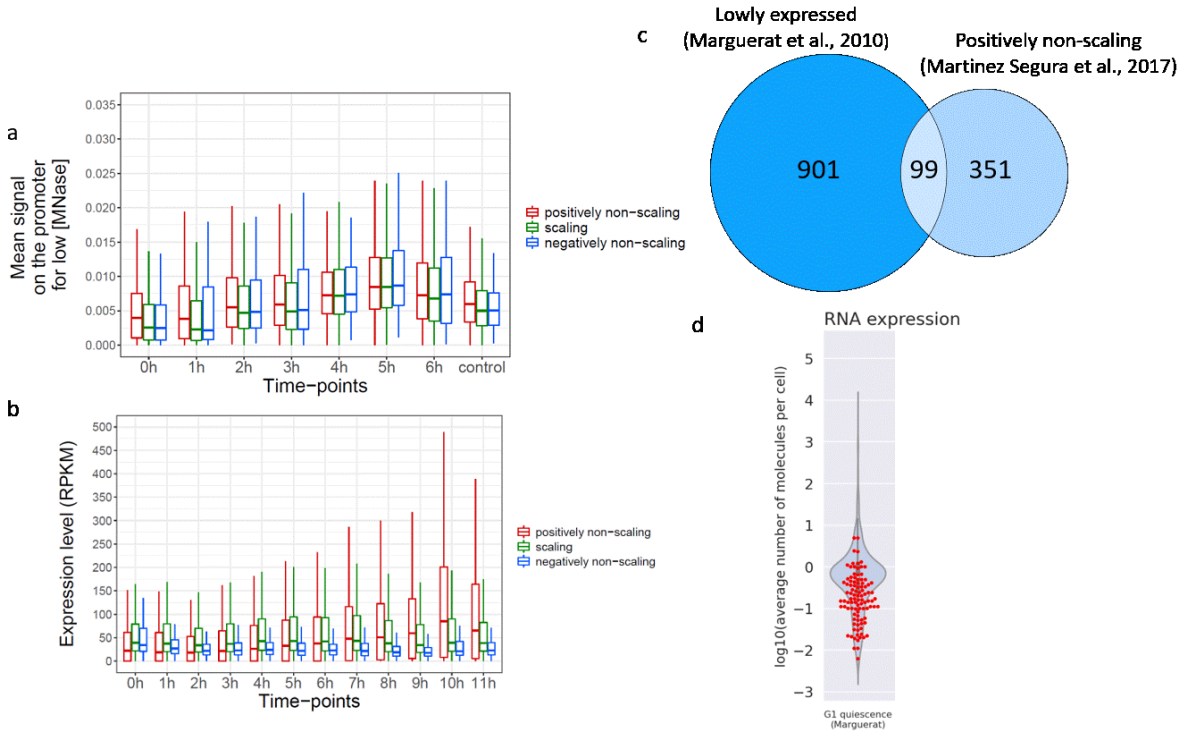


***Figure 4.6: A-b**. Mean signal on the promoters at low MNase concentration **(a)** and expression levels of differentially expressed genes **(b)** while cell size increases without normalising by time-point 0h. **C.** Venn diagram showing overlap between bottom less expressed genes and positively non-scaling genes (Fisher exact test generate a p-value inferior to 0.05). **D.** Expression of the overlapping genes from **(c)**. in context of the other genes for vegetative growth reported by pombase (data from Marguerat et al., 2012). **E.** Mean signal on the promoters at high MNase concentration while cell size increases.*

At a low MNase concentration, the median occupancy on the promoter is higher for positively non-scaling genes than for the other genes at early time-points (~0.004 vs ~0.0025 at 0h for replicate 1 and ~0.0045 vs ~0.0035 at 1h for replicate 2) and control (0.006 vs 0.005 for replicate 1 and 0.0055 vs 0.0045 for replicate 2). However, only the difference between positively non-scaling genes and scaling genes is significant (Table 4.2). This difference is less in both replicates at later time-points but not in the control condition (~0.0085 for all genes at 5h in replicate 1 and ~0.0075 at 5h in replicate 2) (Fig 4.6a and annex 4.4). At later time-points, no significant changes can be seen between the groups of genes (Table 4.2). The significant difference between positively non-scaling genes and scaling genes can be seen once more for the control condition and in both replicates (Table 4.2).

Table 4.2: Comparison tested by unpaired t-test between occupancies of scaling and non-scaling genes at a low MNase concentration.

Low [MNase]			p-value			
Time-points	Group 1	Group 2	Rep 1		Rep 2	
0h	positively non-scaling	negatively non-scaling	0.093	ns	0.0743	ns
	positively non-scaling	scaling	5.90E-05	****	0.0023	**
	negatively non-scaling	scaling	0.525	ns	0.8669	ns
1h	positively non-scaling	negatively non-scaling	0.19	ns	0.06	ns
	positively non-scaling	scaling	8.70E-05	****	7.40E-05	****
	negatively non-scaling	scaling	0.4	ns	0.62	ns
2h	positively non-scaling	negatively non-scaling	0.69	ns	0.61	ns
	positively non-scaling	scaling	0.026	*	3.20E-01	ns
	negatively non-scaling	scaling	0.413	ns	0.99	ns
3h	positively non-scaling	negatively non-scaling	0.921	ns	0.45	ns
	positively non-scaling	scaling	0.024	*	1.20E-01	ns
	negatively non-scaling	scaling	0.258	ns	0.92	ns
4h	positively non-scaling	negatively non-scaling	0.8	ns	0.369	ns
	positively non-scaling	scaling	0.9	ns	1.30E-02	*
	negatively non-scaling	scaling	0.83	ns	0.659	ns
5h	positively non-scaling	negatively non-scaling	0.34	ns	0.95	ns
	positively non-scaling	scaling	0.8	ns	0.53	ns
	negatively non-scaling	scaling	0.36	ns	0.67	ns
6h	positively non-scaling	negatively non-scaling	0.994	ns	0.93	ns
	positively non-scaling	scaling	0.077	ns	0.2	ns
	negatively non-scaling	scaling	0.315	ns	0.43	ns
control	positively non-scaling	negatively non-scaling	0.07796	ns	0.05945	ns
	positively non-scaling	scaling	0.00033	***	1.30E-04	***
	negatively non-scaling	scaling	0.77867	ns	0.8114	ns

Since a different promoter state is visible in normal conditions for positively non-scaling genes, it raises the question whether this group of genes presents distinct levels of expression in normal conditions. To understand this, Figure 4.6b presents RNA levels of each gene category are not normalized by the time-point 0h (Fig 4.6b). Strikingly, one can notice that the levels of positively non-scaling genes are lower at 0h compared to the other genes (20 median rpkm vs

40 for the other genes). At 3h, the median expression of positively and negatively non-scaling genes is similar (~20 rpkm). Finally, from 7h onwards, the positively non-scaling genes start overtaking the other genes in terms of their expression level (Fig 4.6b). This angle of analysis of RNA levels according to cell size without normalizing by time-point 0h offers a new vision of the positively non-scaling genes. Firstly, the expression of the positively non-scaling genes only gets higher at 7h whereas the time-course for the MNase-seq stops at 6h. Therefore, we cannot rule out that changes in chromatin structure for positively non-scaling genes occur after 7h. Secondly, positively non-scaling genes are composed of lowly expressed genes at 0h. Indeed, one fourth of the list of positively non-scaling genes (99 among 451) belongs to the 1000 lowest expressed genes identified by Marguerat and colleagues (Marguerat et al. 2012) which is more than expected by chance (Fig 4.6c). The RNA level of these genes is directly available on Pombase as violin plots showing their lower RNA levels compared to the median expression of all the genes in vegetative growth (Fig 4.6d). The higher signal seen on the promoter of the positively non-scaling genes at a low MNase concentration is thus associated with lower RNA levels. Hence, occupancy on the promoter can show a striking dichotomy: in normal conditions, a high occupancy on the promoter can be associated with low RNA levels; in large cells, a high occupancy on the promoter is systematic (occurs for every gene) and is associated with scaling level (since most of the genes are scaling). So, are the factors occupying the promoter at different sizes the same? Although the MNase-seq is a blind approach and cannot distinguish between the proteins protecting the DNA, the sensitivity of these proteins to MNase can at least be tested, giving an insight onto which kind of interactions are observed on the promoter as a function of cell size.

To understand better the nature of the factors occupying the promoter at different sizes, the size-dependent variation of occupancy on the promoter at high MNase concentration is now considered (Fig 4.6e and annex 4.4). For the three groups of genes, the promoter does not show an increase in occupancy with cell size. However, the signal for positively non-scaling genes is always significantly higher at all time-points compared to the scaling genes (Table 4.3), which was the case at early time-points when using a low MNase, but no longer at late time-points. This suggests that promoters of positively non-scaling genes are bound to factors that are more resistant to MNase than factors protecting the scaling promoters. This trend does not depend on cell size. Therefore, we can assume that these resistant factors are different

from the sensitive factors seen at a low MNase. These resistant factors might induce a lower expression level of these genes in basal conditions.

Table 4.3: Comparison tested by unpaired t-test between occupancies of scaling and non-scaling genes at a high MNase concentration.

High [MNase]			p-value			
Time-points	Group 1	Group 2	Rep 1		Rep 2	
0h	positively non-scaling	negatively non-scaling	0.011	*	0.014	*
	positively non-scaling	scaling	8.40E-05	****	1.40E-05	****
	negatively non-scaling	scaling	0.911	ns	0.782	ns
1h	positively non-scaling	negatively non-scaling	0.06594	ns	0.02	*
	positively non-scaling	scaling	9.80E-04	***	5.50E-05	****
	negatively non-scaling	scaling	0.84284	ns	0.85	ns
2h	positively non-scaling	negatively non-scaling	0.03818	*	0.03647	*
	positively non-scaling	scaling	0.00011	***	5.10E-04	***
	negatively non-scaling	scaling	0.90031	ns	0.90867	ns
3h	positively non-scaling	negatively non-scaling	0.19	ns	0.0814	ns
	positively non-scaling	scaling	0.045	*	4.00E-04	***
	negatively non-scaling	scaling	0.756	ns	0.7387	ns
4h	positively non-scaling	negatively non-scaling	0.1812	ns	0.08636	ns
	positively non-scaling	scaling	0.0041	**	5.10E-04	***
	negatively non-scaling	scaling	0.8955	ns	0.88732	ns
5h	positively non-scaling	negatively non-scaling	0.12074	ns	0.02383	*
	positively non-scaling	scaling	0.00067	***	0.00011	***
	negatively non-scaling	scaling	0.83677	ns	0.8778	ns
6h	positively non-scaling	negatively non-scaling	0.04573	*	0.08459	ns
	positively non-scaling	scaling	0.00017	***	0.00031	***
	negatively non-scaling	scaling	0.94201	ns	0.87172	ns
control	positively non-scaling	negatively non-scaling	0.0091	**	0.015	*
	positively non-scaling	scaling	5.00E-06	****	8.30E-06	****
	negatively non-scaling	scaling	0.9753	ns	0.728	ns

4.4. Other groups of genes show distinct chromatin patterns as a function of cell size

In the previous chapter, one way to analyze the chromatin organization on the promoter was to subtract the signal on the 200 bp upstream of the TSS from a low MNase treatment to the signal on the same locus after a high MNase treatment (chapter 3, Fig 3.7e and annex 3.1e). To remind the reader, this difference reflects the occupancy of MNase-sensitive proteins minus the occupancy of MNase-resistant proteins on the promoters. This metric can thus indicate whether the ratio of these two types of proteins on the promoter changes as the size of the cells increases. This analysis in the previous chapter revealed a higher difference in large cells and even more remarkably, a progressive widening of the signal distribution as the cells elongates (chapter 3, Fig 3.7e). Moreover, a similar trend was observed for occupancy in

MNase-sensitive factors (chapter 3, Fig3.7c). These results demonstrate that the range of increase in MNase-sensitive factors is different depending on the genes, although all genes appear to undergo a higher recruitment efficiency in large cells.

Therefore, genes having a difference of signal higher than 0.0105 at 6h after adding 1NM-PP1 are separated from the rest of the genes. 0.0105 is an empirical threshold that is higher than the highest interquartile value when considering all the genes and both replicates (chapter 3, Fig 3.7e and annex 3.3e). When filtering out the genes corresponding to this setting (that will be colored brown in the next figures), a list of 836 genes is revealed. They are then plotted as a function of time after 1NM-PP1 addition and compared to the rest of the genes (Fig 4.7a and annex 4.5a).

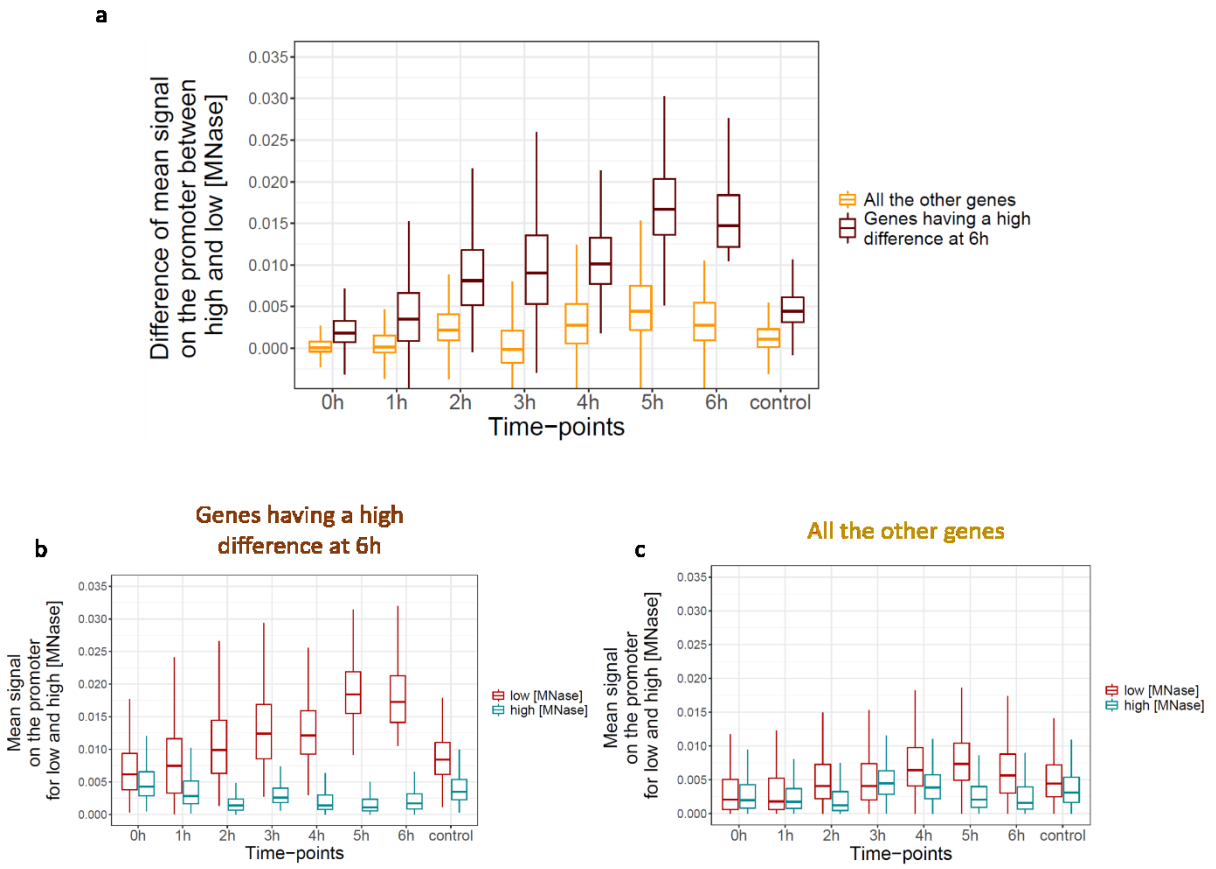


Figure 4.7: Two groups of genes showing differences in the promoter occupancy. A. *Difference of occupancy between a low and a high MNase treatment on the promoter as a function of time after 1NM-PP1 (paired t-test to test differences between the timepoints for both groups and unpaired t-test to test the difference between the groups of genes at all the time-points).* ***B-c.*** *Occupancy on the promoter after low and high MNase treatment for the genes having a high difference at 6h* ***(b)*** *and all the other genes* ***(c)*** *(paired t-test to test differences between the timepoints for both groups and unpaired t-test to test the difference between the groups of genes at all the time-points).*

In replicate 1 (Fig 4.7a and Table 4.4), we can observe that the two groups behave very differently. At 0h, genes having a high difference at 6h already have a higher difference between MNase-sensitive and MNase-resistant factors (~0.002 vs 0 for the rest of the genes) when compared to the other genes. This trend is conserved in the control condition (Fig 4.7a). Then, when the cells elongate, the difference exacerbates between the two groups of genes (0.015 of median occupancy vs 0.0025 for the rest of the genes). In the 2nd replicate (Annex 4.5a and Table 4.4), the median occupancy for the genes having a high difference at 6h is slightly higher than the rest of the genes at 0h. Additionally, the same trend as for replicate 1 can be observed, namely an obvious size-dependent increase in the difference between MNase-sensitive and MNase-resistant factors at their promoters compared to the rest of the genes.

Table 4.4: Comparisons in occupancies between the two gene groups from Figure 4.7a tested by unpaired t-test (Left). Comparisons between all time-points and 0h for the two gene groups from Figure 4.7a tested by paired t-test.

Difference of mean signal on the promoter between low and high [MNase]						Difference of mean signal on the promoter between low and high [MNase]				
			p-value						p-value	
Time-points	Group 1	Group 2	Rep1	Rep2		Gene list	Group 1	Group 2	Rep1	Rep 2
0h	High at 6h	Rest	<2e-16	1.10E-05		Genes having high differences at 6h	0h	1h	<2e-16	<2e-16
1h	High at 6h	Rest	<2e-16	7.00E-08			0h	2h	<2e-16	<2e-16
2h	High at 6h	Rest	<2e-16	<2e-16			0h	3h	<2e-16	<2e-16
3h	High at 6h	Rest	<2e-16	<2e-16			0h	4h	<2e-16	<2e-16
4h	High at 6h	Rest	<2e-16	<2e-16			0h	5h	<2e-16	<2e-16
5h	High at 6h	Rest	<2e-16	<2e-16			0h	6h	<2e-16	<2e-16
6h	High at 6h	Rest	<2e-16	<2e-16			0h	control	<2e-16	<2e-16
control	High at 6h	Rest	<2e-16	<2e-16		All the other genes	0h	1h	<2e-16	< 2e-16
							0h	2h	<2e-16	1.20E-05
							0h	3h	2.90E-09	< 2e-16
							0h	4h	<2e-16	< 2e-16
							0h	5h	<2e-16	< 2e-16
							0h	6h	<2e-16	< 2e-16
							0h	control	<2e-16	< 2e-16

The natural question arising from this result is whether the high difference between MNase-sensitive vs MNase-resistant proteins for this group of genes is due to an increase in occupancy of MNase-sensitive proteins or a decrease of MNase-resistant proteins or a combination of

both. To understand this, the occupancy of these two types of proteins for those genes is represented like in the previous chapter, namely as the mean of occupancy on the promoter after a low and a high MNase treatment (Fig 4.7b, annex 4.5b and Table 4.5).

In both replicates the same trend is observed: an ~3-fold increase in occupancy after a low MNase treatment as cell size increase (from ~0.006 median occupancy at 0h to ~0.017 at 6h) and a slight decrease of occupancy after a high MNase treatment (from 0.004 at 0h for replicate 1 and 1h for replicate 2 to below 0.002 at the latest time-points). This means that these genes undergo a progressive and important size-dependent recruitment of MNase-sensitive proteins on their promoter which is accompanied by a removal of MNase-resistant proteins at this locus.

Table 4.5: Comparison between time-points after a low and a high MNase concentration tested by paired t-test (Left). Comparison between MNase concentration for each time-points tested by unpaired t-test (Right).

Mean signal on the promoter between low and high [MNase] for genes having high occupancy at 6h					Mean signal on the promoter between low and high [MNase] for genes having high occupancy at 6h				
			p-value					p-value	
[MNase]	Group 1	Group 2	Rep1	Rep 2	Time-points	Group 1	Group 2	Rep1	Rep2
low [MNase]	0h	1h	<2e-16	<2e-16	0h	low	high	<2e-16	<2e-16
	0h	2h	<2e-16	<2e-16	1h	low	high	<2e-16	2.30E-07
	0h	3h	<2e-16	<2e-16	2h	low	high	<2e-16	<2e-16
	0h	4h	<2e-16	<2e-16	3h	low	high	<2e-16	<2e-16
	0h	5h	<2e-16	<2e-16	4h	low	high	<2e-16	<2e-16
	0h	6h	<2e-16	<2e-16	5h	low	high	<2e-16	<2e-16
	0h	control	<2e-16	<2e-16	6h	low	high	<2e-16	<2e-16
high [MNase]	0h	1h	<2e-16	< 2e-16	control	low	high	<2e-16	<2e-16
	0h	2h	<2e-16	3.57E-01					
	0h	3h	<2e-16	<2e-16					
	0h	4h	<2e-16	<2e-16					
	0h	5h	<2e-16	0.9412					
	0h	6h	<2e-16	8.00E-08					
	0h	control	<2e-16	< 2e-16					

All the other genes also show an increase in occupancy of MNase-sensitive proteins in the promoter but to a much lesser extend (Fig 4.7c, annex 4.5c and Table 4.6). Moreover, no clear decrease of occupancy of MNase-resistant proteins can be seen throughout the time-course and the median occupancy of MNase-sensitive vs MNase-resistant proteins stay close (0.0015 for MNase-resistant and 0.005 for MNase-sensitive at 6h). This shows a different regulation of chromatin organization around the TSS of the two groups of genes. The list of the 838 genes is enriched in cell communication (p-value = 6.09e-11), regulation of macromolecule metabolic process (p-value = 1.01e-06), RNA biosynthetic process (p-value = 1.91e-05) and is depleted in translation elongation (p-value = 0.00035) (analysis made with the AnGeLi tool like previously (Bitton et al. 2015). Interestingly, the list contains H2A and H2B but not H3 and H4, raising the possibility of a different regulation of chromatin structure at the promoter of those genes when cell size increases.

Table 4.6: Comparison between time-points after a low and a high MNase concentration tested by paired t-test (Left). Comparison between MNase concentration for each time-points tested by unpaired t-test (Right).

Mean signal on the promoter between low and high [MNase] for all the other genes					Mean signal on the promoter between low and high [MNase] for all the other genes				
			p-value					p-value	
[MNase]	Group 1	Group 2	Rep1	Rep 2	Time-points	Group 1	Group 2	Rep1	Rep2
low [MNase]	0h	1h	<2e-16	<2e-16	0h	low	high	0.00018	<2e-16
	0h	2h	<2e-16	<2e-16	1h	low	high	1.80E-15	<2e-16
	0h	3h	<2e-16	<2e-16	2h	low	high	<2e-16	<2e-16
	0h	4h	<2e-16	0.07735	3h	low	high	1.50E-13	<2e-16
	0h	5h	<2e-16	<2e-16	4h	low	high	<2e-16	<2e-16
	0h	6h	<2e-16	<2e-16	5h	low	high	<2e-16	<2e-16
	0h	control	<2e-16	<2e-16	6h	low	high	<2e-16	<2e-16
high [MNase]	0h	1h	1.40E-11	<2e-16	control	low	high	<2e-16	<2e-16
	0h	2h	<2e-16	<2e-16					
	0h	3h	<2e-16	<2e-16					
	0h	4h	<2e-16	<2e-16					
	0h	5h	7.20E-08	<2e-16					
	0h	6h	2.70E-06	<2e-16					
	0h	control	<2e-16	<2e-16					

The next question is thus whether these genes are differentially expressed in normal conditions or during the time-course. To test this, the overlap between the two replicates of the genes having a high difference at 6h is then considered and analyzed by RNA-seq (Fig 4.8a). Strikingly, the two groups of genes adopt quite different transcriptional states (Fig 4.8b). The 836 genes having the high size-dependent occupancy of the MNase-sensitive proteins show at 0h and throughout the time-course a doubled-median of RNA levels (50 rpkm vs 25 rpkm for the rest of the genes) (Martínez Segura, 2017). Therefore, the group presenting higher expression might require increased recruitment of MNase-sensitive factors at their promoter and not MNase-resistant factors.

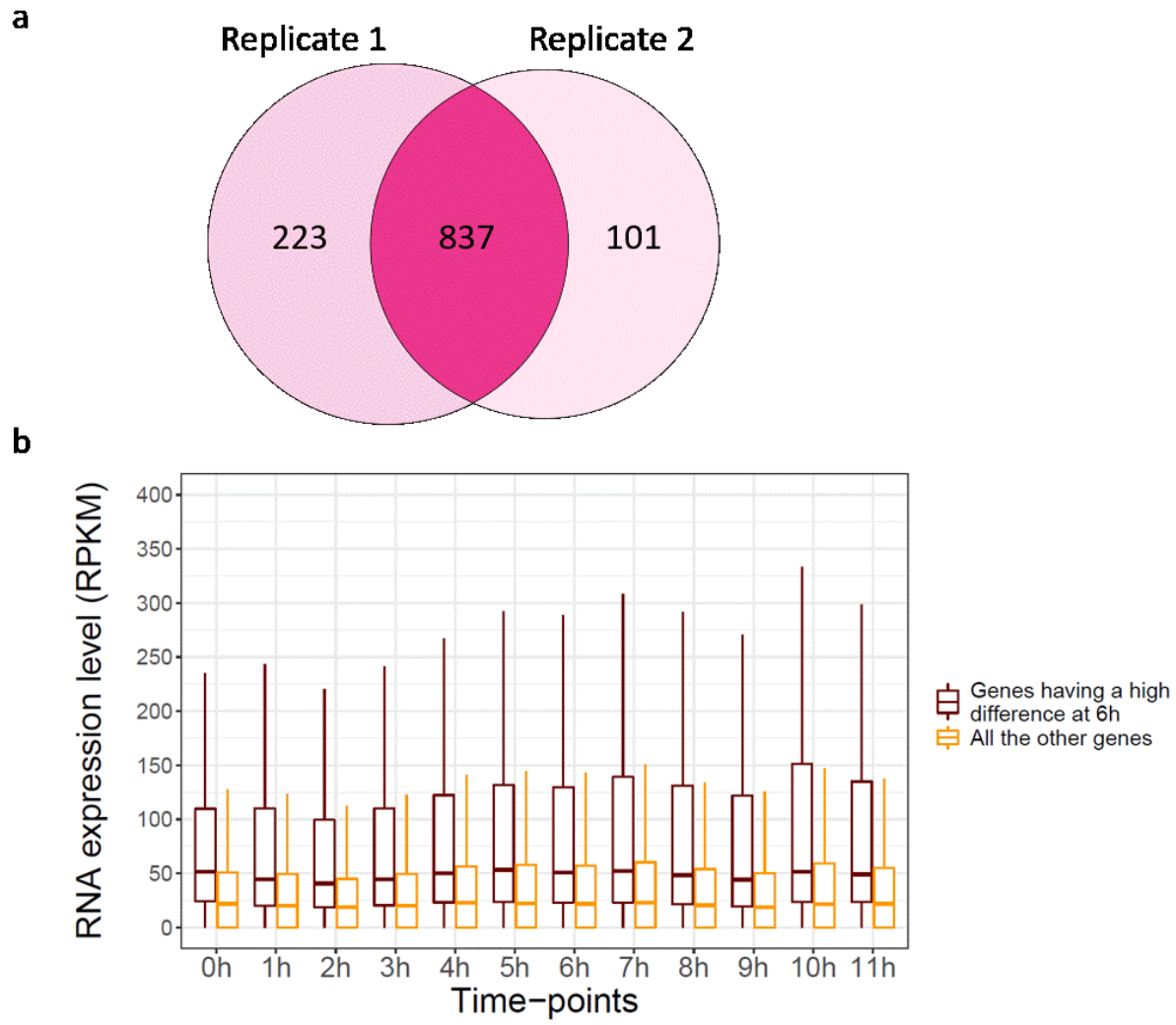


Figure 4.8: Gene expression of the two groups of genes showing differences in promoter occupancy. ***A.*** *Overlap of genes having a high difference at 6h between the two replicates (fisher exact test gives a significant overlap).* ***B.*** *RNA-seq showing the RNA levels of two groups of genes. The 2 groups of genes are statistically different at all time-points (comparison tested by unpaired t-test).*

4.5. Discussion

To summarize this second chapter of results, I showed a clustering analysis where most genes exhibit a higher occupancy of MNase-sensitive proteins / complexes at their promoter upon cell size increase (Fig 4.9). Each cluster can be distinguished by the particular chromatin organization around the TSS of the genes in normal proliferating cells.

Cluster 1 gathers genes having a relatively clear NDR at the promoter in basal conditions (Fig 4.2a, b and annex 4.2a, b), which is poorly affected by higher MNase treatment.

Still in basal conditions, cluster 2 promoters are occupied by a or several factors that are MNase-sensitive for the one localized just upstream of the TSS and MNase-resistant for the most upstream of the TSS. The footprint left by the latter is similar between normal and large cells; this means that this particular chromatin structure does not depend on cell size (Fig 4.2c, annex 4.2c). Since nucleosomes are usually resistant to MNase, one could propose that the region upstream of the promoter defined herein is indeed packaged into nucleosomes. One question here is thus whether these putative nucleosomes are involved in the modulation of transcription levels: is their presence preventing, enhancing or indifferent to transcription? In *S. pombe*, the regulation of the fbp1 gene shows how the chromatin conformation far upstream of the TSS can participate in transcription regulation (Umeda et al., 2018; Hirota et al., 2008). Briefly, under glucose rich conditions, the region upstream of fbp1 is packaged into nucleosomes, leading to a repressive chromatin structure and resistant fragments after MNase digestion. Despite this repressive state of the gene and chromatin structure, the RNAPII occupies an upstream locus far from the fbp1 TSS and initiates the transcription of ncRNAs. This model is an interesting example of how regions further upstream of the TSS could still expose a footprint even after a high MNase treatment. After further genome-wide investigation by tiling arrays, other loci were described to have the same pattern of transcription upstream of a coding gene (Hirota et al. 2008). These genes, the UPF0300 family protein 6, meu10, msy1 and tgd1 are all contained in cluster 2, suggesting that those genes may bear the mark of such regulation even after treatment with a high MNase concentration. Furthermore, some ncRNAs are found around regulatory regions of the coding genes such as the promoter (Takemata and Ohta, 2017) and their transcription by RNAPII promote the local loading of TFs close to the promoter of the coding genes. It is noteworthy that cells from the control condition, which were starved from glucose for 6h, still present the MNase-resistant footprint upstream of the promoter (Fig 4.2f and annex 4.2f), further analysis at the single gene level to confirm the presence of a nucleosome at these loci. Alternatively, the specific pattern of cluster 2 could be explained by the presence of alternative TSS that could recruit the transcription machinery upstream of the canonical TSS and consequently leaves a footprint

after MNase digestion. However, the large existence of such alternative TSS has been disputed in *S. pombe* (Thodberg et al. 2019).
Interestingly, the histone genes that are organized in pairs are found in this list, including hta1, htb1, hht1, hht2, hhf1, hhf2 and hhf3. An analysis at a single gene level indeed confirmed the presence of a MNase resisting particle upstream of the genes. In the case of hht2 and hhf2, the putative nucleosome upstream of the TSS might, at least for the histone genes found in cluster 2, correspond to the +1 nucleosome of the adjacent pair. Once again, a closer analysis at the single-gene level for the other genes is required to confirm this hypothesis.

The promoters gathered in cluster 3 resemble promoters from cluster 1, namely without binding of proteins at a low and a high MNase concentration (Fig 4.2a, d, g and annex 4.2a, d, g). Strikingly, they present a high occupancy of the +1 nucleosome. This list contains many ribosomal protein genes and is enriched with cytoplasmic translation, which were indeed described to possess a strongly positioned +1 nucleosome (Rawal et al., 2018; Nocetti and Whitehouse, 2016; Kubik et al., 2015; Knight et al., 2014; Mavrich et al., 2008b). As was observed on the transcription level from figure 4.4b, those genes were also described as highly expressed (Mivelaz et al., 2020; Metzl-Raz et al., 2017; Kubik et al., 2015; Marguerat et al., 2012). One possible scenario is that the phased +1 nucleosome over those genes emerge from the action of GRFs and RSC which exposes the TBP binding site by shifting the +1 nucleosome aside (Kubik et al., 2018). This could in turn drive the high transcription level of those genes. Moreover, these studies also depict the characteristic architecture laying at the promoter of those genes as containing a “fragile” nucleosome. Although the role of these mysterious nucleosomal particles remained to be defined, scientists have speculated that they contribute to the downstream transcription of the gene by the RNAPII (Kubik et al., 2015; Pradhan et al., 2015; Knight et al., 2014). A “fragile” nucleosome does not appear in these data, showing that an even lower MNase concentration might be needed to uncover their presence.

Nevertheless, all three clusters present clearly an increase in recruitment at the promoter upon cell size increase (Fig and annex 4.2 and Fig and annex 4.3). The size-dependence of this phenomenon is particularly appreciated for cluster 2, for which the promoters were already occupied in basal condition. One could hypothesize that, for promoters contained in cluster 2, the factors in normal-sized cells could serve as a scaffold for further and more efficient

recruitment upon cell size increase. For all three clusters, the increase in occupancy upon size increase can emerge from two scenarios: more and more promoters are occupied within each cluster; alternatively, more and more cells exhibit proteins at the promoter for a gene, increasing the probability of signal at each location. Thanks to the heatmap (Fig and annex 4.1 and 4.2), one can see that the signal throughout the promoters intensifies with cell size increase, consistent with the second scenario. Single-cell MNase-seq is required to confirm this hypothesis. The size-dependent changes of occupancy of MNase-sensitive factors also depend on the basal chromatin organization (e.g., for cluster 2, since the occupancy is already higher than the other clusters at 0h, the increase associated with the cell elongation is also higher).
For each cluster, an interesting prospect is to search for common motifs on the promoter by using the MEME tool (Bailey et al., 2009)/ This analysis could point towards possible candidates that are present at the promoter in a cell size-dependent fashion and which could be subsequently tested by ChIP-seq. The state of chromatin organization for each cluster might be associated with specific patterns of gene expression since RNA levels of each group present different median rpkm.

Despite the higher expression levels of cluster 3 throughout the time-course, the distinct chromatin states of each cluster are not responsible for the differential scaling of the genes while cell size increases (negatively non-scaling and positively non-scaling genes). Nevertheless, positively non-scaling promoters contain genes that are lowly expressed in normal conditions (this analysis and (Martínez Segura 2017)). Since this group is enriched in catabolic processes, one can speculate that expression of these genes is less needed in a normal state of proliferation and at normal size. In addition, these genes present a different chromatin structure on their promoter; they are bound with MNase-resistant factors that are not size-dependent. In normal-sized cells, it is possible that these factors are involved in transcription regulation in order to maintain these genes in a lowly expressed state. Since these genes are expected to increase their transcription in large cells, we can conclude that such size-independent factors are not playing a key role in adjustment of transcription as cell size increases until 6h after addition of 1NM-PP1. However, after 7h, these genes start being expressed at a very high level, even overtaking the level of scaling genes. How could these genes be expressed that high after 7h? It is possible that the MNase-resistant factors that were associated with low expression at 0h are now removed from the promoter, enabling a higher

range of transcription rate compared to other genes. Such a regulation, if they involved histone proteins, is reminiscent of the different state of chromatin structure on the promoter of inducible genes, which allow them to be expressed at a much higher range than constitutive genes (introduction and (Tirosh and Barkai, 2008)). Consequently, a direct prospect of this work would be to conduct a low and a high MNase treatment on cells from 7h to 11h after addition of 1NM-PP1. Such an experiment would help understanding if MNase-resistant factors are indeed removed from promoters of positively non-scaling genes when those ones start having expression levels superior to the rest. Finally, although all genes recruit more and more MNase-sensitive proteins on their promoter as a function of cell size, all genes are not doing so in the same range. In other words, the ratio of MNase-sensitive vs MNase-resistant fragment do not change equally for all the genes during the time-course. A particular set of genes exhibits a more important recruitment of MNase-sensitive proteins concomitantly with a slight removal of MNase-resistant proteins while cell size increases (Fig 4.7 and annex 4.5). Besides, this trend is associated with RNA levels that are higher than for the other genes and this occurs in basal conditions as well as throughout the time-course (Fig 4.8). This last result paves the way for a potential regulation of transcription scaling of highly expressed genes through chromatin organization on the promoter. It is possible that this group of genes requires specific proteins such as TFs or chromatin remodelers that can in turn participate in the destabilization of MNase-resistant complexes. We can imagine a model whereby chromatin remodelers are recruited at the promoters of these genes and destabilize nucleosomes upstream of the TSS. This action might promote the efficient recruitment of the TF(s) and RNAPII necessary to promote the high transcriptional state of these genes and consequently, their transcriptional scaling to cell size. Again, this might involve the formation of a “fragile” nucleosome. But why would some genes exhibit a different range of promoter occupancy as a function of cell size, although their expression is scaling? A recent study (Lanz et al., 2021) shows a continuum of scaling behaviors across the proteome, meaning that all the genes exhibit a non-scaling behavior to a certain extend. Other studies shed light on the effect of the promoter sequence on the scaling behavior of the gene, since the trans factors regulating the promoter strength are assumed to be in excess (Lin and Wang, 2021; Sun et al., 2020). More precisely, the promoter has a specific ability to recruit RNAPII and TF depending on its sequence (Lin and Wang, 2021). This would explain why transcription scaling occurs within a spectrum. However, not all the proteins regulating transcription might be in excess: the histones have a particular

transcription regulation. In addition to being expressed in S-phase (Eriksson et al., 2012; Takayama and Takahashi, 2007), their promoter shows an increased affinity for TFs and/or RNAPII, which would probably leave a footprint when treated with a low MNase treatment In line with this, it was reported earlier that histones were part of cluster 2, which shows high promoter occupancy in basal condition compared to the other clusters. Are genes from cluster 2 following the same rule, namely having a relatively high promoter strength for RNAPII and TFs, GRFs or chromatin remodeling? Firstly, the cluster 2 promoters are not saturated since we can see an increase in occupancy at this locus when cell size increases. But are these factors truly efficient for transcription? Indeed, their function could be impaired by the formation of factor aggregates that could lead to the high footprint seen in large cells. However, it does not seem to be the case since we do not observe a sub-scaling behavior for this group of genes, but a constant concentration of their transcripts.

For now, the MNase-seq data in cells of increasing size do not allow to rule out the possibility that histone proteins, forming a “fragile” nucleosome, are occluding the promoter of these genes at large size. That is why I will explore this possibility in the next chapter that will thus focus on histone proteins.

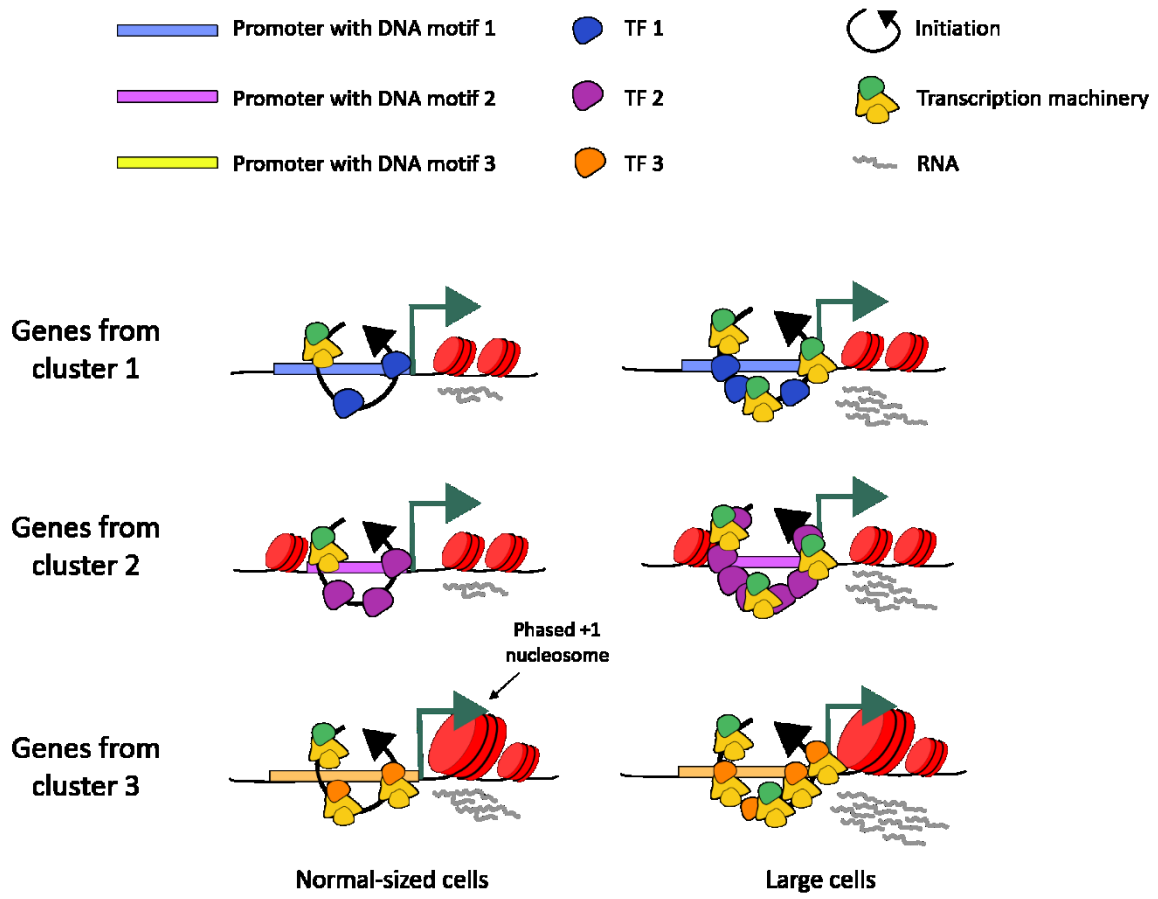
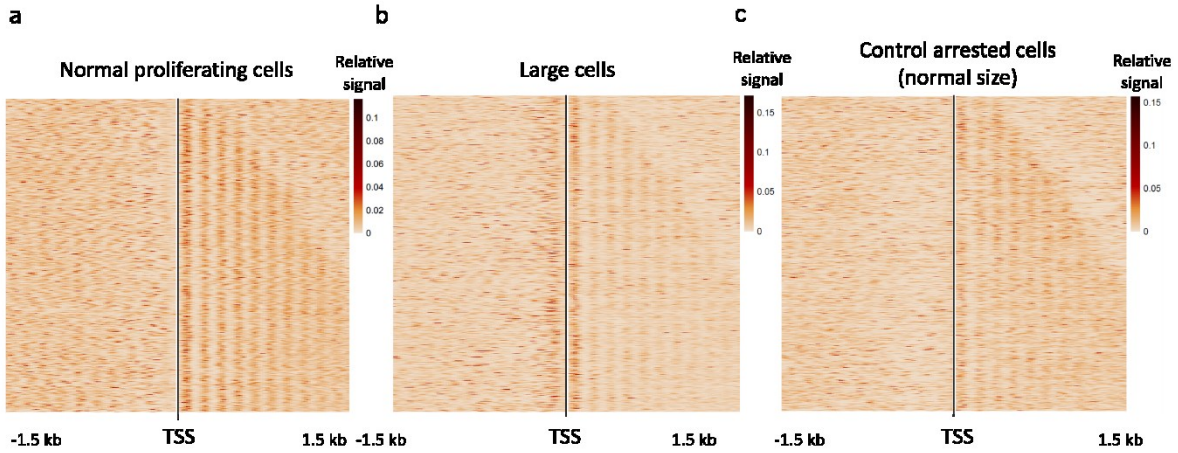
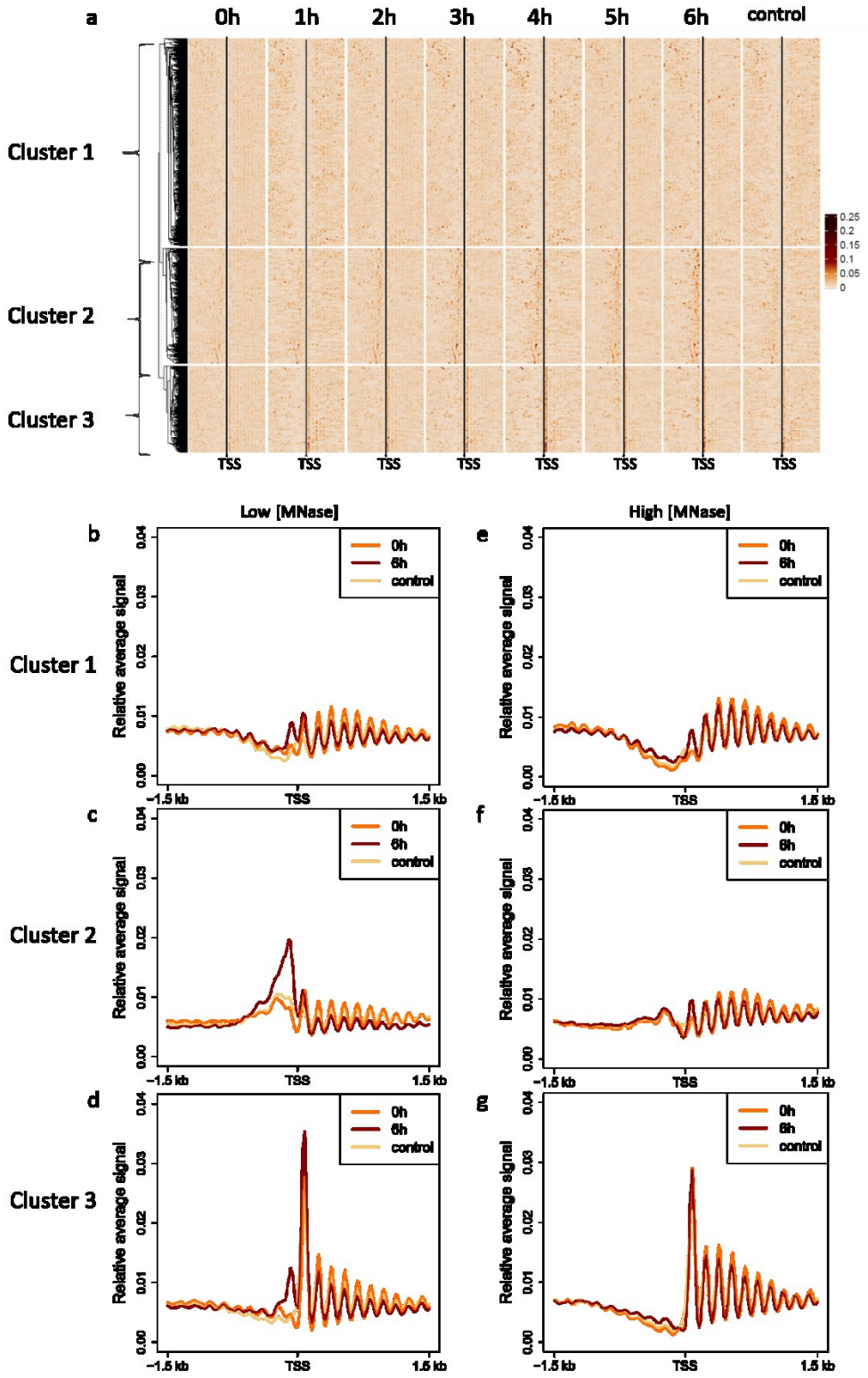


Figure 4.9: Hypothetical model explaining the profiles from each cluster after MNase treatment. *For each cluster, a global increase in occupancy occurs on the promoter upon cell size increase. Cluster 1 is the largest cluster and present the same architecture as seen in chapter 3 for the average profile for normal and large cells. Cluster 2 exhibit a relatively high occupancy in basal conditions on the promoter that gets even more occupied upon cell size increase. Cluster 3, enriched in ribosomal protein genes and in translation functions, shows a phase +1 nucleosome and is associated with high RNA levels in basal conditions and throughout* cdc2-asM17 *elongation (Amalia Martinez Segura, 2017). Cluster 3 also present an increase in occupancy on the promoter as cell size increases. The different promoter architecture might depend on specific DNA motif which recruits distinct types of TFs.*

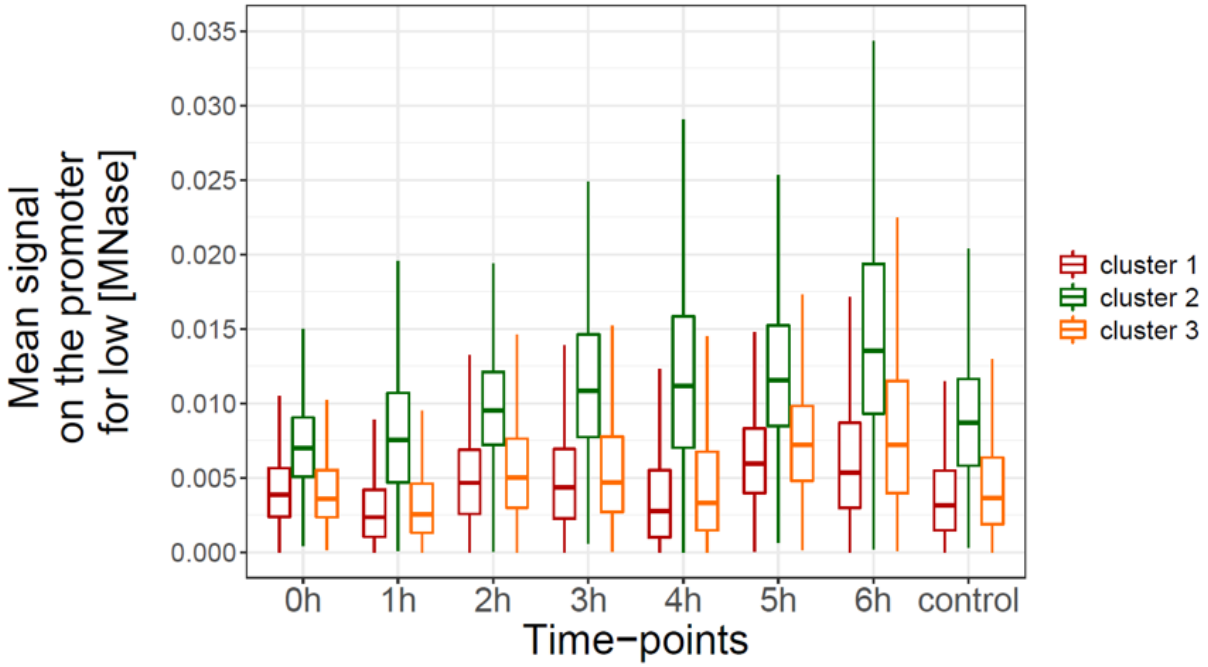
ANNEXES



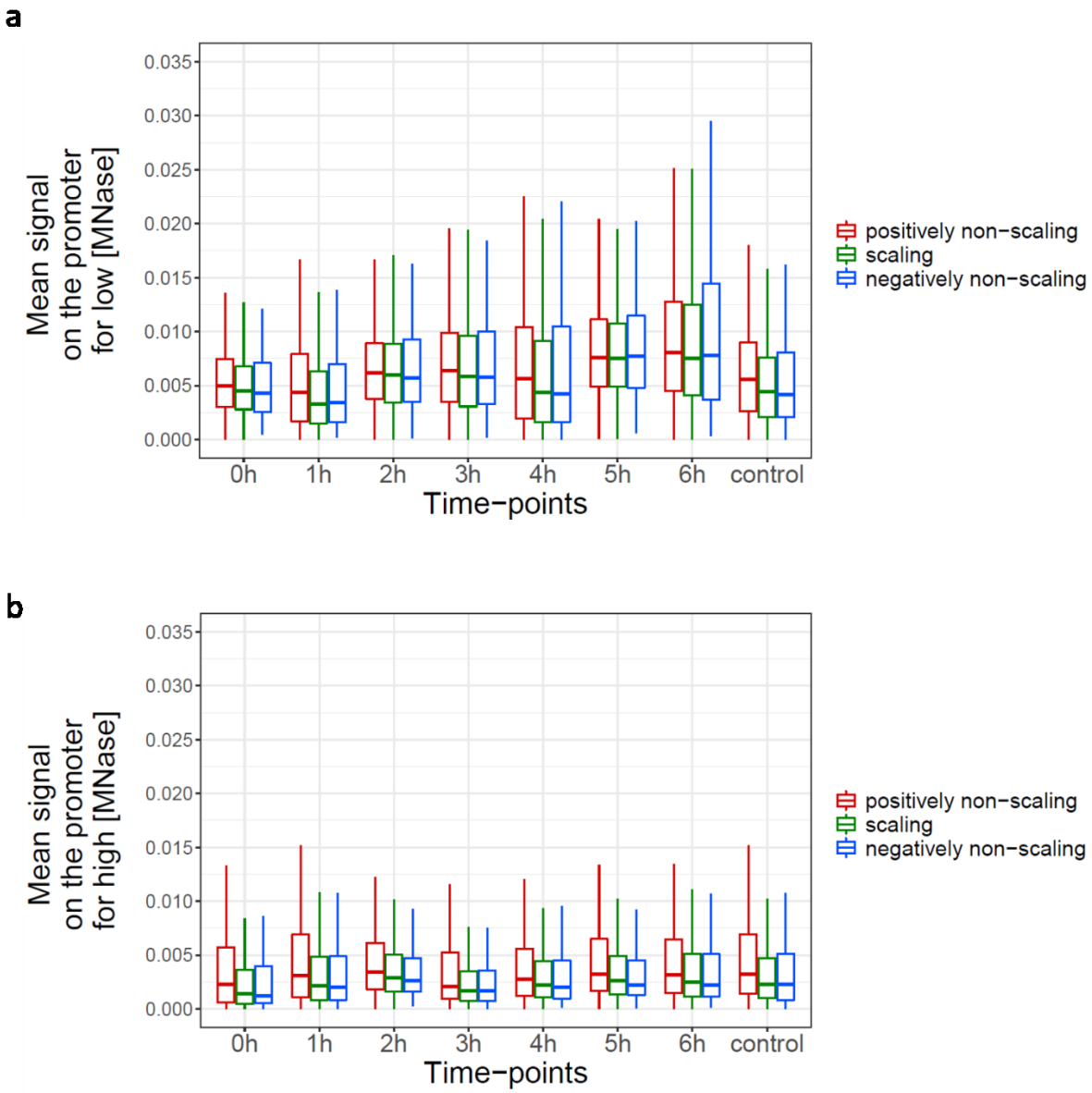
Annex 4.1: Detailed heatmap of the chromatin profile 1.5 kb around the TSS, obtained at low MNase concentration and for cells of different sizes*: normal proliferating cells* ***(a)****, large cells arrested with 1NM-PP1 for 6h* ***(b),*** *normal-sized cells arrested 6h with 1NM-PP1 and deprived of glucose for 6h* ***(c).***



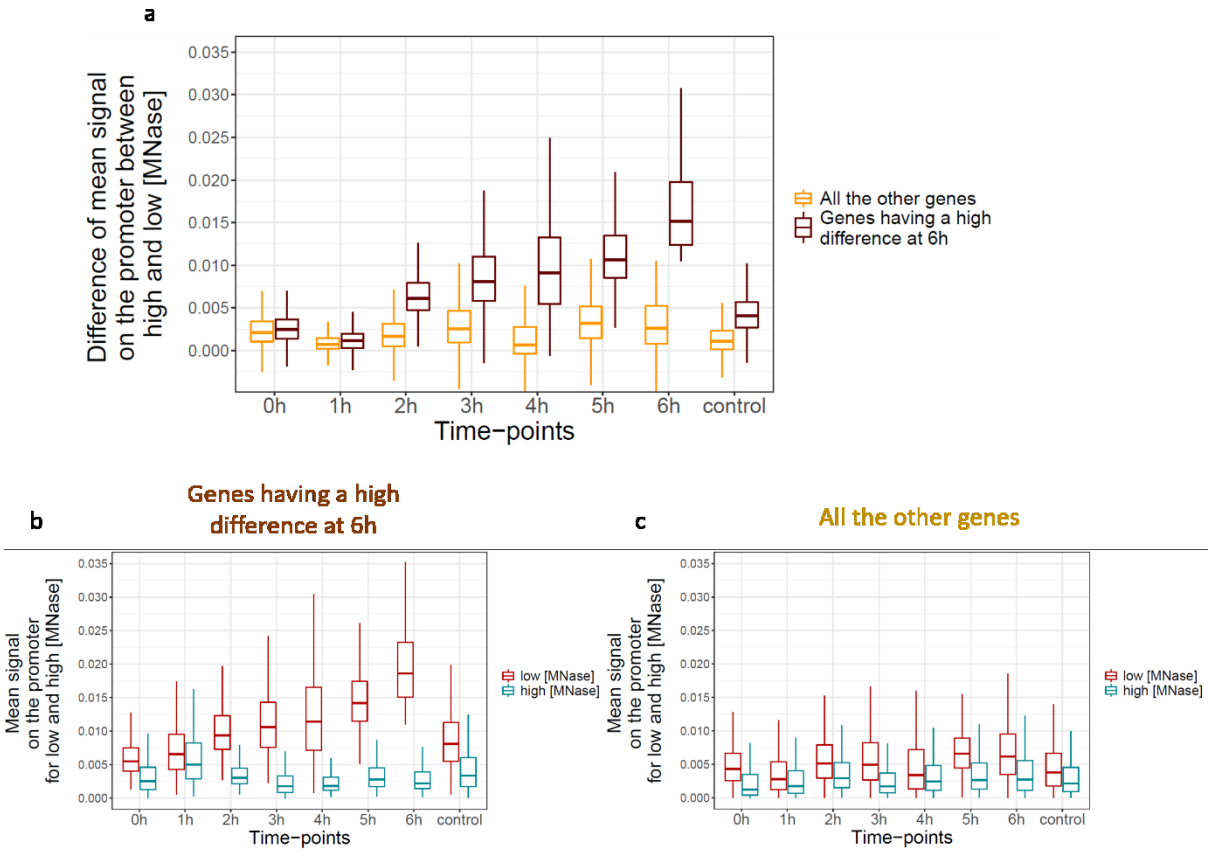
Annex 4.2: Three clusters of genes having specific chromatin organization profile 1.5 kb around the TSS obtained at low MNase concentration and for cells of different sizes. A. *Detailed heatmap for cell population each hour after addition of 1NM-PP1 until 6h and finally 6h with 1NM-PP1 and without glucose (control)* ***(a). B-g.*** *Average profile from a low MNase concentration of cluster 1* ***(b)****, cluster 2* ***(c)*** *and cluster 3* ***(d)*** *and from a high MNase concentration of cluster 1* ***(e)****, cluster 2* ***(f)*** *and cluster 3* ***(g)*** *for normal proliferating cells (0h), large cells (6h with 1NM-PP1) and control (6h with 1NM-PP1 and without glucose).*



Annex 4.3: Mean signal on the promoter for low MNase concentration and for each time-point after addition of 1NM-PP1 and for the control condition



***Annex 4.4: Mean signal on the promoters**, at a low MNase concentration **(a)** and at a high MNase concentration **(b)**.*



Annex 4.5: Two groups of genes showing differences in the promoter occupancy. A. *Difference of occupancy between a low and a high MNase treatment on the promoter as a function of time after 1NM-PP1 (paired t-test to test differences between the timepoints for both groups and unpaired t-test to test the difference between the groups of genes at all the time-points).* ***B-c.*** *Occupancy on the promoter after low and high MNase treatment for the genes having a high difference at 6h* ***(b)*** *and all the other genes* ***(c)*** *(paired t-test to test differences between the timepoints for both groups and unpaired t-test to test the difference between the groups of genes at all the time-points).*

Chapter 5: Expression of two histones H2A and H3 and genome-wide occupancy of H3 as a function of cell size

With the MNase-seq experiments, I could explore the global chromatin organization for cells of increasing size. This approach was fruitful as it permits to distinguish a singular pattern on the promoter of larger cells, namely a large and high footprint of a, or several, factor(s) that is/are MNase-sensitive. In the second chapter of results, I show the reader that such a pattern is pervasive as it occurs on all the genes, although diverse groups of genes initially present distinct chromatin architectures. However, the size-dependent increase in signal on the promoter at a low MNase concentration depends on the initial state of chromatin around the TSS. In addition, genes that tend to be more expressed than the rest in basal conditions, and throughout the time-course, also show a higher occupancy of MNase-sensitive proteins concomitantly to a slight decrease of MNase-resistant factors at their promoters when the cells get larger. The natural question emerging from such work is whether the peak seen on the promoter of large cells is composed of histones or non-histone proteins. Even if the peak is fragile because of its sensitivity to MNase, the proteins underlying this peak could be non-canonical nucleosomes (Brahma and Henikoff, 2019; Kubik et al., 2015; Knight et al., 2014; Xi et al., 2011). Moreover, I observed in the first chapter of results, a sudden drop in promoter accessibility at intermediate time-points (3h or 4h depending on the replicates, chapter 3, **Fig 7a and annex 7a**), which might be explained by the presence of histones. Hence, it is important to explore the behavior of histone proteins. Here, I choose to report the level of two histones, H3 and H2A, as a function of cell size and by two different approaches: microscopy and western blot. Why study those histones? H3 can be heavily modified on its tail and these PTM are widely studied (Taylor and Young, 2021; Chan and Maze, 2020; Jenuwein and Allis, 2001). In addition, the tools for studying H3, such as antibodies, are very abundant. H2A is the histone that was used previously in our lab and for which we have a strain enabling its tracking by microscopy (Sun et al. 2020). Then, I focus on the difference of H3 localization on the DNA between normal-sized and large cells.

5.1. The change of histone abundance as a function of cell size

The abundance of histones as a function of cell size is an aspect that has interested many scientists who measured the levels of histones in *S. cerevisiae* and in human cells (Claude et al., 2021; Jiang and Zhang, 2021; Lanz et al., 2021; Swaffer et al., 2021a). The outcome of these

analyses is that the level of histones does not scale with cell size, as opposed to most proteins. This intriguing result was also reported in *S. pombe* by Amalia Martínez Segura who performed proteomics on the *cdc2-asM17* strain for 11 hours after addition of 1NM-PP1 (Martínez Segura 2017). More precisely, she analyzed the proteome of the elongating *cdc2-asM17* and, adding standards peptides at a known concentration in each sample, she could report the absolute quantity of each peptide contained in the proteome (Martínez Segura 2017). From these normalized mass spectrometry data, I have represented in figure 1 the abundance for the four core histones H2A, H2B, H3 and H4 with time after 1NM-PP1, relatively to time-point 0h (**Fig 5.1**). Very clearly, the abundance of all four histones decreases when the cells become larger. However, at early time-points, H2A abundance increases and reaches its maximum at 3h. The fact that histone proteins belong to the class of sub-scaling proteins was indeed previously observed (Claude et al., 2021; Lanz et al., 2021; Swaffer et al., 2021a), but their abundance level was believed to remain constant, since it is supposed to match the genomic content (Claude et al., 2021). It is noteworthy that these findings were reported in *S. cerevisiae*, thus leaving space for a different size-dependent histone regulation in *S. pombe*. Nevertheless, these intriguing proteomic results generated by Amalia Martínez Segura merit further investigations to perceive better the regulation of histones as a function of cell size.

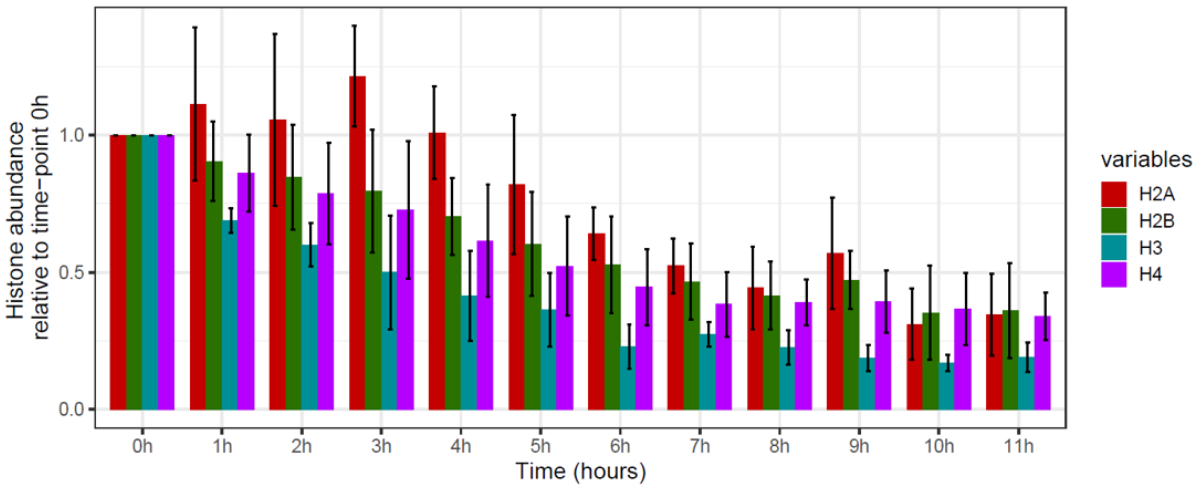


Figure 5.1: Mass spectrometry data from Martínez Segura (Martínez Segura, 2017) showing the abundance for the four core histones as a function of cell size.

5.2. The abundance of histones H2A, H3 and of the acetylation on the 9th lysine of H3

- H2A

Another way to understand the dynamics of histone abundance as a function of cell size is to detect these proteins by microscopy, at a single-cell level. Our group possesses two strains I could cross with the *cdc2-asM17* strain to quantify the level of H2A alpha in cells of increasing size. Both strains encode a fusion protein, the first one containing H2A alpha fused to the GFP and the second one containing UCH2 (a marker of the nuclear envelope) fused to the mcherry protein. After selecting the resulting SBM320 mutant (see method) with the wanted genetic background (*cdc2-asM17:: uch2-mcherry:: hta2-GFP*), the cells are grown in the same conditions as before, 6h with 1NM-PP1. At each hour, a picture is taken to visualize consecutively the two fluorophores fused with the proteins of interest (Fig 5.2a, b). After acquiring enough pictures to quantify the signal, the kinetics of H2A normalized by Uch2 can be represented as a function of cell size (Fig 5.2c). However, because the analysis is done at the single cell level and the imaged cells are not the same throughout the time-course, I decided not to normalize the signal by the time-point 0h. A similar trend as from proteomics is deduced from the quantification: a signal augmentation at 3h and a progressive decrease until 6h. However, no statistical difference is reported between 0h, 1h and 2h (Table 5.1). Therefore, it is quite clear that H2A undergoes a specific regulation during the time-course so that its expression level in the nucleus can rise at 3h and decrease below the level from basal conditions at larger sizes.

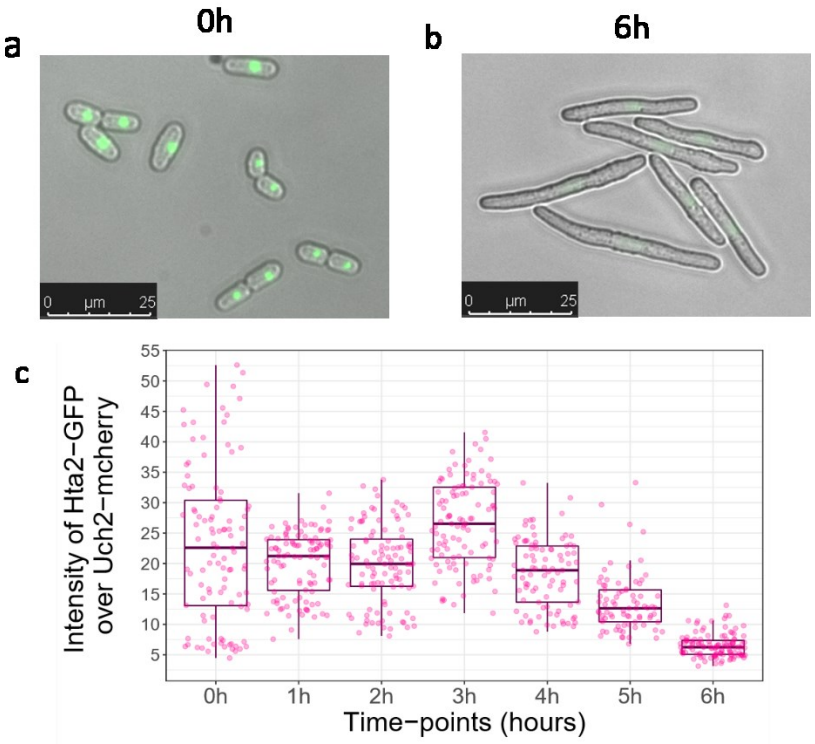


Figure 5.2: Figure 3: Abundance of H2A as a function of cell size. A*. Microscopy image at 0h merging the channel showing H2A-GFP with the brightfield channel.* ***B.*** *Microscopy image at 6h after 1NM-PP1 merging the channel showing H2A-GFP with the brightfield channel.* ***C.*** *Quantification of the GFP signal normalised by the mcherry signal for every time-point after 1NM-PP1. The comparisons are tested by Wilcoxon test.*

Unfortunately, no control condition is available for this experiment, and we cannot rule out that this specific regulation is an effect of the cell cycle. How could this change of histone abundance be due to the cell-cycle? As explained in the introduction, Cdc2 is the kinase responsible for the progression of the cell cycle; it acts in G1 as well as in G2. Therefore, if a proportion of the population is in the G1 phase when 1NM-PP1 is given to the cells, they will first go through S-phase before arresting in G2. Since the histones are synthesized in S phase, the quantity of histones of these cells will increase after those cells pass through S-phase. To understand if a substantial proportion of cells is indeed going through S-phase before 3h, I used the RNA-seq data created by Amalia Martínez Segura (Martínez Segura 2017) to represent the normalized expression of a group of genes involved in the G1-S transition (Fig 5.3).

Table 5.1: Comparison of H2A beta intensities from Figure 5.2c tested by Wilcoxon test.

Time-points	group1	group2	p-value	
0h	0h	6h	< 2e-16	****
	0h	1h	0.1406	ns
	0h	2h	0.1528	ns
	0h	3h	0.0045	**
	0h	4h	0.0178	*
	0h	5h	5.60E-07	****
6h	6h	1h	< 2e-16	****
	6h	2h	< 2e-16	****
	6h	3h	< 2e-16	****
	6h	4h	< 2e-16	****
	6h	5h	< 2e-16	****
1h	1h	2h	0.5362	ns
	1h	3h	1.40E-10	****
	1h	4h	0.0235	*
	1h	5h	6.60E-14	****
2h	2h	3h	2.20E-10	****
	2h	4h	0.1611	ns
	2h	5h	4.70E-11	****
3h	3h	4h	1.60E-13	****
	3h	5h	< 2e-16	****
4h	4h	5h	1.40E-08	****

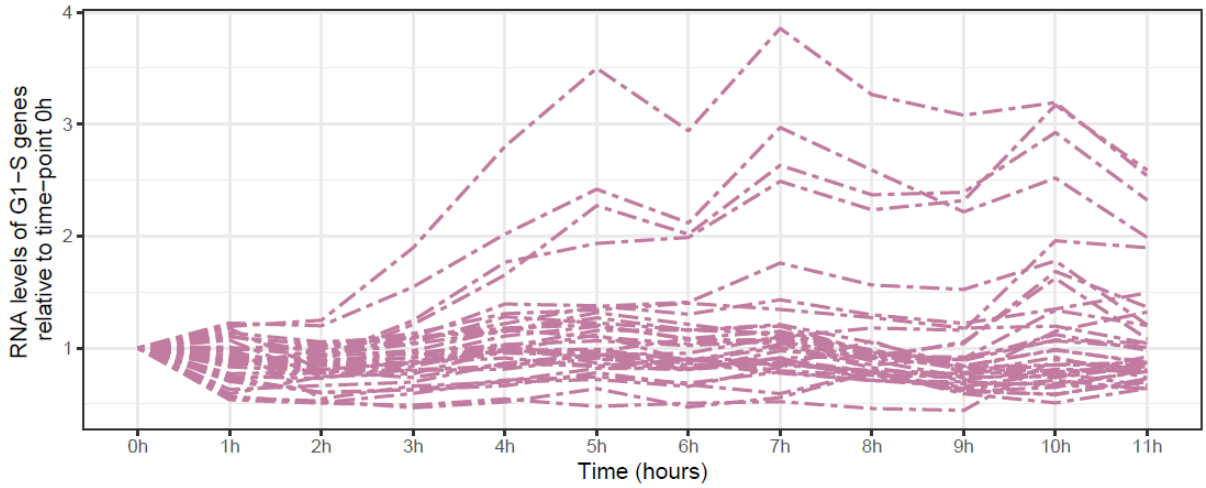


Figure 5.3: RNA levels of genes involved in the G1/S transition during the time-course after 1NM-PP1.

For most of those genes, no clear increase around 3h can be witnessed, except for four genes (cig2, nrm1, rep2 and yox1) which level increases from 2h. Therefore, the theoretical cells undergoing S phase represent a small fraction of the population and thus could not account for the significant increase in H2A in the nucleus at 3h (Table 5.1). Therefore, one can hypothesize that this increase around 3h is associated with a certain size (~20 um at 2h), rather than a specific stage of the cell cycle.

- H3 and H3K9ac

H3 is another core histone forming the nucleosome and although its abundance was assessed by Martínez Segura by proteomics (**Fig 5.1**), the result will be confirmed here by western blot. By loading the same concentration of proteins in each well, western blots often consist in assessing the change of concentration of a protein of interest. However, this time, the method consists in visualizing the protein amount for the same number of cells, and not the protein amount for the same concentration of proteins (**Fig 5.4**).

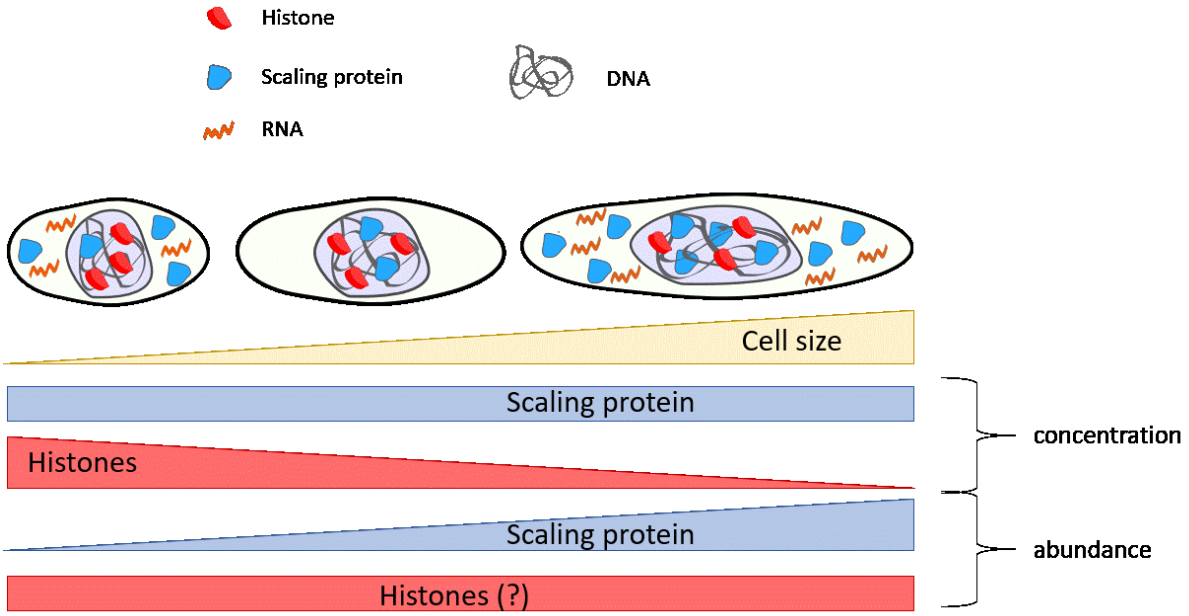


Figure 5.4: Representation of the distinction between concentration and abundance of proteins in the context of cell size increase. *The concentration of the scaling proteins remains constant while cell size increases whereas its abundance increase. Histone concentration decreases while cell size increases and previous findings showed that their abundance remains stable in absence of DNA replication (Claude, Martinez).*

This distinction is key as it allows us to measure the change in histone abundance with cell size, rather than the change in histone concentration. To this end, the proteins for the same number of *cdc2-asM17* cells are first extracted in the same volume, at each time-point after 1NM-PP1. The total protein measurement by Bradford assay (**Fig 5.5a**) confirms that increasing protein quantities are present as the cells are growing in size. Consistently with the increase in cell length (four times more after 6 hours with 1NM-PP1) and because *S. pombe* width is supposed to stay constant during its growth (Wood and Nurse, 2015), the total protein abundance is four

times more at 6h compared to 0h. This illustrates the principle of scaling at the protein level. To assess the change in protein abundance with cell size increase, protein amounts corresponding to the same number of cells are deposited in each well of the western blot gel for each time-point after 1NM-PP1. After transferring the proteins from the gel to the membrane and after staining the membrane with Ponceau Red, a gradient of signal is visible from 0h to 6h after 1NM-PP1 (Fig 5.5b). This reflects the increase in total protein with cell size increase, as seen with the Bradford assay and thus proves that protein amounts corresponding to the same number of cells were deposited on the gel at each time-point.

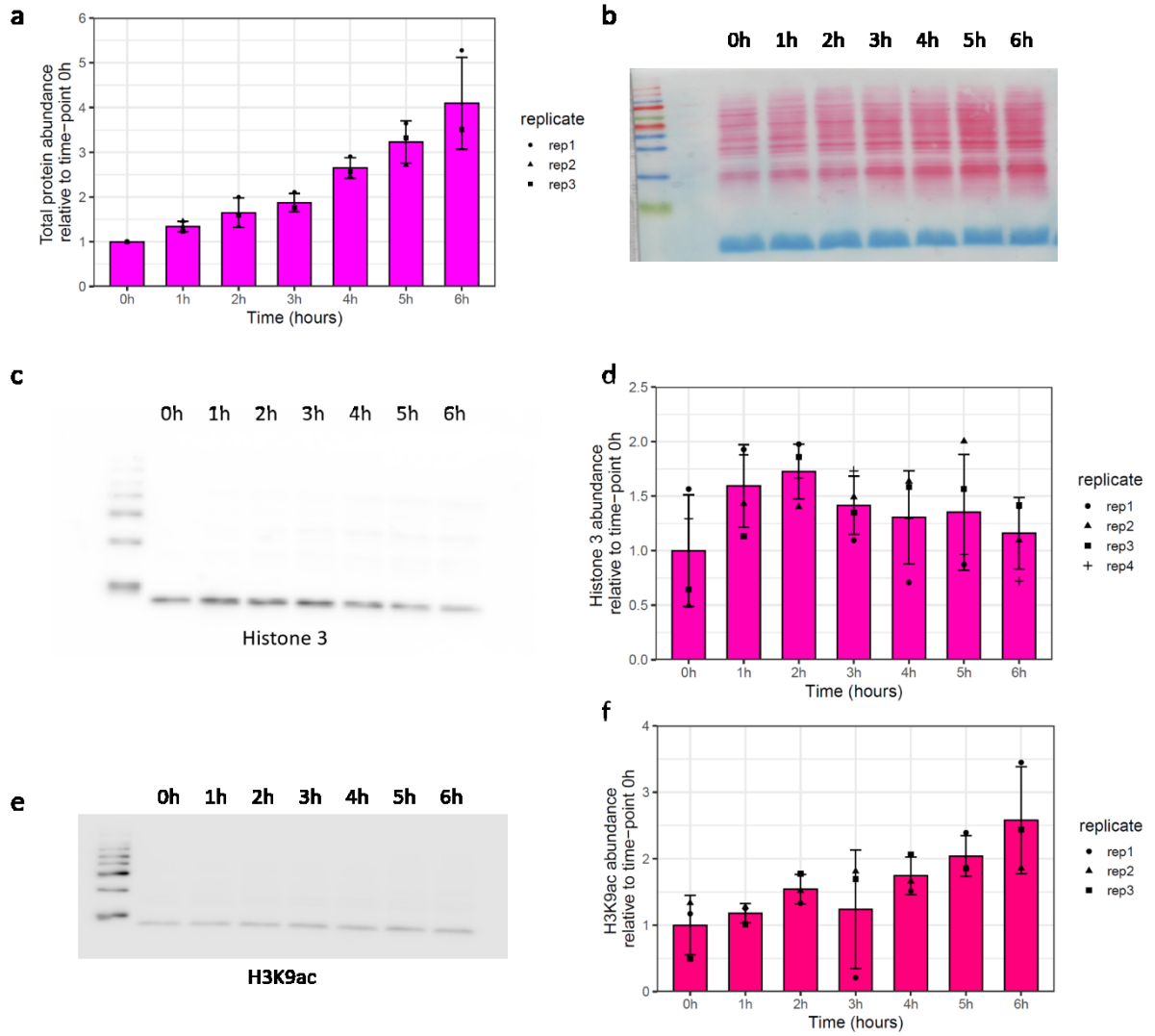


Figure 5.5: Abundance of H3 and H3K9ac as a function of cell size. A*. Total protein abundance measured by Bradford assay for 6h after addition of 1NM-PP1.* ***B.*** *Ponceau Red showing the abundance of total proteins after transferring the proteins from the gel to the nitrocellulose membrane and for 6h after 1NM-PP1****. C.*** *Western blot using anti-H3 antibody on the samples from* ***(b). D.*** *Quantification of the western blot on c. with normalisation by the average of all the signal and by the time-point 0h.* ***E.*** *Western blot with an anti-H3K9ac antibody on the samples from* ***(b). F.*** *Quantification of the western blot on* ***(e)*** *with normalisation by the average of all the signal and by the time-point 0h*

On this same membrane, a detection of histone 3 is conducted with an anti- H3 antibody (Fig 5.5c). The corresponding quantification of histone 3 abundance with cell size increase is depicted on figure 5.5d (Fig 5.5d). This analysis comprises four biological replicates. For every time-point, the intensity signal is normalized by the average signal of all time-points and then by time-point 0h. The trend of H3 abundance increases at the beginning of the time-course:

the average of H3 signal at 2h is 1.75 more than at 0h. Then, it progressively decreases but not enough to reach the level at 0h: the average signal at 6h is slightly higher than 1. This is reminiscent of the trend seen for H2A on the proteomics and by microscopy (Fig 5.1 and 5.2c). However, considering the four replicates and the error bars, the abundance of histone 3 is mainly stable with cell size increase, contrasting with the progressive decrease of H3 seen by mass spectrometry. Therefore, since a clear increase at 3h was witnessed only for H2A, one can imagine that H2A and H3 undergo different regulations throughout the time-course. Moreover, this increase in H2A does not seem to emerge from the passage of the cells through the S phase since all the genes involved in G1/S transition are not more expressed before 3h.

Histone 3 can be chemically modified on its tail and some of those post-translational modifications, such as acetylation, attest to transcription activation (Gates et al., 2017; Kratz et al., 2010; Sinha et al., 2006; Pokholok et al., 2005; Hebbes et al., 1988). Even more importantly, acetylation, in collaboration with other factors and PTMs, can potentiate transcription (Donczew et al., 2020; Jang et al., 2005; Matangkasombut and Buratowski, 2003; Tjian and Maniatis, 1994). Acetylation on the 9th residue of histone 3, H3K9ac, is usually associated with transcription initiation (Gates et al., 2017; Karmodiya et al., 2012; Rhee and Pugh, 2012; Kratz et al., 2010; Sinha et al., 2006; Pokholok et al., 2005). Since transcription initiation rates scale with cell size (Swaffer et al., 2021b; Padovan-Merhar et al., 2015; Sun et al., 2020), does this modification abundance change with cell size? After probing the membrane with an antibody directed against H3K9ac (Fig 5.5e), the same quantification as before was done (Fig 5.5f). Since H3 abundance mainly remains stable, the quantification of H3K9ac is not normalized by the H3 abundance. Strikingly, H3K9ac increases proportionally to cell size (Fig 5.5f). Except the 3h time-point, all the other time-points show a higher signal than the precedent one, indicating the progressive deposition of acetylation on the 9th lysine of histone 3 as the cells are growing in size. At 6h, the level of H3K9ac abundance is 2.5 more than at 0h. This new result suggests that the kinetics of a modification of H3, H3K9ac, is dependent on cell size.

As mentioned before, H3K9ac can be found in the promoter of active genes (Gates et al., 2017; Karmodiya et al., 2012; Rhee and Pugh, 2012; Kratz et al., 2010; Sinha et al., 2006; Pokholok et al., 2005). Since this PTM scales with cell size, one could hypothesize that it is linked to the

chromatin structure at the promoter of cells of increasing size (chapter 3 and 4). Moreover, although the size-dependent changes of H3 abundance are not very clear (diminution observed by Mass Spectrometry and constant abundance seen by western blot), its distribution around the TSS can be deciphered by ChIP-seq. In other words, can we establish a difference of H3 enrichment between large and proliferating cells? Answering this question will finally settle the matter of whether the fragile peak, appearing as cell size increases, is composed of H3 and is thus likely to be a "fragile" nucleosome.

5.3. The fragile peak seen on the promoter of large cells contains H3 for few genes

To understand whether the chromatin structure on the promoter of large cells is composed of histone proteins, a ChIP-seq using the same antibody anti-H3 as for the western blot is conducted on normal proliferating cells and large arrested cells (see chapter 2). This approach has been used in the past to confirm that fragile nucleosomes were histone-based DNA/protein particles (Kubik et al., 2015; Xi et al., 2011). The results are represented as the log2 ratio of signal from large cells over normal cells, 1.5 kb on both sides of the TSS, in bin of 20 bp, like previously with the MNase-seq. Then, the signal of all the genes is aligned on the TSS and the average signal for all the genes is calculated (Fig 5.6a).

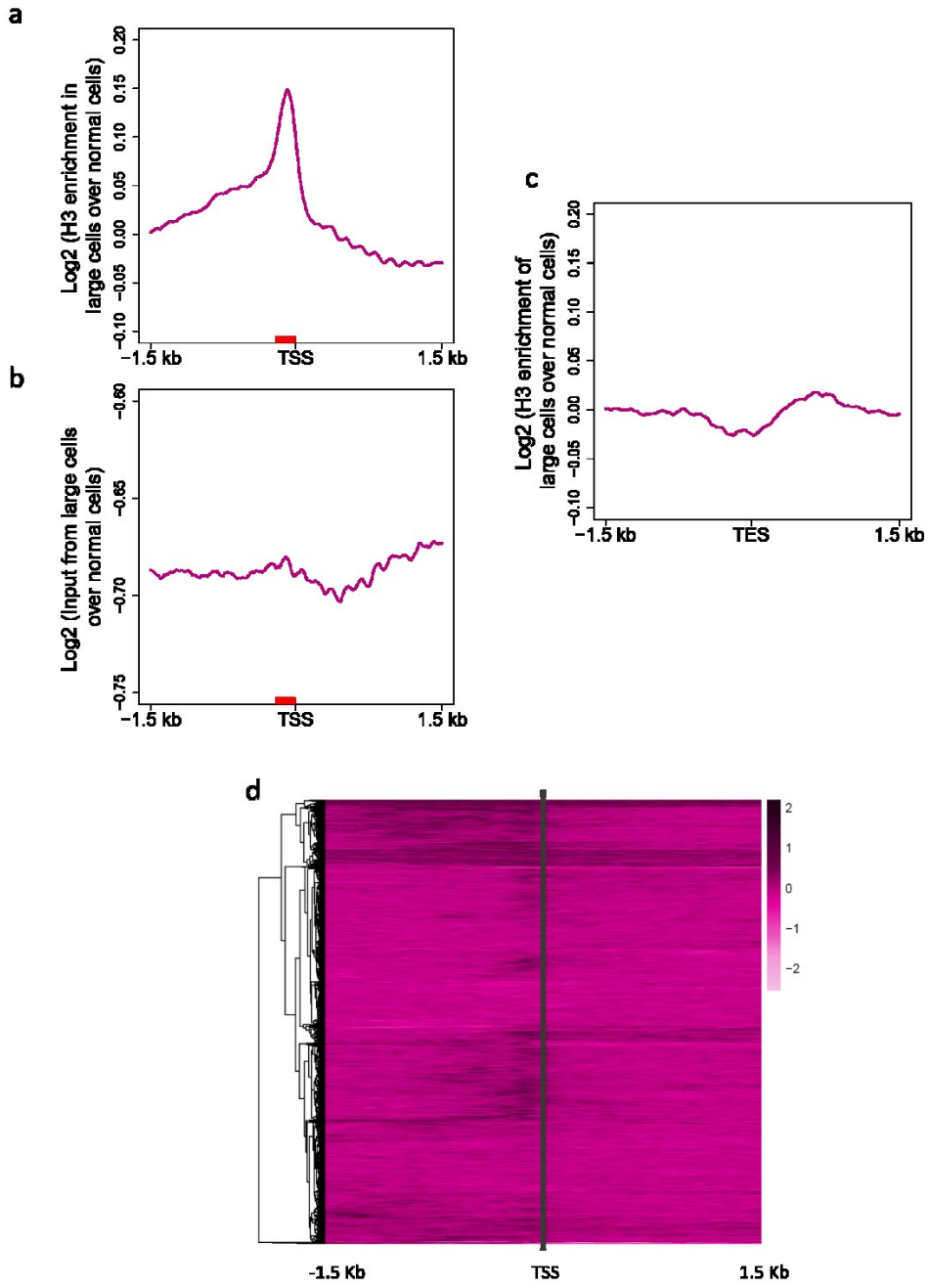


Figure 5.6: Log2 ratio of H3 enrichment between normal sized-cells and large arrested cells with 1NM-PP1. A-b. *Average profile 1.5 kb centred the TSS* ***(a)*** *and TES* ***(c)*** *showing the log2 ratio of H3 occupancy on large cells over normal-sized cells (a), (c) and showing the log2 ratio of input from large cells over input from normal-sized cells* ***(b). D*** *Heatmap showing the log2 ratio of H3 occupancy on normal-sized cells over large cells for all the genes clustered together by hierarchical clustering.*

This average profile is the average of three biological replicates that are presented separately on annex 5.1a, b and c (annex 5.1a, b and c). The promoter of large cells exhibits an obvious increase in signal, meaning that large cells have, on average, more H3 at this locus compared to normal-sized cells. To rule out that this profile is due to chromatin shearing, input samples

are visualized on Figure 5.6b and annex 5.1d, e and f for the three replicates separately (**Fig 5.6 and annex 5.1d, e, f**). Like the IP profile, the average profiles correspond to the log2 ratio of input signal from large cells over normal-size cells (**Fig 5.6b and annex 5.6d, e, f**). Importantly, the input profiles do not exhibit a similar pattern to the IP profiles (Fig 5.6a, b and annex 5.6), showing that the relative increase in signal on the promoter of large cells likely stems from an enrichment of H3 at this locus. However, when looking at the y-axis, we can see that the peak only reaches 0.15, meaning that the signal is only ~1.11 times higher in large cells compared to normal cells. To discriminate if this small effect is due to the noise of the experiment or if it is biologically relevant, the average profile around the TES is now considered (**Fig 5.6c**). The reader can observe that no such profile can be witnessed for this region. Moreover, the pattern of the slight increase in H3 on the promoter is seen for all three replicates separately (**Annex 5a, b and c**), showing the reproducibility of the increase in H3 on the promoter of large cells. Because of the small effect (only ~1.11 times more H3), we can imagine that such consistent effect is due to the enrichment of histone 3 only on few promoters in large cells rather than on all the promoters. Thus, the signal for all the genes is then visualized on a heatmap in which the genes are grouped by hierarchical clustering from the pheatmap function in R (**Fig 5.6d**). Consistent with the second scenario, few groups of genes present a higher H3 occupancy on the promoter in large cells compared to small cells whereas other genes do not show any clear differences between signal on the promoter and signal on the gene body. Therefore, for a majority of genes, the ChIP-seq experiment does not allow to observe the enrichment of H3 at the promoter of large cells, suggesting that in large cells, H3 occupancy cannot account for the fragile peak seen after MNase-seq, at least in a large subset of genes. Instead, those promoters might bear, at a low MNase concentration, the footprint left by the transcription machinery as well as the PIC and chromatin remodeling.

Then, to extract the genes having an enrichment of histone 3 occupancy on the promoter of large cells, the genes presenting a fold change of more than 2 in any bin within 200 bp upstream from the TSS are selected. This means that for those genes, H3 occupies the promoter twice more in large cells compared to normal-sized cells on average. The average profile for the 67 selected genes is presented on figure 5.7 (**Fig 5.7**) for the IP samples (**Fig 5.7a**) and the input (**Fig5.7b**).

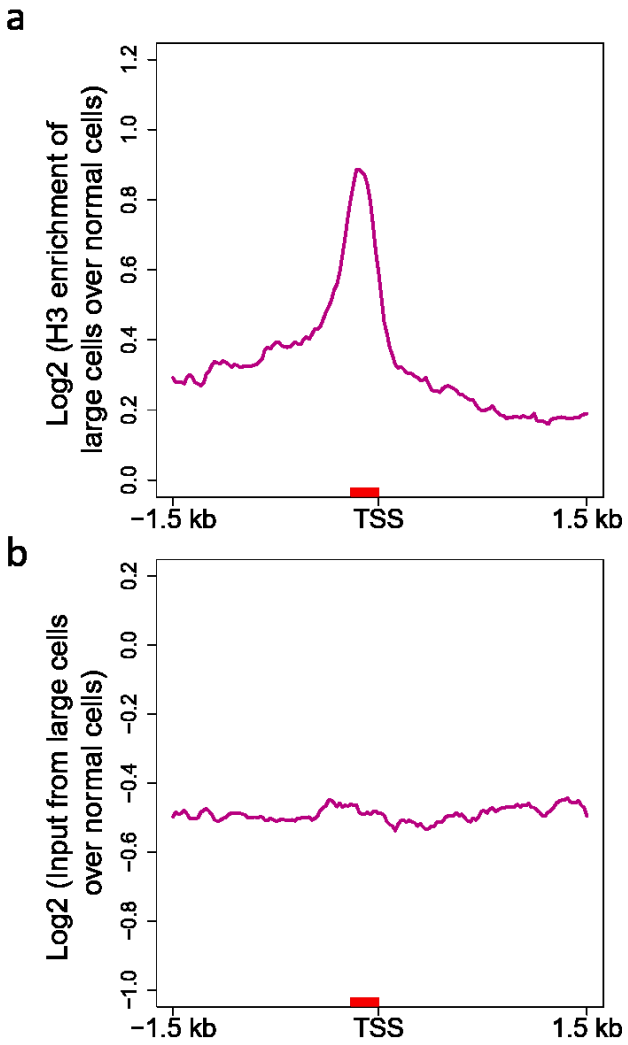


Figure 5.7: Characteristic of chromatin architecture 1.5 kb around the TSS for genes having a fold change of more than one within 200 bp upstream of the TSS. A-b. *Average profile 1.5 kb centred the TSS showing the log2 ratio of H3 occupancy on large cells over normal-sized cells* ***(a)****, and showing the log2 ratio of input from large cells over input from normal-sized cells* ***(b)***

This time, by looking at the y-axis, the reader can see that the average fold-change of H3 occupancy in large cells compared to normal-sized cells is much higher than for all the genes (the peak on the promoter reaches 0.9 for the selected genes vs ~0.15 for all the genes) (Fig 5.7a). This result reflects the higher frequency of H3 binding at the promoter in large cells compared to normal-sized cells. Using Angeli from Bahler's website, no significant enrichment was found from this list of genes (Bitton et al., 2015). It contains, for example, ribosomal genes, genes coding for proteins of the cell wall, genes coding for ncRNAs and so on so forth. Therefore, these genes cannot be linked by their function, but it would be interesting to look at their chromatin organization profile during the time-course.

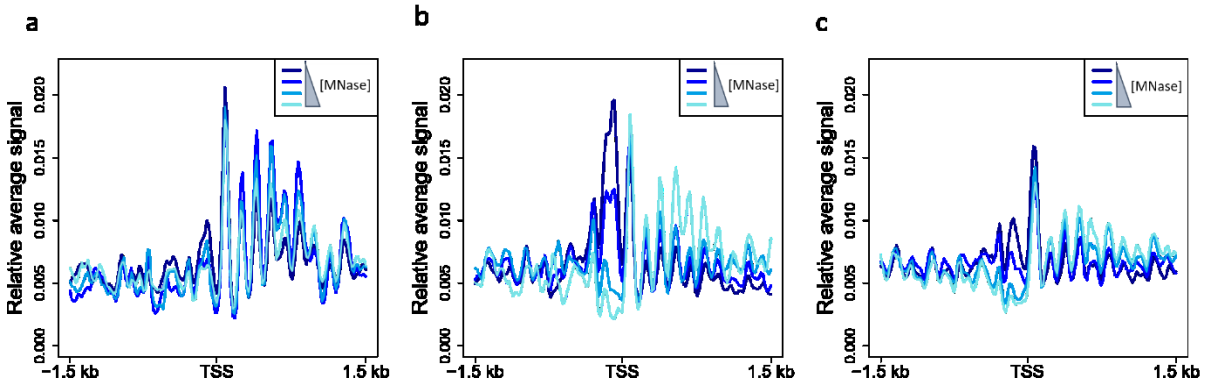


Figure 5.8: Average profile 1.5 kb on both side of the TSS for genes having a high H3 enrichment on promoter of large cells; *in normal sized-cells* ***(a),*** *large cells* ***(b)*** *and control cells* ***(c)*** *after an increasing MNase treatment.*

Their average chromatin profile at 0h, 6h and for the control condition is represented in figure 5.8 b, c, d and for 1.5 kb each side of the TSS and for several MNase concentrations (Fig 5.8a, b, c and annex 5.2a, b, c). For normal proliferating cells (Fig 5.8a and annex 5.2a), we notice an oscillating signal of a great amplitude, reaching around 0.02 of normalized signal vs less than ~0.012 for all the genes (chapter 3, Fig 3.1a, Fig 3.6a). At a low MNase concentrations, the promoter exhibits a peak which size is comparable to the one seen on the promoter of large cells in the same conditions for all the genes (at a low MNase concentration) (Fig 5.8a, annex 5.2a and chapter 3, Fig 3.6b). A similar pattern is visible on the promoter of the control cells (Fig 5.8c and annex 5.2c). However, the height of this peak doubles on the promoter of large cells, reaching 0.02 vs 0.012 for all the genes (Fig 5.8c, annex 5.2b and chapter 3, fig 3.6b). From the ChIP-seq results, this category of genes presents an enrichment of H3 on the promoter after four times increase in size. Therefore, the peak reflecting the enrichment of H3 at large size (Fig 5.7a) overlaps the high peak seen on the promoter of large cells (Fig 5.8b and annex 5.2b). However, at a high MNase concentration, the peak disappears, like for the average profile of all the genes (chapter 3, Fig 3.6b). This means that the H3 detected more specifically in large cells compared to normal-sized cells is not incorporated into a canonical nucleosome that would be resistant to a high MNase treatment. This opens the possibility that this subset of promoters accommodates a "fragile" nucleosome upon cell size increase. In the literature, such nucleosomal particles were found at growth and highly regulated genes, such as ribosomal protein genes (Kubik et al., 2015; Knight et al., 2014; Xi et al., 2011). Therefore, the

natural question emerging from the present study is whether this subset of genes showing an enrichment of “fragile” nucleosomes upon size increase is indeed more expressed than the rest of the genes. Furthermore, are these genes presenting a differential expression while cell size increases?

5.4. Expression levels of the genes having high enrichment of H3 on the promoter

To test whether the expression level of the genes showing an enrichment of H3 at the promoter of large cells, I used once more the RNA-seq data created and analyzed by Amalia Martínez Segura (Martínez Segura 2017). I represent here the RNA levels for this group side by side to the rest of the genes (**Fig 5.9a**). Strikingly, the group of genes showing an enrichment of H3 on the promoter also shows higher median expression levels compared to the rest of the genes (~50 rpkm at early time-points vs ~slightly more than 25 rpkm for the rest of the genes). In addition, we can observe an increasing trend for the median expression of the group of genes of interest (from ~50 rpkm at early time-points and gradually reaches ~80 rpkm at 7h). A higher expression was also observed for genes having a higher occupancy of MNase-sensitive proteins at the promoter of large cells (**chapter 4, Fig 4.8**). Are these two groups sharing some genes? A Venn diagram shows the overlap between these two lists (**Fig 5.9b**). Half of the list of genes having an enrichment of H3 at the promoter in large cells is common with the other list of genes having higher occupancy of MNase-sensitive proteins at the promoter of large cells. Besides, this overlap, contains more genes than expected by chance (Fisher exact test generates a p-value inferior to 0.05. Altogether, these results suggest a specific regulation of chromatin organization at the promoter for some genes that maintain high expression levels while cell size increases. This regulation involves an enrichment of H3 at the promoter that is not yet incorporated into a nucleosome.

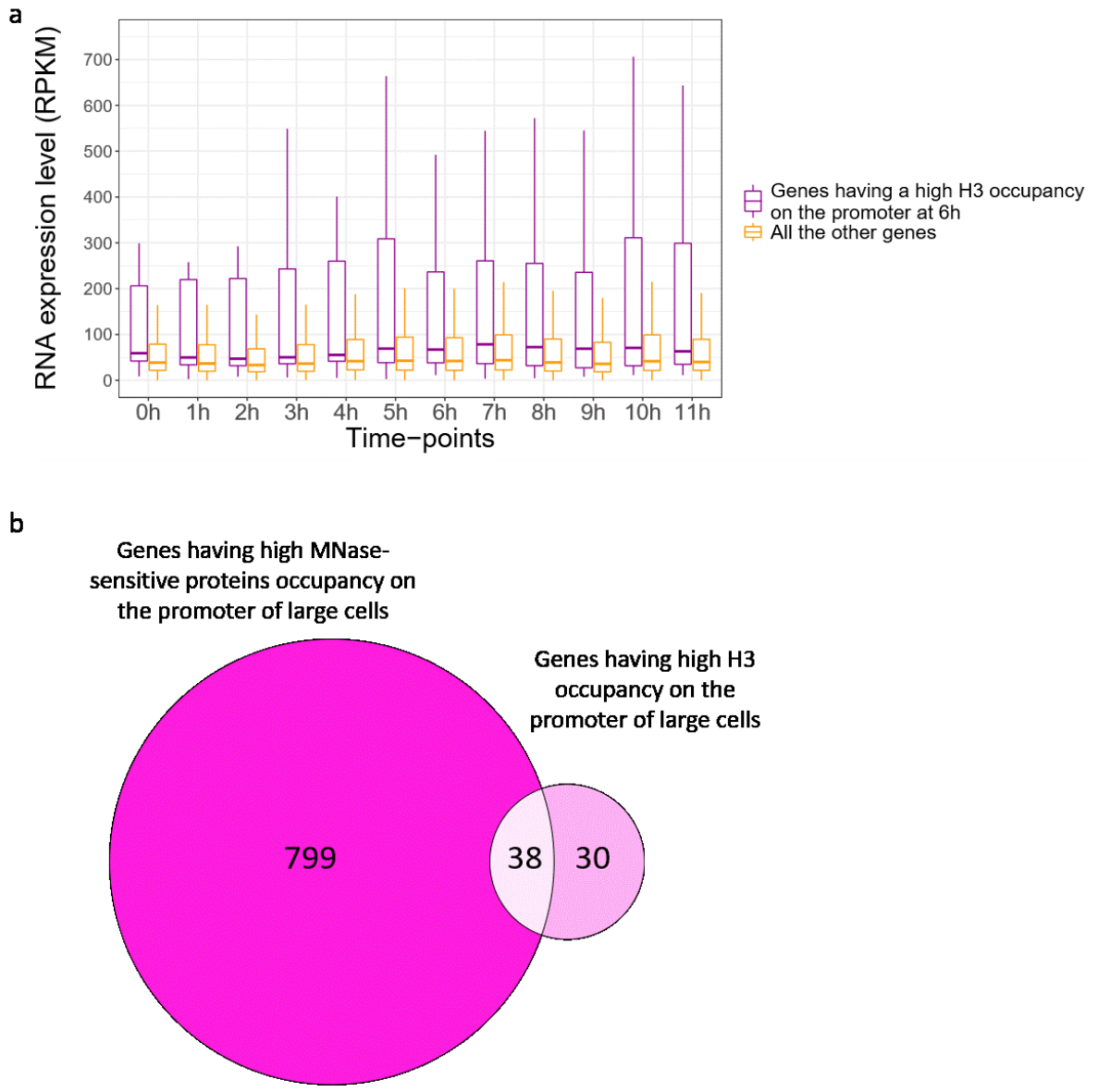


Figure 5.9: A. *RNA levels of genes having an enrichment of H3 on the promoter of large cells and RNA levels of all the other genes.* ***B.*** *Venn diagram showing the overlap between genes having high MNase-sensitive proteins occupancy on the promoter of large cells and genes having high H3 occupancy on the promoter of large cells (Fisher exact test gives a p-value inferior to 0.05).*

5.5. Discussion

This last chapter of results focuses on dynamics of histones as a function of cell size: their abundance, their modifications, and their binding to the DNA. Firstly, the abundance of the four core histones measured with mass spectrometry by Amalia Martínez Segura (Martínez Segura 2017) is going down as cell size increases (Fig 5.1). However, this gradual decrease is not clear for H2A which abundance first increases before decreasing below the basal level. This differential regulation is also seen when measuring the abundance of these two histones separately (Fig 5.2c, and 5.5c, d). However, the trend of H3 abundance was not the same when observed by western blot and mass spectrometry, although it is different from H2A trend with both techniques. Since there was a good conservation of H2A trend between mass spectrometry and microcopy (Fig 5.1 and 5.2c), one prospect of this work would be to test H3 levels also by microscopy instead of western blot.

On the one hand, H2A alpha levels in the nucleus were reported here by microscopy thanks to GFP fused to hta2 and mcherry fused to Uch2, a marker of the nuclear envelop. The genomic organization of hta2 (coding for H2A beta) is distinct from hta1 (coding for H2A alpha) and the other histone genes coding for H3, H4 and H2B. While hta2 is isolated similarly to pht1 (coding for H2A.Z), the other histone genes form pairs of transcribing units of opposing direction (chapter 4, Fig 4.4). Besides, hta2 was not associated with a size-independent nucleosome upstream of the promoter, as it was suggested for the core histones all contained in cluster 2, including H3 (chapter 4, Fig 4.2 and Fig4.4). Therefore, these distinct organizations, associated with specific chromatin state upstream of the promoter, could explain why those proteins do not share a common regulation while cell size increases.

On the other hand, the two histone types, H3 and H2A, are part of the gene list having high RNA levels and a progressively higher occupancy of MNase-sensitive proteins on the promoter as cell size increases (chapter 4, Fig 4.7). On the contrary, H3 and H4 were not identified in this list. Altogether, these observations lead to the proposition of a model wherein the distinct chromatin state of H3 and H2A in basal conditions and upon cell size increase defines their differential regulation as a function of cell size.

Finally, in the literature, different observations report different exchange rate of H2A-H2B compared to H3-H4 in the nucleosome (~8-40 fold higher for H2A-H2B), even in a condition of cell cycle arrest in which no DNA replication occurs (Yaakov et al., 2021; Kulaeva et al., 2009; Venkatesh and Workman, 2015). One could propose that such differences would require a differential level of the H2A-H2B and H3-H4 dimers. However, since the mRNA copies of H2A-H2B versus those of H3-H4 are not differently expressed (Marguerat et al., 2012), this does not seem to be the case, at least in basal conditions.

Since the four core histones are constituting the nucleosomes, changes in histones abundance could reflect a different state of the nucleosomes on the DNA. Moreover, H2A levels were reported exclusively in the nucleus thanks to Uch2 signal which marks the nuclear envelope (Sun et al., 2020). The highest increase in H2A expression occurs at 3h, a time-point at which occurs a global decrease in promoter and +1 nucleosome locus accessibility (**Chapter 3, Fig 3.10a**). Hence, one can imagine that higher H2A amounts influences global DNA accessibility in the vicinity of the TSS. When cells become larger, mass spectrometry as well as microscopy identify a significant decrease in H2A levels (**Fig 5.1, Fig 5.2, Table.5.1**). Therefore, one possibility is that the decrease of H2A indeed modifies the state of the nucleosome in large cells close to the TSS or in the gene body, although for the latter, such changes are not detected by MNase-seq.

H3 does not remain completely unchanged while cell size increases since it accumulates an acetylation on the 9th lysine of its tail (**Fig 5.5f**). This PTM has been found in the promoter of active genes, thereby providing a marker for transcription initiation (Gates et al., 2017; Rhee and Pugh, 2012; Kratz et al., 2010; Sinha et al., 2006; Pokholok et al., 2005). Hence, we can easily imagine that an increase in H3K9ac is specifically enriched close to the promoter of all genes upon cell size increase. A ChIP-seq probing the genome wide localization of H3K9ac will be needed to confirm this idea. Importantly, changes in chromatin structure of larger cells occur at the promoter. We can thus hypothesize that H3K9ac, in collaboration with other PTM, destabilizes the nucleosomes close to the TSS to potentiate the scaling of transcription initiation. Inhibiting the Gcn5-mediated acetylation on H3 might help go beyond correlation

between acetylation abundance and cell size and give insights into the mechanism of transcription initiation scaling.

How could cells of increasing size promote higher levels of H3K9ac? One interesting idea is that large cells contain higher levels of acetyl-coA. Acetyl-coA would then be used by HATs such as Gcn5 which would in turn increase the level of H3K9ac. In line with this, large cells might contain higher levels of ATP molecules due to the proliferation of mitochondria during the *cdc2-asM17* elongation (Dischinger et al., 2008). Chromatin remodelers, like SWI/SNF or RSC, which are contacting acetylated lysines on histones, can use the energy provided by the ATP. This last scenario provides a molecular mechanism whereby large cells could sense their size by their metabolic status and communicate it to the genome in order to adjust transcription initiation.

Finally, Histone 3 enrichment does not dramatically change between large and small cells (Fig 5.6a), consistently with the stable abundance of H3 (Fig 5.5e, f). This is in accordance with a recent study (Swaffer et al., 2021b) proposing a model where scaling is mainly limited by RNAPII and occurs in an invariant chromatin accessibility. However, some genes do present a clear increase in H3 occupancy at the promoter upon size increase (Fig 5.6d and Fig5.7a). The increase in H3 enrichment in large cells occurs at the same region as does the peak seen on large cells by MNase-seq, namely on the promoter. Therefore, the scaling of those specific genes might not be exclusively explained by the model described above (Swaffer et al., 2021b) and may involve a local change of H3 occupancy at the promoter.

Since the peak on the promoter of large cells seen by MNase-seq is sensitive to high levels of MNase concentration, I concluded the presence of H3 in those cells at the promoter could not be incorporated into a canonical nucleosome. The latter would be indeed resistant to high level of MNase like nucleosomes found in the gene body.

A destabilized nucleosome (or “fragile” nucleosome) on the promoter was reported several times in different organisms (Brahma and Henikoff, 2019; Iwafuchi-Doi, 2019; Mueller et al., 2017; Kubik et al., 2015) and occurs in certain types of genes such as highly expressed genes and growth genes. Those nucleosomal particle were indeed described as a novel signature for transcription activation (Brahma and Henikoff, 2020; Iwafuchi-Doi, 2019). Such particles might be present for those genes in normal cells but at a low frequency (eg, less normal cells have

them compared to largel cells) (Fig 5.10). MNase-seq at a single-cell level is required to further grasp this concept (Lai et al., 2018).

Furthermore, half of the genes having a higher H3 occupancy at the promoter in large cells also show higher RNA levels compared to the rest of the genes, in basal conditions as well as with cell size increase (Fig 5.9). In addition, this group of genes also exhibit higher occupancies of MNase-sensitive factors on the promoter of large cells. In addition to H3, those MNase-sensitive factors could be non-histone proteins such as the RNAPII, TF or chromatin remodelers, all working together to maintain the concentration of mRNA in large cells. Therefore, is this specific chromatin architecture with a "fragile" nucleosome preserving the scaling of high RNA levels? Answering this question will definitely improve our knowledge about transcription scaling regulation and, importantly, about the gene-dependent strategies to achieve it.

To this aim, a recent method called CUT&RUN (Cleavage Under Targets and Release Using Nuclease) could be used to map specific chromatin aspects, such as chemically modified histones, TFs, GRF or chromatin remodelers and "fragile" nucleosomes (Skene and Henikoff, 2017). Briefly, a DNA-binding protein is targeted by a specific antibody coupled with MNase. The MNase cleavage operate both side of the binding site of the interest protein. The method which principle resemble ChIP-seq supposedly present a higher resolution than the latter. To illustrate this, CUT&RUN efficiently uncovered the presence of "fragile" nucleosome on the promoter in normal condition which were invisible after traditional ChIP-seq.

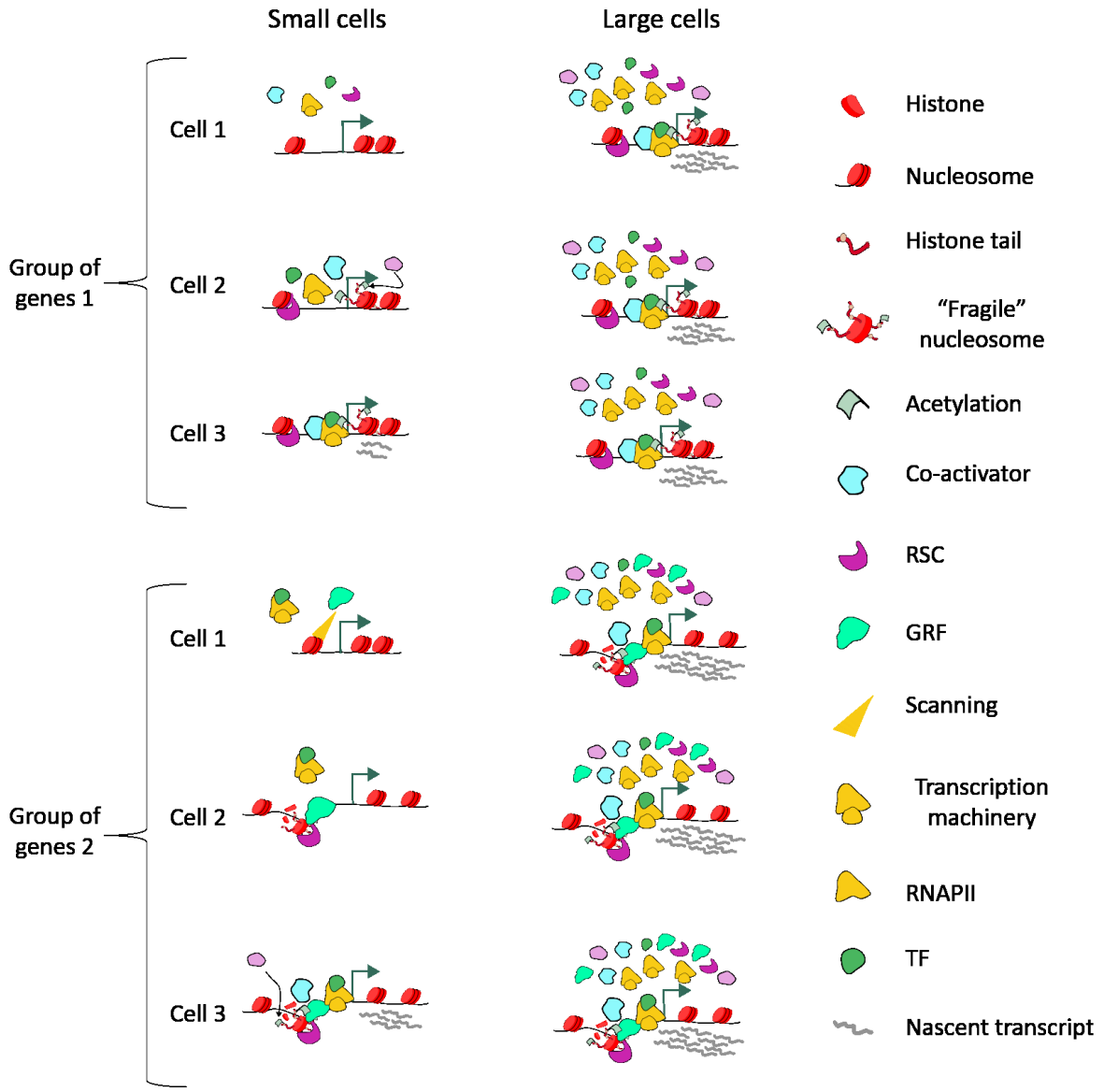
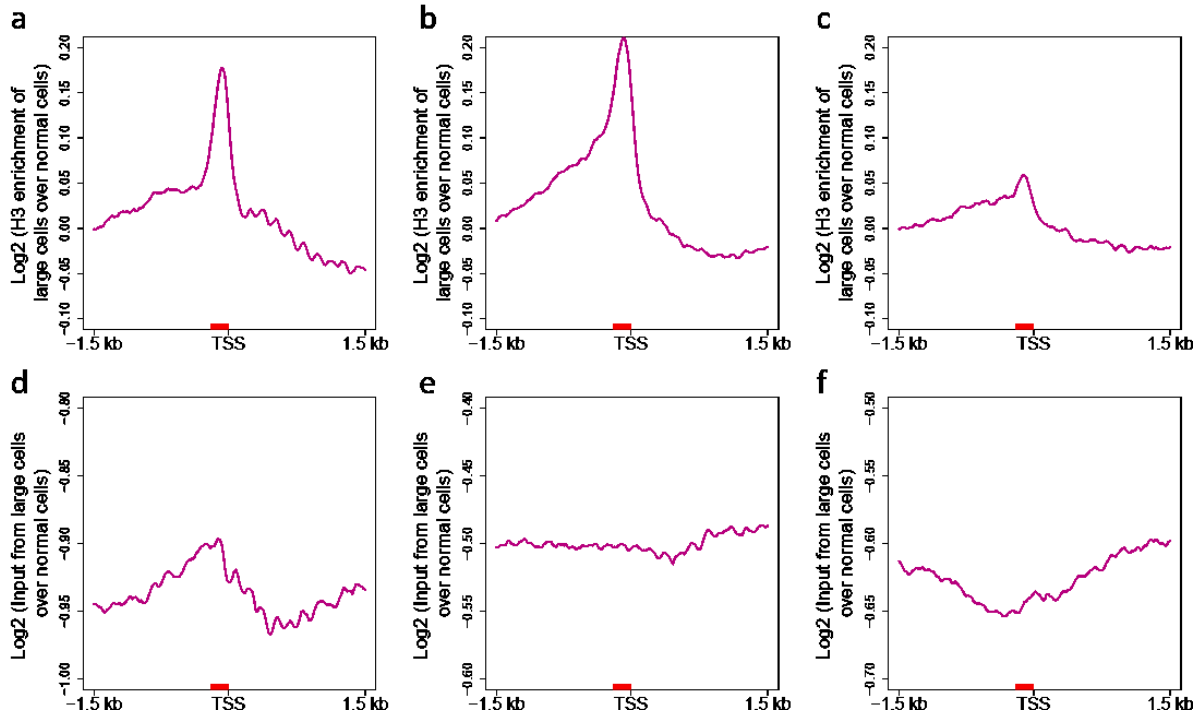
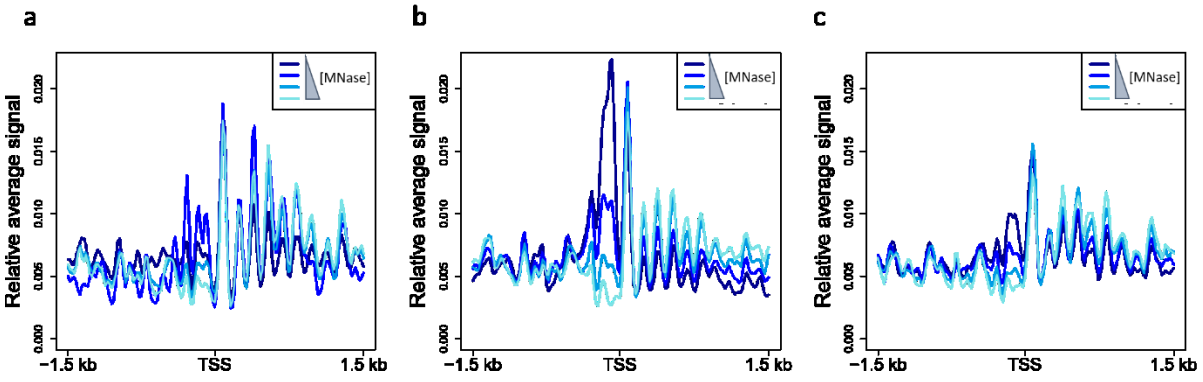


Figure 5.10: Hypothetical model showing the difference of chromatin architecture between two groups of genes and across different cell sizes. *The two groups of genes presented here are transcribed by the RNAPII. Other factors (co-activators having HAT activities, TFs, RSC, GTFs, GRFs, etc) intervene during transcription initiation (Haberle and Stark 2018, Rafal Donczew et al.,2020). The events driving transcription initiation depends on the availability of the RNAPII (Sun et al., 2020) and probably on the associated factors. Because of the size-dependant accumulation of the RNAPII and the associated factors, the frequency of recruitment of the transcription machinery is lower in a population of normal-sized cells compare to a population of large cells. Group 1 of genes are most coding genes and can be distinguished from group 2 by their promoter architecture. Group 2 of genes have the particularity to host a "fragile" nucleosome at their promoter. The factors able to disrupt the nucleosome to make it "fragile" are more numerous in large cells. Therefore, the presence of a "fragile" nucleosome at the promoter is more frequent in those large cells. Some genes contained in group 2 are associated with high RNA levels, suggesting that the corresponding promoter architecture might favour transcription activation.*

ANNEXES



Annex 5.1: ***A-c.*** *Average profile cantered on the TSS representing the log2 ratio of H3 enrichment from large cells over normal-sized cells for replicate 1* ***(a)****, 2* ***(b)*** *and 3* ***(c)****.* ***D-f.*** *Average profile cantered on the TSS representing the log2 ratio of input from large cells over normal-sized cells for replicate 1* ***(d)****, 2* ***(e)*** *and 3* ***(f).***



Annex 5.2: Average profile 1.5 kb on both side of the TSS for genes having a high H3 enrichment on promoter of large cells; *in normal sized-cells* ***(a),*** *large cells* ***(b)*** *and control cells* ***(c)*** *after an increasing MNase treatment.*

Chapter 6: General discussion

6.1. Summary

In this thesis, I investigated the link between chromatin structure and cell size. To this aim, I used the fission yeast *Schizosaccharomyces pombe* as its size is easily trackable being rod shape. Besides, I took advantage of the knowledge and the tools available in our group, such as the *cdc2-asM17* strain that can be arrested in G2 with an ATP analogue (1NM-PP1) and that reaches four times its normal size after 6h of 1NM-PP1 (Aoi et al., 2014). This fantastic model can thus provide different increasing sizes each hour after adding 1NM-PP1. To understand how the chromatin structure changes as cell size increases, I used the MNase-seq approach on a cell population of increasing cell size. It consists in digesting the DNA of the cell population with several increasing MNase concentrations. This technique led to the characterization of the proteins occupying differentially the DNA depending on the size of the cells. I then focused on the dynamics of histone regulation as a function of cell size with different approaches.

6.2. Conclusions

Five main conclusions arise from this dissertation:

- Cell elongation is associated with a substantial increase in MNase-sensitive protein occupancy at all promoters. The increase is dependent on the initial chromatin state of the gene. However, no change is witnessed on the gene body.

- Although such an increase in MNase-sensitive proteins occupancy is occurring for all promoters, the range of this increase depends on the genes and the genes that have a higher increase in occupancy also tend to be more expressed in basal conditions and while cell size increases.

- Chromatin structure at the promoter while cell size increase was not associated with differential gene expression while cell size increase. However, positively non-scaling genes

tend to be less expressed in basal condition and this is associated with a higher occupancy of MNase-sensitive as well as MNase-resistant proteins on the promoter.

- Histone proteins are subject to different regulations during cell size elongation: while H3 abundance mostly remains stable, H3K9ac and H2A abundance are differently regulated.

- Although H3 abundance appears stable, its genome-wide localization varies between normal-sized and large cells. More specifically, few genes exhibit a higher enrichment at the promoter of large cells compared to normal cells, which is enough to be seen in the average profile.

6.3. Outlook

Reflection on this work can be placed in the context of transcription scaling to cell size. In other words, to what extent could these changes cooperate with RNAPII to induce transcription scaling? It was previously demonstrated that the concentration of RNAPII in the nucleus was proportional to cell size (Swaffer et al., 2021b; Sun et al., 2020) and more importantly, that scaling of mRNA amount was mediated by a size-dependent increase in initiation rate. Therefore, one assumption is that the MNase-sensitive factors appearing on the promoter as cell size increase could be RNAPII itself. A ChIP-seq using an antibody against one subunit of the RNPAII would be needed to confirm this intuition in the elongating *cdc2-asM17*. An attempt at such an experiment was made in cells cultivated without 1NM-PP1 and cells cultivated 6h with 1NM-PP1 but, because of the poor quality of the data, the results were not conclusive (data not shown).

Furthermore, as reported in the introduction, the RNAPII does not operate alone to initiate transcription. It needs TFs and co-activators which can, in the case of chromatin remodelers and pioneer TF, destabilize locally the chromatin. Interestingly, those factors were previously detected on the promoter of the genes only after mild MNase digestion, showing that they are good candidates for the identification of the MNase-sensitive factors occupying the promoter

in a cell size-dependent manner (Brahma and Henikoff, 2020, 2019; Chereji et al., 2017; Kubik et al., 2015).

Finally, the research presented here identifies different sets of gene which exhibit different chromatin state on the promoters of larger cells, although an increase in MNase-sensitive factors operates for all genes. This might reflect different promoter architecture depending on the genes as cell size increases. The typical architecture might then participate in sustaining the mRNA concentration for each group of genes (Fig 6.1). This observation could be completed by the search of common DNA motif within each sets of genes.

Altogether, this study provides a groundwork to understand how transcription scaling occurs in cells of varying size.

6.4. Digression

MNase-seq data typically reports nucleosome organization on the average gene through an oscillating signal. This wave-like pattern can be characterized by several metrics like the height of wave (corresponding to nucleosome occupancy) or the spacing of the wave (NRL) (Fig 1.8). A similar wave-like representation can make us think of sound properties in the time domain. Indeed, in a sound representation, the height of the wave, or amplitude, defines the intensity of the sound. The spacing between two oscillations, the wavelength, across all the peaks is related to the frequency of the sound, namely if it is high or low pitch.

The crazy idea of transforming my MNase-seq data into sounds came during a discussion with Aubin Fleiss. It was just an idea until Leo Martin, engineer in CEA (Centre d'Energie Atomique) and musician, assured me it was possible and easy enough so he could do it.

After I provided him with the signal corresponding to the average profiles of large cells for the four MNase concentrations, he did an interpolation to increase the number of data points: 11025 instead of 600. One average profile indeed contained 150 values which corresponds to 1500 bp divided in 20 bp bins.

The data needed to be converted into .wav 16 bit so that it could give a sound. The conversion includes several steps: normalization of the values between -0.5 and 0.5, multiplication of the values by a factor so they become integer. The first part of the profile (the fuzzy nucleosomes) is excluded due to the low amplitude of the wave. Finally, each average profile, starting by the low MNase concentration to the highest MNase concentration, is repeated five times one after the others with a 0.5 second of silence between the profiles. Voila!

Chromatin architecture in large cells

Or: https://soundcloud.com/user-9243378/chromatin-architecture-in-large-cells

References

Allard, C.A.H., Opalko, H.E., Liu, K.-W., Medoh, U., and Moseley, J.B. (2018). Cell size–dependent regulation of Wee1 localization by Cdr2 cortical nodes. J. Cell Biol. *217*, 1589–1599.

Allard, C.A.H., Opalko, H.E., and Moseley, J.B. (2019). Stable Pom1 clusters form a glucose-modulated concentration gradient that regulates mitotic entry. ELife *8*, e46003.

Allegra, P., Sterner, R., Clayton, D.F., and Allfrey, V.G. (1987). Affinity chromatographic purification of nucleosomes containing transcriptionally active DNA sequences. J. Mol. Biol. *196*, 379–388.

Allshire, R.C., and Ekwall, K. (2015). Epigenetic Regulation of Chromatin States in Schizosaccharomyces pombe. Cold Spring Harb. Perspect. Biol. *7*, a018770.

Amodeo, A.A., and Skotheim, J.M. (2016). Cell-Size Control. Cold Spring Harb. Perspect. Biol. *8*, a019083.

Angermayr, M., Oechsner, U., Gregor, K., Schroth, G.P., and Bandlow, W. (2002). Transcription initiation in vivo without classical transactivators: DNA kinks flanking the core promoter of the housekeeping yeast adenylate kinase gene, AKY2, position nucleosomes and constitutively activate transcription. Nucleic Acids Res. *30*, 4199–4207.

Angermayr, M., Schwerdffeger, K., and Bandlow, W. (2003). A nucleosome-free dG-dC-rich sequence element promotes constitutive transcription of the essential yeast RIO1 gene. Biol. Chem. *384*, 1287–1292.

Aoi, Y., Kawashima, S.A., Simanis, V., Yamamoto, M., and Sato, M. (2014). Optimization of the analogue-sensitive Cdc2/Cdk1 mutant by in vivo selection eliminates physiological limitations to its use in cell cycle analysis. Open Biol. *4*, 140063.

Ashworth, W., Stoney, P.N., and Yamamoto, T. (2019). States of decay: The systems biology of mRNA stability. Curr. Opin. Syst. Biol. *15*, 48–57.

Backman, S.A., Stambolic, V., Suzuki, A., Haight, J., Elia, A., Pretorius, J., Tsao, M.-S., Shannon, P., Bolon, B., Ivy, G.O., et al. (2001). Deletion of Pten in mouse brain causes seizures, ataxia and defects in soma size resembling Lhermitte-Duclos disease. Nat. Genet. *29*, 396–403.

Bahar Halpern, K., Caspi, I., Lemze, D., Levy, M., Landen, S., Elinav, E., Ulitsky, I., and Itzkovitz, S. (2015). Nuclear Retention of mRNA in Mammalian Tissues. Cell Rep. *13*, 2653–2662.

Bähler, J., and Pringle, J.R. (1998). Pom1p, a fission yeast protein kinase that provides positional information for both polarized growth and cytokinesis. Genes Dev. *12*, 1356–1370.

Bai, L., and Morozov, A.V. (2010). Gene regulation by nucleosome positioning. Trends Genet. *26*, 476–483.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. *37*, W202–W208.

Bakel, H. van, Tsui, K., Gebbia, M., Mnaimneh, S., Hughes, T.R., and Nislow, C. (2013). A Compendium of Nucleosome and Transcript Profiles Reveals Determinants of Chromatin Architecture and Transcription. PLOS Genet. *9*, e1003479.

Baldi, S., Korber, P., and Becker, P.B. (2020). Beads on a string—nucleosome array arrangements and folding of the chromatin fiber. Nat. Struct. Mol. Biol. *27*, 109–118.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Baptista, T., Grünberg, S., Minoungou, N., Koster, M.J.E., Timmers, H.T.M., Hahn, S., Devys, D., and Tora, L. (2017). SAGA Is a General Cofactor for RNA Polymerase II Transcription. Mol. Cell *68*, 130-143.e5.

Barnum, K.J., and O'Connell, M.J. (2014). Cell Cycle Regulation by Checkpoints. In Cell Cycle Control, E. Noguchi, and M.C. Gadaleta, eds. (New York, NY: Springer New York), pp. 29–40.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-Resolution Profiling of Histone Methylations in the Human Genome. Cell *129*, 823–837.

Basehoar, A.D., Zanton, S.J., and Pugh, B.F. (2004). Identification and distinct regulation of yeast TATA box-containing genes. Cell *116*, 699–709.

Battich, N., Stoeger, T., and Pelkmans, L. (2015). Control of Transcript Variability in Single Mammalian Cells. Cell *163*, 1596–1610.

Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407–412.

Berry, S., Müller, M., and Pelkmans, L. (2021). Nuclear RNA concentration coordinates RNA production with cell size in human cells (Molecular Biology).

Billon, P., and Côté, J. (2012). Precise deposition of histone H2A.Z in chromatin for genome expression and maintenance. Biochim. Biophys. Acta BBA - Gene Regul. Mech. *1819*, 290–302.

Bitton, D.A., Schubert, F., Dey, S., Okoniewski, M., Smith, G.C., Khadayate, S., Pancaldi, V., Wood, V., and Bähler, J. (2015). AnGeLi: A Tool for the Analysis of Gene Lists from Fission Yeast. Front. Genet. *6*, 330.

Black, J.C., Choi, J.E., Lombardo, S.R., and Carey, M. (2006). A Mechanism for Coordinating Chromatin Modification and Preinitiation Complex Assembly. Mol. Cell *23*, 809–818.

Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., Furey, T.S., and Crawford, G.E. (2008). High-resolution mapping and characterization of open chromatin across the genome. Cell *132*, 311–322.

Brahma, S., and Henikoff, S. (2019). RSC-Associated Subnucleosomes Define MNase-Sensitive Promoters in Yeast. Mol. Cell *73*, 238-249.e3.

Brahma, S., and Henikoff, S. (2020). Epigenome Regulation by Dynamic Nucleosome Unwrapping. Trends Biochem. Sci. *45*, 13–26.

Bryant, G.O., Prabhu, V., Floer, M., Wang, X., Spagna, D., Schreiber, D., and Ptashne, M. (2008). Activator Control of Nucleosome Occupancy in Activation and Repression of Transcription. PLOS Biol. *6*, e317.

Cadart, C., Monnier, S., Grilli, J., Sáez, P.J., Srivastava, N., Attia, R., Terriac, E., Baum, B., Cosentino-Lagomarsino, M., and Piel, M. (2018). Size control in mammalian cells involves modulation of both growth rate and cell cycle duration. Nat. Commun. *9*, 3275.

Cai, L., Sutter, B.M., Li, B., and Tu, B.P. (2011). Acetyl-CoA Induces Cell Growth and Proliferation by Promoting the Acetylation of Histones at Growth Genes. Mol. Cell *42*, 426–437.

Calder, W.A. (1996). Size, Function, and Life History (Courier Corporation).

Campos, H., Boeing, W.J., Dungan, B.N., and Schaub, T. (2014a). Cultivating the marine microalga Nannochloropsis salina under various nitrogen sources: Effect on biovolume yields, lipid content and composition, and invasive organisms. Biomass Bioenergy *66*, 301–307.

Campos, M., Surovtsev, I.V., Kato, S., Paintdakhi, A., Beltran, B., Ebmeier, S.E., and Jacobs-Wagner, C. (2014b). A constant size extension drives bacterial cell size homeostasis. Cell *159*, 1433–1446.

Challal, D., Barucco, M., Kubik, S., Feuerbach, F., Candelli, T., Geoffroy, H., Benaksas, C., Shore, D., and Libri, D. (2018). General Regulatory Factors Control the Fidelity of Transcription by Restricting Non-coding and Ectopic Initiation. Mol. Cell *72*, 955-969.e7.

Chan, J.C., and Maze, I. (2020). Nothing Is Yet Set in (Hi)stone: Novel Post-Translational Modifications Regulating Chromatin Function. Trends Biochem. Sci. *45*, 829–844.

Chen, X., Qi, Y., Wu, Z., Wang, X., Li, J., Zhao, D., Hou, H., Li, Y., Yu, Z., Liu, W., et al. (2021). Structural insights into preinitiation complex assembly on core promoters. Science.

Chen, Y., Zhao, G., Zahumensky, J., Honey, S., and Futcher, B. (2020). Differential Scaling of Gene Expression with Cell Size May Explain Size Control in Budding Yeast. Mol. Cell *78*, 359-370.e6.

Chereji, R.V., Kan, T.-W., Grudniewska, M.K., Romashchenko, A.V., Berezikov, E., Zhimulev, I.F., Guryev, V., Morozov, A.V., and Moshkin, Y.M. (2016). Genome-wide profiling of

nucleosome sensitivity and chromatin accessibility in Drosophila melanogaster. Nucleic Acids Res. *44*, 1036–1051.

Chereji, R.V., Ocampo, J., and Clark, D.J. (2017). MNase-Sensitive Complexes in Yeast: Nucleosomes and Non-histone Barriers. Mol. Cell *65*, 565-577.e3.

Chereji, R.V., Bryson, T.D., and Henikoff, S. (2019). Quantitative MNase-seq accurately maps nucleosome occupancy levels. Genome Biol. *20*, 198.

Chung, H.-R., Dunkel, I., Heise, F., Linke, C., Krobitsch, S., Ehrenhofer-Murray, A.E., Sperling, S.R., and Vingron, M. (2010). The Effect of Micrococcal Nuclease Digestion on Nucleosome Positioning Data. PLOS ONE *5*, e15754.

Clapier, C.R., Iwasa, J., Cairns, B.R., and Peterson, C.L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nat. Rev. Mol. Cell Biol. *18*, 407–422.

Claude, K.-L., Bureik, D., Chatzitheodoridou, D., Adarska, P., Singh, A., and Schmoller, K.M. (2021). Transcription coordinates histone amounts and genome content. Nat. Commun. *12*, 4202.

Conlon, I., and Raff, M. (1999). Size Control in Animal Development. Cell *96*, 235–244.

Conlon, I., and Raff, M. (2003). Differences in the way a mammalian cell and yeast cells coordinate cell growth and cell-cycle progression. J. Biol. *2*, 7.

Costanzo, M., Nishikawa, J.L., Tang, X., Millman, J.S., Schub, O., Breitkreuz, K., Dewar, D., Rupes, I., Andrews, B., and Tyers, M. (2004). CDK Activity Antagonizes Whi5, an Inhibitor of G1/S Transcription in Yeast. Cell *117*, 899–913.

Coudreuse, D., and Nurse, P. (2010). Driving the cell cycle with a minimal CDK control network. Nature *468*, 1074–1079.

DeGennaro, C.M., Alver, B.H., Marguerat, S., Stepanova, E., Davis, C.P., Bähler, J., Park, P.J., and Winston, F. (2013). Spt6 Regulates Intragenic and Antisense Transcription, Nucleosome Positioning, and Histone Modifications Genome-Wide in Fission Yeast. Mol. Cell. Biol. *33*, 4779–4792.

Denny, M.W., Daniel, T.L., and Koehl, M. a. R. (1985). Mechanical Limits to Size in Wave-Swept Organisms. Ecol. Monogr. *55*, 69–102.

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., Zhou, M.-M., and Zhou, M.-M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. Nature *399*, 491–496.

Dischinger, S., Krapp, A., Xie, L., Paulson, J.R., and Simanis, V. (2008). Chemical genetic analysis of the regulatory role of Cdc2p in the S. pombe septation initiation network. J. Cell Sci. *121*, 843–853.

Donachie, W.D. (1968). Relationship between Cell Size and Time of Initiation of DNA Replication. Nature *219*, 1077–1079.

Donczew, R., Warfield, L., Pacheco, D., Erijman, A., and Hahn, S. (2020). Two roles for the yeast transcription coactivator SAGA and a set of genes redundantly regulated by TFIID and SAGA. ELife *9*, e50109.

Dorsey, S., Tollis, S., Cheng, J., Black, L., Notley, S., Tyers, M., and Royer, C.A. (2018). G1/S Transcription Factor Copy Number Is a Growth-Dependent Determinant of Cell Cycle Commitment in Yeast. Cell Syst. *6*, 539-554.e11.

Elliott, S.G. (1983). Coordination of growth with cell division: Regulation of synthesis of RNA during the cell cycle of the fission yeast Schizosaccharomyces pombe. Mol. Gen. Genet. MGG *192*, 204–211.

Elliott, S.G., and McLaughlin, C.S. (1978). Rate of macromolecular synthesis through the cell cycle of the yeast Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *75*, 4384–4388.

Eriksson, P.R., Ganguli, D., Nagarajavel, V., and Clark, D.J. (2012). Regulation of histone gene expression in budding yeast. Genetics *191*, 7–20.

Evrin, C., Serra-Cardona, A., Duan, S., Mukherjee, P.P., Zhang, Z., and Labib, K.P.M. (2022). Spt5 histone binding activity preserves chromatin during transcription by RNA polymerase II. EMBO J. e109783.

Facchetti, G., Chang, F., and Howard, M. (2017). Controlling cell size through sizer mechanisms. Curr. Opin. Syst. Biol. *5*, 86–92.

Facchetti, G., Knapp, B., Flor-Parra, I., Chang, F., and Howard, M. (2019). Reprogramming Cdr2-Dependent Geometry-Based Cell Size Control in Fission Yeast. Curr. Biol. *29*, 350-358.e4.

Fantes, P.A. (1977). Control of cell size and cycle time in Schizosaccharomyces pombe. J. Cell Sci. *24*, 51–67.

Fantes, P., and Nurse, P. (1977). Control of cell size at division in fission yeast by a growth-modulated size control over nuclear division. Exp. Cell Res. *107*, 377–386.

Farnung, L., Ochmann, M., Engeholm, M., and Cramer, P. (2021). Structural basis of nucleosome transcription mediated by Chd1 and FACT. Nat. Struct. Mol. Biol. *28*, 382–387.

Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.-P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Müller, S., Pawson, T., et al. (2012). Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. Cell *149*, 214–231.

Floer, M., Wang, X., Prabhu, V., Berrozpe, G., Narayan, S., Spagna, D., Alvarez, D., Kendall, J., Krasnitz, A., Stepansky, A., et al. (2010). A RSC/Nucleosome Complex Determines Chromatin Architecture and Facilitates Activator Binding. Cell *141*, 407–418.

Foster, J.B. (1964). Evolution of Mammals on Islands. Nature *202*, 234–235.

Fraser, R.S., and Nurse, P. (1979). Altered patterns of ribonucleic acid synthesis during the cell cycle: a mechanism compensating for variation in gene concentration. J. Cell Sci. *35*, 25–40.

Fraser, R.S.S., and Nurse, P. (1978). Novel cell cycle control of RNA synthesis in yeast. Nature *271*, 726–730.

Galitski, T., Saldanha, A.J., Styles, C.A., Lander, E.S., and Fink, G.R. (1999). Ploidy Regulation of Gene Expression. Science *285*, 251–254.

Ganguli, D., Chereji, R.V., Iben, J.R., Cole, H.A., and Clark, D.J. (2014). RSC-dependent constructive and destructive interference between opposing arrays of phased nucleosomes in yeast. Genome Res. *24*, 1637–1649.

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. *8*, 113–126.

Gates, L.A., Shi, J., Rohira, A.D., Feng, Q., Zhu, B., Bedford, M.T., Sagum, C.A., Jung, S.Y., Qin, J., Tsai, M.-J., et al. (2017). Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation. J. Biol. Chem. *292*, 14456–14472.

Gerganova, V., Floderer, C., Archetti, A., Michon, L., Carlini, L., Reichler, T., Manley, S., and Martin, S.G. (2019). Multi-phosphorylation reaction and clustering tune Pom1 gradient mid-cell levels according to cell size. ELife *8*, e45983.

Giaimo, B.D., Ferrante, F., Herchenröther, A., Hake, S.B., and Borggrefe, T. (2019). The histone variant H2A.Z in gene regulation. Epigenetics Chromatin *12*, 37.

Givens, R.M., Lai, W.K.M., Rizzo, J.M., Bard, J.E., Mieczkowski, P.A., Leatherwood, J., Huberman, J.A., and Buck, M.J. (2012). Chromatin architectures at fission yeast transcriptional promoters and replication origins. Nucleic Acids Res. *40*, 7176–7189.

Godde, J.S., and Widom, J. (1992). Chromatin structure of Schizosaccharomyces pombe. A nucleosome repeat length that is shorter than the chromatosomal DNA length. J. Mol. Biol. *226*, 1009–1025.

Godin, M., Delgado, F.F., Son, S., Grover, W.H., Bryan, A.K., Tzur, A., Jorgensen, P., Payer, K., Grossman, A.D., Kirschner, M.W., et al. (2010). Using buoyant mass to measure the growth of single cells. Nat. Methods *7*, 387–390.

Golding, I., Paulsson, J., Zawilski, S.M., and Cox, E.C. (2005). Real-Time Kinetics of Gene Activity in Individual Bacteria. Cell *123*, 1025–1036.

Gómez, E.B., and Forsburg, S.L. (2004). Analysis of the fission yeast Schizosaccharomyces pombe cell cycle. Methods Mol. Biol. Clifton NJ *241*, 93–111.

Gossett, A.J., and Lieb, J.D. (2012). In Vivo Effects of Histone H3 Depletion on Nucleosome Occupancy and Position in Saccharomyces cerevisiae. PLOS Genet. *8*, e1002771.

Gould, K.L., and Nurse, P. (1989). Tyrosine phosphorylation of the fission yeast cdc2+ protein kinase regulates entry into mitosis. Nature *342*, 39–45.

Grant, P.A., Duggan, L., Côté, J., Roberts, S.M., Brownell, J.E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C.D., Winston, F., et al. (1997). Yeast Gcn5 functions in two multisubunit

complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. Genes Dev. *11*, 1640–1650.

Grant, P.A., Eberharter, A., John, S., Cook, R.G., Turner, B.M., and Workman, J.L. (1999). Expanded Lysine Acetylation Specificity of Gcn5 in Native Complexes*. J. Biol. Chem. *274*, 5895–5900.

Gu, Y., and Oliferenko, S. (2021). The principles of cellular geometry scaling. Curr. Opin. Cell Biol. *68*, 20–27.

Haberle, V., and Stark, A. (2018). Eukaryotic core promoters and the functional basis of transcription initiation. Nat. Rev. Mol. Cell Biol. *19*, 621–637.

Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. Cell *55*, 1137–1145.

Hansson, K.-A., Eftestøl, E., Bruusgaard, J.C., Juvkam, I., Cramer, A.W., Malthe-Sørenssen, A., Millay, D.P., and Gundersen, K. (2020). Myonuclear content regulates cell size with similar scaling properties in mice and humans. Nat. Commun. *11*, 6288.

Hartwell, L.H. (1974). Saccharomyces cerevisiae cell cycle. BACTERIOL REV *38*, 35.

Hebbes, T.R., Thorne, A.W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J. *7*, 1395–1402.

Heldt, F.S., Lunstone, R., Tyson, J.J., and Novák, B. (2018). Dilution and titration of cell-cycle regulators may control cell size in budding yeast. PLOS Comput. Biol. *14*, e1006548.

Henikoff, S., Henikoff, J.G., Sakai, A., Loeb, G.B., and Ahmad, K. (2009). Genome-wide profiling of salt fractions maps physical properties of chromatin. Genome Res. *19*, 460–469.

Hewish, D.R., and Burgoyne, L.A. (1973). Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. Biochem. Biophys. Res. Commun. *52*, 504–510.

Hirota, K., Miyoshi, T., Kugou, K., Hoffman, C.S., Shibata, T., and Ohta, K. (2008). Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. Nature *456*, 130–134.

Hirschhorn, J.N., Brown, S.A., Clark, C.D., and Winston, F. (1992). Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. Genes Dev. *6*, 2288–2298.

Hocher, A., Rojec, M., Swadling, J.B., Esin, A., and Warnecke, T. (2019). The DNA-binding protein HTa from Thermoplasma acidophilum is an archaeal histone analog.

Hong, L., Schroth, G.P., Matthews, H.R., Yau, P., and Bradbury, E.M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 “tail” to DNA. J. Biol. Chem. *268*, 305–314.

Hörz, W., and Altenburger, W. (1981). Sequence specific cleavage of DNA by micrococcal nuclease. Nucleic Acids Res. *9*, 2643–2658.

Hu, S., Chen, X., Liao, J., Chen, Y., Zhao, C., and Zhang, Y. (2017). CAM: A quality control pipeline for MNase-seq data. PLOS ONE *12*, e0182771.

Huang, D., Kaluarachchi, S., Dyk, D. van, Friesen, H., Sopko, R., Ye, W., Bastajian, N., Moffat, J., Sassi, H., Costanzo, M., et al. (2009). Dual Regulation by Pairs of Cyclin-Dependent Protein Kinases and Histone Deacetylases Controls G1 Transcription in Budding Yeast. PLOS Biol. *7*, e1000188.

Huisinga, K.L., and Pugh, B.F. (2004). A Genome-Wide Housekeeping Role for TFIID and a Highly Regulated Stress-Related Role for SAGA in Saccharomyces cerevisiae. Mol. Cell *13*, 573–585.

Ietswaart, R., Rosa, S., Wu, Z., Dean, C., and Howard, M. (2017). Cell-Size-Dependent Transcription of FLC and Its Antisense Long Non-coding RNA COOLAIR Explain Cell-to-Cell Expression Variation. Cell Syst. *4*, 622-635.e9.

Ioshikhes, I.P., Albert, I., Zanton, S.J., and Pugh, B.F. (2006). Nucleosome positions predicted through comparative genomics. Nat. Genet. *38*, 1210–1215.

Iwafuchi-Doi, M. (2019). The mechanistic basis for chromatin regulation by pioneer transcription factors. WIREs Syst. Biol. Med. *11*, e1427.

Haldane, J.B.S. (1928). On being the right size. In Possible Worlds, pp.18-26. London: Chatto and Windus.

Jacobson, R.H., Ladurner, A.G., King, D.S., and Tjian, R. (2000). Structure and Function of a Human TAFII250 Double Bromodomain Module. Science *288*, 1422–1425.

Jain, D., Nemec, S., Luxey, M., Gauthier, Y., Bemmo, A., Balsalobre, A., and Drouin, J. (2018). Regulatory integration of Hox factor activity with T-box factors in limb development. Development *145*, dev159830.

Jang, M.K., Mochizuki, K., Zhou, M., Jeong, H.-S., Brady, J.N., and Ozato, K. (2005). The Bromodomain Protein Brd4 Is a Positive Regulatory Component of P-TEFb and Stimulates RNA Polymerase II-Dependent Transcription. Mol. Cell *19*, 523–534.

Jeffers, T.E., and Lieb, J.D. (2017). Nucleosome fragility is associated with future transcriptional response to developmental cues and stress in C. elegans. Genome Res. *27*, 75–86.

Jenuwein, T., and Allis, C.D. (2001). Translating the Histone Code. Science *293*, 1074–1080.

Jeronimo, C., Watanabe, S., Kaplan, C.D., Peterson, C.L., and Robert, F. (2015). The Histone Chaperones FACT and Spt6 Restrict H2A.Z from Intragenic Locations. Mol. Cell *58*, 1113–1123.

Jiang, Z., and Zhang, B. (2021). On the role of transcription in positioning nucleosomes. PLOS Comput. Biol. *17*, e1008556.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark “nucleosome-free regions” of active promoters and other regulatory regions. Nat. Genet. *41*, 941–945.

Johnson, A., and Skotheim, J.M. (2013). Start and the Restriction Point. Curr. Opin. Cell Biol. *25*, 10.1016/j.ceb.2013.07.010.

Johnsson, A. (2009). Characterization of Gcn5 Histone Acetyltransferase in Schizosaccharomyces Pombe. Ph.D. Karolinska Institutet (Sweden).

Johnston, G.C., Pringle, J.R., and Hartwell, L.H. (1977). Coordination of growth with cell division in the yeast Saccharomyces cerevisiae. Exp. Cell Res. *105*, 79–98.

Jones, A.R., Band, L.R., and Murray, J.A.H. (2019). Double or Nothing? Cell Division and Cell Size Control. Trends Plant Sci. *24*, 1083–1093.

Jorgensen, P., and Tyers, M. (2004). How Cells Coordinate Growth and Division. Curr. Biol. *14*, R1014–R1027.

Jun, S., and Taheri-Araghi, S. (2015). Cell-size maintenance: universal strategy revealed. Trends Microbiol. *23*, 4–6.

Jung, Y.H., Sauria, M.E.G., Lyu, X., Cheema, M.S., Ausio, J., Taylor, J., and Corces, V.G. (2017). Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes. Cell Rep. *18*, 1366–1382.

Juven-Gershon, T., and Kadonaga, J.T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev. Biol. *339*, 225–229.

Kakui, Y., Barrington, C., Barry, D.J., Gerguri, T., Fu, X., Bates, P.A., Khatri, B.S., and Uhlmann, F. (2020). Fission yeast condensin contributes to interphase chromatin organization and prevents transcription-coupled DNA damage. Genome Biol. *21*, 272.

Kanno, T., Kanno, Y., Siegel, R.M., Jang, M.K., Lenardo, M.J., and Ozato, K. (2004). Selective Recognition of Acetylated Histones by Bromodomain Proteins Visualized in Living Cells. Mol. Cell *13*, 33–43.

Kaplan, C.D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. Science *301*, 1096–1099.

Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. BMC Genomics *13*, 424.

Keifenheim, D., Sun, X.-M., D’Souza, E., Ohira, M.J., Magner, M., Mayhew, M.B., Marguerat, S., and Rhind, N. (2017). Size-Dependent Expression of the Mitotic Activator Cdc25 Suggests a Mechanism of Size Control in Fission Yeast. Curr. Biol. *27*, 1491-1497.e4.

Kelly, T.K., Liu, Y., Lay, F.D., Liang, G., Berman, B.P., and Jones, P.A. (2012). Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. Genome Res. *22*, 2497–2506.

Kempe, H., Schwabe, A., Crémazy, F., Verschure, P.J., and Bruggeman, F.J. (2015). The volumes and transcript counts of single cells reveal concentration homeostasis and capture biological noise. Mol. Biol. Cell *26*, 797–804.

Kepler, T.B., and Elston, T.C. (2001). Stochasticity in Transcriptional Regulation: Origins, Consequences, and Mathematical Representations. Biophys. J. *81*, 3116–3136.

Kleijn, I.T., Martínez-Segura, A., Bertaux, F., Saint, M., Kramer, H., Shahrezaei, V., and Marguerat, S. (2021). Growth-Rate Dependent And Nutrient-Specific Gene Expression Resource Allocation In Fission Yeast.

Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. Nat. Rev. Genet. *20*, 207–220.

Knapp, B.D., Odermatt, P., Rojas, E.R., Cheng, W., He, X., Huang, K.C., and Chang, F. (2019). Decoupling of Rates of Protein Synthesis from Cell Expansion Leads to Supergrowth. Cell Syst. *9*, 434-445.e6.

Knezetic, J.A., and Luse, D.S. (1986). The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. Cell *45*, 95–104.

Knight, B., Kubik, S., Ghosh, B., Bruzzone, M.J., Geertz, M., Martin, V., Dénervaud, N., Jacquet, P., Ozkan, B., Rougemont, J., et al. (2014). Two distinct promoter architectures centered on dynamic nucleosomes control ribosomal protein gene transcription. Genes Dev. *28*, 1695–1709.

Koch, A.L., and Schaechter, M.Y. 1962 A Model for Statistics of the Cell Division Process. Microbiology *29*, 435–454.

Koch, C., Schleiffer, A., Ammerer, G., and Nasmyth, K. (1996). Switching transcription on and off during the yeast cell cycle: Cln/Cdc28 kinases activate bound transcription factor SBF (Swi4/Swi6) at start, whereas Clb/Cdc28 kinases displace it from the promoter in G2. Genes Dev. *10*, 129–141.

Korber, P., and Barbaric, S. (2014). The yeast PHO5 promoter: from single locus to systems biology of a paradigm for gene regulation through chromatin. Nucleic Acids Res. *42*, 10888–10902.

Korber, P., and Hörz, W. (2004). In Vitro Assembly of the Characteristic Chromatin Organization at the Yeast PHO5 Promoter by a Replication-independent Extract System *. J. Biol. Chem. *279*, 35113–35120.

Kornberg, R.D. (1974). Chromatin Structure: A Repeating Unit of Histones and DNA. Science *184*, 868–871.

Kornberg, R.D., and Lorch, Y. (1999). Twenty-Five Years of the Nucleosome, Fundamental Particle of the Eukaryote Chromosome. Cell *98*, 285–294.

Kornberg, R.D., and Lorch, Y. (2020). Primary Role of the Nucleosome. Mol. Cell *79*, 371–375.

Kratz, A., Arner, E., Saito, R., Kubosaki, A., Kawai, J., Suzuki, H., Carninci, P., Arakawa, T., Tomita, M., Hayashizaki, Y., et al. (2010). Core promoter structure and genomic context reflect histone 3 lysine 9 acetylation patterns. BMC Genomics *11*, 257.

Krebs, A.R., Imanci, D., Hoerner, L., Gaidatzis, D., Burger, L., and Schübeler, D. (2017). Genome-wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters. Mol. Cell *67*, 411-422.e4.

Krietenstein, N., Wal, M., Watanabe, S., Park, B., Peterson, C.L., Pugh, B.F., and Korber, P. (2016). Genomic Nucleosome Organization Reconstituted with Pure Proteins. Cell *167*, 709-721.e12.

Kristjuhan, A., Walker, J., Suka, N., Grunstein, M., Roberts, D., Cairns, B.R., and Svejstrup, J.Q. (2002). Transcriptional Inhibition of Genes with Severe Histone H3 Hypoacetylation in the Coding Region. Mol. Cell *10*, 925–933.

Kubik, S., Bruzzone, M.J., Jacquet, P., Falcone, J.-L., Rougemont, J., and Shore, D. (2015). Nucleosome Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast. Mol. Cell *60*, 422–434.

Kubik, S., Bruzzone, M.J., Albert, B., and Shore, D. (2017). A Reply to “MNase-Sensitive Complexes in Yeast: Nucleosomes and Non-histone Barriers,” by Chereji et al. Mol. Cell *65*, 578–580.

Kubik, S., O’Duibhir, E., de Jonge, W.J., Mattarocci, S., Albert, B., Falcone, J.-L., Bruzzone, M.J., Holstege, F.C.P., and Shore, D. (2018). Sequence-Directed Action of RSC Remodeler and General Regulatory Factors Modulates +1 Nucleosome Position to Facilitate Transcription. Mol. Cell *71*, 89-102.e5.

Kulaeva, O.I., and Studitsky, V.M. (2010). Mechanism of histone survival during transcription by RNA polymerase II. Transcription *1*, 85–88.

Kulaeva, O.I., Gaykalova, D.A., Pestov, N.A., Golovastov, V.V., Vassylyev, D.G., Artsimovitch, I., and Studitsky, V.M. (2009). Mechanism of chromatin remodeling and recovery during passage of RNA polymerase II. Nat. Struct. Mol. Biol. *16*, 1272–1278.

Kwon, C.-H., Zhu, X., Zhang, J., Knoop, L.L., Tharp, R., Smeyne, R.J., Eberhart, C.G., Burger, P.C., and Baker, S.J. (2001). Pten regulates neuronal soma size: a mouse model of Lhermitte-Duclos disease. Nat. Genet. *29*, 404–411.

Labib, K., and Moreno, S. (1996). rum1: a CDK inhibitor regulating G1 progression in fission yeast. Trends Cell Biol. *6*, 62–66.

Lai, B., Gao, W., Cui, K., Xie, W., Tang, Q., Jin, W., Hu, G., Ni, B., and Zhao, K. (2018). Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. Nature *562*, 281–285.

Lantermann, A., Strålfors, A., Fagerström-Billai, F., Korber, P., and Ekwall, K. (2009). Genome-wide mapping of nucleosome positions in Schizosaccharomyces pombe. Methods *48*, 218–225.

Lantermann, A.B., Straub, T., Strålfors, A., Yuan, G.-C., Ekwall, K., and Korber, P. (2010). Schizosaccharomyces pombe genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of Saccharomyces cerevisiae. Nat. Struct. Mol. Biol. *17*, 251–257.

Lanz, M.C., Zatulovskiy, E., Swaffer, M.P., Zhang, L., Zhang, S., You, D.S., Marinov, G., McAlpine, P., Elias, J.E., and Skotheim, J.M. (2021). Increasing cell size remodels the proteome and promotes senescence (Cell Biology).

Lee, C.-K., Shibata, Y., Rao, B., Strahl, B.D., and Lieb, J.D. (2004). Evidence for nucleosome depletion at active regulatory regions genome-wide. Nat. Genet. *36*, 900–905.

Lee, D.Y., Hayes, J.J., Pruss, D., and Wolffe, A.P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell *72*, 73–84.

Lee, W., Tillo, D., Bray, N., Morse, R.H., Davis, R.W., Hughes, T.R., and Nislow, C. (2007). A high-resolution atlas of nucleosome occupancy in yeast. Nat. Genet. *39*, 1235–1244.

Lenhard, B., Sandelin, A., and Carninci, P. (2012). Metazoan promoters: emerging characteristics and insights into transcriptional regulation. Nat. Rev. Genet. *13*, 233–245.

Lenski, R.E., and Travisano, M. (1994). Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc. Natl. Acad. Sci. *91*, 6808–6814.

Li, B., Carey, M., and Workman, J.L. (2007). The Role of Chromatin during Transcription. Cell *128*, 707–719.

Li, G., Liu, S., Wang, J., He, J., Huang, H., Zhang, Y., and Xu, L. (2014). ISWI proteins participate in the genome-wide nucleosome distribution in Arabidopsis. Plant J. *78*, 706–714.

Li, H., Hou, J., Bai, L., Hu, C., Tong, P., Kang, Y., Zhao, X., and Shao, Z. (2015a). Genome-wide analysis of core promoter structures in Schizosaccharomyces pombe with DeepCAGE. RNA Biol. *12*, 525–537.

Li, Q., Rycaj, K., Chen, X., and Tang, D.G. (2015b). Cancer stem cells and cell size: A causal link? Semin. Cancer Biol. *35*, 191–199.

Lickwar, C.R., Mueller, F., Hanlon, S.E., McNally, J.G., and Lieb, J.D. (2012). Genome-wide protein–DNA binding dynamics suggest a molecular clutch for transcription factor function. Nature *484*, 251–255.

Lieleg, C., Krietenstein, N., Walker, M., and Korber, P. (2015). Nucleosome positioning in yeasts: methods, maps, and mechanisms. Chromosoma *124*, 131–151.

Lin, J., and Amir, A. (2018). Homeostasis of protein and mRNA concentrations in growing cells. Nat. Commun. *9*, 4496.

Lin, J., and Wang, Q. (2021). Heterogeneous recruitment abilities to RNA polymerases generate nonlinear scaling of gene expression level with cell volume.

Lin, J.C., Jeong, S., Liang, G., Takai, D., Fatemi, M., Tsai, Y.C., Egger, G., Gal-Yam, E.N., and Jones, P.A. (2007). Role of Nucleosomal Occupancy in the Epigenetic Silencing of the MLH1 CpG Island. Cancer Cell *12*, 432–444.

Litsios, A., Huberts, D.H.E.W., Terpstra, H.M., Guerra, P., Schmidt, A., Buczak, K., Papagiannakis, A., Rovetta, M., Hekelaar, J., Hubmann, G., et al. (2019). Differential scaling between G1 protein production and cell size dynamics promotes commitment to the cell division cycle in budding yeast. Nat. Cell Biol. *21*, 1382–1392.

Liu, M.-J., Seddon, A.E., Tsai, Z.T.-Y., Major, I.T., Floer, M., Howe, G.A., and Shiu, S.-H. (2015). Determinants of nucleosome positioning and their influence on plant gene expression. Genome Res. *25*, 1182–1195.

Lock, A., Rutherford, K., Harris, M.A., Hayles, J., Oliver, S.G., Bähler, J., and Wood, V. (2019). PomBase 2018: user-driven reimplementation of the fission yeast database provides rapid and intuitive access to diverse, interconnected information. Nucleic Acids Res. *47*, D821–D827.

Lorch, Y., LaPointe, J.W., and Kornberg, R.D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell *49*, 203–210.

Lord, P.G., and Wheals, A.E. (1983). Rate of cell cycle initiation of yeast cells when cell size is not a rate-determining factor. J. Cell Sci. *59*, 183–201.

Louder, R.K., He, Y., López-Blanco, J.R., Fang, J., Chacón, P., and Nogales, E. (2016). Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. Nature *531*, 604–609.

Lozoya, O.A., Wang, T., Grenet, D., Wolfgang, T.C., Sobhany, M., Silva, D.G. da, Riadi, G., Chandel, N., Woychik, R.P., and Santos, J.H. (2019). Mitochondrial acetyl-CoA reversibly regulates locus-specific histone acetylation and gene expression. Life Sci. Alliance *2*.

Luzzati, V., and Nicolaïeff, A. (1963). The structure of nucleohistones and nucleoprotamines. J. Mol. Biol. *7*, 142–163.

Marguerat, S., and Bähler, J. (2012). Coordinating genome expression with cell size. Trends Genet. *28*, 560–565.

Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R., and Bähler, J. (2012). Quantitative Analysis of Fission Yeast Transcriptomes and Proteomes in Proliferating and Quiescent Cells. Cell *151*, 671–683.

Martin, S.G., and Berthelot-Grosjean, M. (2009). Polar gradients of the DYRK-family kinase Pom1 couple cell length with the cell cycle. Nature *459*, 852–856.

Martínez Segura, A. (2017). Coordination between cell size and global gene expression. Department of Clinical Sciences PhD Thesis, Imperial College London.

Martinez-Campa, C., Politis, P., Moreau, J.-L., Kent, N., Goodall, J., Mellor, J., and Goding, C.R. (2004). Precise Nucleosome Positioning and the TATA Box Dictate Requirements for the Histone H4 Tail and the Bromodomain Factor Bdf1. Mol. Cell *15*, 69–81.

Martínez-Reyes, I., Diebold, L.P., Kong, H., Schieber, M., Huang, H., Hensley, C.T., Mehta, M.M., Wang, T., Santos, J.H., Woychik, R., et al. (2016). TCA Cycle and Mitochondrial Membrane Potential Are Necessary for Diverse Biological Functions. Mol. Cell *61*, 199–209.

Mason, J.A., and Mellor, J. (1997). Isolation of nuclei for chromatin analysis in fission yeast. Nucleic Acids Res. *25*, 4700–4701.

Matangkasombut, O., and Buratowski, S. (2003). Different Sensitivities of Bromodomain Factors 1 and 2 to Histone H4 Acetylation. Mol. Cell *11*, 353–363.

Matangkasombut, O., Buratowski, R.M., Swilling, N.W., and Buratowski, S. (2000). Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. Genes Dev. *14*, 951–962.

Mavrich, T.N., Jiang, C., Ioshikhes, I.P., Li, X., Venters, B.J., Zanton, S.J., Tomsho, L.P., Qi, J., Glaser, R., Schuster, S.C., et al. (2008a). Nucleosome organization in the Drosophila genome. Nature *453*, 358–362.

Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. (2008b). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. Genome Res. *18*, 1073–1083.

Metzl-Raz, E., Kafri, M., Yaakov, G., Soifer, I., Gurvich, Y., and Barkai, N. (2017). Principles of cellular resource allocation revealed by condition-dependent proteome profiling. ELife *6*, e28034.

Mieczkowski, J., Cook, A., Bowman, S.K., Mueller, B., Alver, B.H., Kundu, S., Deaton, A.M., Urban, J.A., Larschan, E., Park, P.J., et al. (2016). MNase titration reveals differences between nucleosome occupancy and chromatin accessibility. Nat. Commun. *7*, 11485.

Miettinen, T.P., and Björklund, M. (2016). Cellular Allometry of Mitochondrial Functionality Establishes the Optimal Cell Size. Dev. Cell *39*, 370–382.

Minnoye, L., Marinov, G.K., Krausgruber, T., Pan, L., Marand, A.P., Secchia, S., Greenleaf, W.J., Furlong, E.E.M., Zhao, K., Schmitz, R.J., et al. (2021). Chromatin accessibility profiling methods. Nat. Rev. Methods Primer *1*, 1–24.

Mirzayans, R., Andrais, B., and Murray, D. (2018). Roles of Polyploid/Multinucleated Giant Cancer Cells in Metastasis and Disease Relapse Following Anticancer Treatment. Cancers *10*, 118.

Mitchison, J.M., and Creanor, J. (1971). Induction synchrony in the fission yeast Schizosaccharomyces pombe. Exp. Cell Res. *67*, 368–374.

Mivelaz, M., Cao, A.-M., Kubik, S., Zencir, S., Hovius, R., Boichenko, I., Stachowicz, A.M., Kurat, C.F., Shore, D., and Fierz, B. (2020). Chromatin Fiber Invasion and Nucleosome Displacement by the Rap1 Transcription Factor. Mol. Cell *77*, 488-500.e9.

Modi, S., Vargas-Garcia, C.A., Ghusinga, K.R., and Singh, A. (2017). Analysis of Noise Mechanisms in Cell-Size Control. Biophys. J. *112*, 2408–2418.

Monahan, B.J., Villén, J., Marguerat, S., Bähler, J., Gygi, S.P., and Winston, F. (2008). Fission yeast SWI/SNF and RSC complexes show compositional and functional differences from budding yeast. Nat. Struct. Mol. Biol. *15*, 873–880.

Mondesert, O., McGowan, C.H., and Russell, P. (1996). Cig2, a B-type cyclin, promotes the onset of S in Schizosaccharomyces pombe. Mol. Cell. Biol. *16*, 1527–1533.

Monds, R.D., Lee, T.K., Colavin, A., Ursell, T., Quan, S., Cooper, T.F., and Huang, K.C. (2014). Systematic Perturbation of Cytoskeletal Function Reveals a Linear Scaling Relationship between Cell Geometry and Fitness. Cell Rep. *9*, 1528–1537.

Moreno, S., and Nurse, P. (1994). Regulation of progression through the G1 phase of the cell cycle by the rum1+ gene. Trends Genet. *10*, 150.

Moseley, J.B., Mayeux, A., Paoletti, A., and Nurse, P. (2009). A spatial gradient coordinates cell size and mitotic entry in fission yeast. Nature *459*, 857–860.

Moyle-Heyrman, G., Zaichuk, T., Xi, L., Zhang, Q., Uhlenbeck, O.C., Holmgren, R., Widom, J., and Wang, J.-P. (2013). Chemical map of Schizosaccharomyces pombe reveals species-specific features in nucleosome positioning. Proc. Natl. Acad. Sci. *110*, 20158–20163.

Mueller, B., Mieczkowski, J., Kundu, S., Wang, P., Sadreyev, R., Tolstorukov, M.Y., and Kingston, R.E. (2017). Widespread changes in nucleosome accessibility without changes in nucleosome occupancy during a rapid transcriptional induction. Genes Dev. *31*, 451–462.

Müller, M., Pelkmans, L., and Berry, S. (2021). High content genome-wide siRNA screen to investigate the coordination of cell size and RNA production. Sci. Data *8*, 162.

Muñoz, S., Minamino, M., Casas-Delucchi, C.S., Patel, H., and Uhlmann, F. (2019). A Role for Chromatin Remodeling in Cohesin Loading onto Chromosomes. Mol. Cell *74*, 664-673.e5.

Murawska, M., Schauer, T., Matsuda, A., Wilson, M.D., Pysik, T., Wojcik, F., Muir, T.W., Hiraoka, Y., Straub, T., and Ladurner, A.G. (2020). The Chaperone FACT and Histone H2B Ubiquitination Maintain S. pombe Genome Architecture through Genic and Subtelomeric Functions. Mol. Cell *77*, 501-513.e7.

Nabilsi, N.H., Deleyrolle, L.P., Darst, R.P., Riva, A., Reynolds, B.A., and Kladde, M.P. (2014). Multiplex mapping of chromatin accessibility and DNA methylation within targeted single molecules identifies epigenetic heterogeneity in neural stem cells and glioblastoma. Genome Res. *24*, 329–339.

Nagai, S., Davis, R.E., Mattei, P.J., Eagen, K.P., and Kornberg, R.D. (2017). Chromatin potentiates transcription. Proc. Natl. Acad. Sci. *114*, 1536–1541.

Navarro, F.J., and Nurse, P. (2012). A systematic screen reveals new elements acting at the G2/M cell cycle control. Genome Biol. *13*, R36.

Neurohr, G.E., Terry, R.L., Lengefeld, J., Bonney, M., Brittingham, G.P., Moretto, F., Miettinen, T.P., Vaites, L.P., Soares, L.M., Paulo, J.A., et al. (2019). Excessive Cell Growth Causes Cytoplasm Dilution And Contributes to Senescence. Cell *176*, 1083-1097.e18.

Nguyen, V.Q., Ranjan, A., Liu, S., Tang, X., Ling, Y.H., Wisniewski, J., Mizuguchi, G., Li, K.Y., Jou, V., Zheng, Q., et al. (2021). Spatiotemporal coordination of transcription preinitiation complex assembly in live cells. Mol. Cell *81*, 3560-3575.e6.

Nicholson, D. (2020). On Being the Right Size, Revisited: The Problem with Engineering Metaphors in Molecular Biology. pp. 40–68.

Nocetti, N., and Whitehouse, I. (2016). Nucleosome repositioning underlies dynamic gene expression. Genes Dev. *30*, 660–672.

Norton, V.G., Imai, B.S., Yau, P., and Bradbury, E.M. (1989). Histone acetylation reduces nucleosome core particle linking number change. Cell *57*, 449–457.

Novák, B. (2013). Pom1 is not the size ruler. Cell Cycle *12*, 3463–3464.

Nurse, P. (1975). Genetic control of cell size at cell division in yeast. Nature *256*, 547–551.

Nurse, P., and Bissett, Y. (1981). Gene required in G1 for commitment to cell cycle and in G2 for control of mitosis in fission yeast. Nature *292*, 558–560.

Oda, A., Takemata, N., Hirata, Y., Miyoshi, T., Suzuki, Y., Sugano, S., and Ohta, K. (2015). Dynamic transition of transcription and chromatin landscape during fission yeast adaptation to glucose starvation. Genes Cells *20*, 392–407.

Olins, A.L., and Olins, D.E. (1974). Spheroid Chromatin Units (ν Bodies). Science *183*, 330–332.

Ozsolak, F., Song, J.S., Liu, X.S., and Fisher, D.E. (2007). High-throughput mapping of the chromatin structure of human promoters. Nat. Biotechnol. *25*, 244–248.

Padovan-Merhar, O., Nair, G.P., Biaesch, A.G., Mayer, A., Scarfone, S., Foley, S.W., Wu, A.R., Churchman, L.S., Singh, A., and Raj, A. (2015). Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcriptional Mechanisms. Mol. Cell *58*, 339–352.

Pan, K.Z., Saunders, T.E., Flor-Parra, I., Howard, M., and Chang, F. (2014). Cortical regulation of cell size by a sizer cdr2p. ELife *3*, e02040.

Payne, J.L., McClain, C.R., Boyer, A.G., Brown, J.H., Finnegan, S., Kowalewski, M., Krause, R.A., Lyons, S.K., McShea, D.W., Novack-Gottshall, P.M., et al. (2011). The evolutionary consequences of oxygenic photosynthesis: a body size perspective. Photosynth. Res. *107*, 37–57.

Pende, M., Kozma, S.C., Jaquet, M., Oorschot, V., Burcelin, R., Le Marchand-Brustel, Y., Klumperman, J., Thorens, B., and Thomas, G. (2000). Hypoinsulinaemia, glucose intolerance and diminished β-cell size in S6K1-deficient mice. Nature *408*, 994–997.

Pérez-Hidalgo, L., and Moreno, S. (2016). Nutrients control cell size. Cell Cycle *15*, 1655–1656.

Pointner, J., Persson, J., Prasad, P., Norman-Axelsson, U., Strålfors, A., Khorosjutina, O., Krietenstein, N., Svensson, J.P., Ekwall, K., and Korber, P. (2012). CHD1 remodelers regulate nucleosome spacing in vitro and align nucleosomal arrays over gene coding regions in S. pombe. EMBO J. *31*, 4388–4403.

Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E., et al. (2005). Genome-wide Map of Nucleosome Acetylation and Methylation in Yeast. Cell *122*, 517–527.

Pollard, T.D., Weihing, R.R., and Adelman, M.R. (1974). Actin And Myosin And Cell Movemen. CRC Crit. Rev. Biochem. *2*, 1–65.

Pradhan, S.K., Xue, Y., and Carey, M.F. (2015). Fragile Nucleosomes Influence Pol II Promoter Function. Mol. Cell *60*, 342–343.

Price, S.A., and Hopkins, S.S.B. (2015). The macroevolutionary relationship between diet and body mass across mammals. Biol. J. Linn. Soc. *115*, 173–184.

Pritchard, D.K., and Schubiger, G. (1996). Activation of transcription in Drosophila embryos is a gradual process mediated by the nucleocytoplasmic ratio. Genes Dev. *10*, 1131–1142.

Quintales, L., Vázquez, E., and Antequera, F. (2015). Comparative analysis of methods for genome-wide nucleosome cartography. Brief. Bioinform. *16*, 576–587.

Rach, E.A., Winter, D.R., Benjamin, A.M., Corcoran, D.L., Ni, T., Zhu, J., and Ohler, U. (2011). Transcription Initiation Patterns Indicate Divergent Strategies for Gene Regulation at the Chromatin Level. PLOS Genet. *7*, e1001274.

Raj, A., and van Oudenaarden, A. (2008). Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. Cell *135*, 216–226.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165.

Rando, O.J. (2007). Global patterns of histone modifications. Curr. Opin. Genet. Dev. *17*, 94–99.

Rando, O.J. (2010). Chapter 5 - Genome-Wide Mapping of Nucleosomes in Yeast. In Methods in Enzymology, (Academic Press), pp. 105–118.

Rando, O.J., and Winston, F. (2012). Chromatin and Transcription in Yeast. Genetics *190*, 351–387.

Rao, C.V., Wolf, D.M., and Arkin, A.P. (2002). Control, exploitation and tolerance of intracellular noise. Nature *420*, 231–237.

Raser, J.M., and O'Shea, E.K. (2005). Noise in Gene Expression: Origins, Consequences, and Control. Science *309*, 2010–2013.

Rawal, Y., Chereji, R.V., Qiu, H., Ananthakrishnan, S., Govind, C.K., Clark, D.J., and Hinnebusch, A.G. (2018). SWI/SNF and RSC cooperate to reposition and evict promoter nucleosomes at highly expressed genes in yeast. Genes Dev. *32*, 695–710.

Reid, M.A., Dai, Z., and Locasale, J.W. (2017). The impact of cellular metabolism on chromatin dynamics and epigenetics. Nat. Cell Biol. *19*, 1298–1306.

Rhee, H.S., and Pugh, B.F. (2012). Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature *483*, 295–301.

Rizzo, J.M., Mieczkowski, P.A., and Buck, M.J. (2011). Tup1 stabilizes promoter nucleosome positioning and occupancy at transcriptionally plastic genes. Nucleic Acids Res. *39*, 8803–8819.

Rizzo, J.M., Bard, J.E., and Buck, M.J. (2012). Standardized collection of MNase-seq experiments enables unbiased dataset comparisons. BMC Mol. Biol. *13*, 15.

Robert, L., Hoffmann, M., Krell, N., Aymerich, S., Robert, J., and Doumic, M. (2014). Division in Escherichia coli is triggered by a size-sensing rather than a timing mechanism. BMC Biol. *12*, 17.

Rodríguez-López, M., Gonzalez, S., Hillson, O., Tunnacliffe, E., Codlin, S., Tallada, V.A., Bähler, J., and Rallis, C. (2020). The GATA Transcription Factor Gaf1 Represses tRNAs, Inhibits Growth, and Extends Chronological Lifespan Downstream of Fission Yeast TORC1. Cell Rep. *30*, 3240-3249.e4.

Russell, P., and Nurse, P. (1987). Negative regulation of mitosis by wee1+, a gene encoding a protein kinase homolog. Cell *49*, 559–567.

Sainsbury, S., Bernecky, C., and Cramer, P. (2015). Structural basis of transcription initiation by RNA polymerase II. Nat. Rev. Mol. Cell Biol. *16*, 129–143.

Saint, M., Bertaux, F., Tang, W., Sun, X.-M., Game, L., Köferle, A., Bähler, J., Shahrezaei, V., and Marguerat, S. (2019). Single-cell imaging and RNA sequencing reveal patterns of gene expression heterogeneity during fission yeast growth and adaptation. Nat. Microbiol. *4*, 480–491.

Santos, J.H. (2021). Mitochondria signaling to the epigenome: A novel role for an old organelle. Free Radic. Biol. Med. *170*, 59–69.

Sato, S., Burgess, S.B., and McIlwain, D.L. (1994). Transcription and Motoneuron Size. J. Neurochem. *63*, 1609–1615.

Saucedo, L.J., and Edgar, B.A. (2002). Why size matters: altering cell size. Curr. Opin. Genet. Dev. *12*, 565–571.

Schmidt, E.E., and Schibler, U. (1995). Cell size regulation, a mechanism that controls cellular RNA accumulation: consequences on regulation of the ubiquitous transcription factors Oct1 and NF-Y, and the liver-enriched transcription factor DBP. J Cell 128–467.

Schmoller, K.M., Turner, J.J., Kõivomägi, M., and Skotheim, J.M. (2015). Dilution of the cell cycle inhibitor Whi5 controls budding-yeast cell size. Nature *526*, 268–272.

Scott, M., and Hwa, T. (2011). Bacterial growth laws and their applications. Curr. Opin. Biotechnol. *22*, 559–565.

Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z., and Hwa, T. (2010). Interdependence of Cell Growth and Gene Expression: Origins and Consequences. Science *330*, 1099–1102.

Seel, A., Padovani, F., Finster, A., Mayer, M., Bureik, D., Osman, C., Klecker, T., and Schmoller, K.M. (2021). Regulation with cell size ensures mitochondrial DNA homeostasis during cell growth.

Shahrezaei, V., and Marguerat, S. (2015). Connecting growth with gene expression: of noise and numbers. Curr. Opin. Microbiol. *25*, 127–135.

Shahrezaei, V., and Swain, P.S. (2008). The stochastic nature of biochemical networks. Curr. Opin. Biotechnol. *19*, 369–374.

Sharon, E., Dijk, D. van, Kalma, Y., Keren, L., Manor, O., Yakhini, Z., and Segal, E. (2014). Probing the effect of promoters on noise in gene expression using thousands of designed sequences. Genome Res. *24*, 1698–1706.

Shivaswamy, S., Bhinge, A., Zhao, Y., Jones, S., Hirst, M., and Iyer, V.R. (2008). Dynamic Remodeling of Individual Nucleosomes Across a Eukaryotic Genome in Response to Transcriptional Perturbation. PLOS Biol. *6*, e65.

Si, F., Le Treut, G., Sauls, J.T., Vadia, S., Levin, P.A., and Jun, S. (2019). Mechanistic Origin of Cell-Size Control and Homeostasis in Bacteria. Curr. Biol. *29*, 1760-1770.e7.

Sinha, I., Wirén, M., and Ekwall, K. (2006). Genome-wide patterns of histone modifications in fission yeast. Chromosome Res. *14*, 95–105.

Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. ELife *6*, e21856.

Smith, F.A., and Lyons, S.K. (2011). How big should a mammal be? A macroecological look at mammalian body size over space and time. Philos. Trans. R. Soc. B Biol. Sci. *366*, 2364–2378.

Soriano, I., Quintales, L., and Antequera, F. (2013). Clustered regulatory elements at nucleosome-depleted regions punctuate a constant nucleosomal landscape in Schizosaccharomyces pombe. BMC Genomics *14*, 813.

Stevens, K.M., Swadling, J.B., Hocher, A., Bang, C., Gribaldo, S., Schmitz, R.A., and Warnecke, T. (2020). Histone variants in archaea and the evolution of combinatorial chromatin complexity. Proc. Natl. Acad. Sci. *117*, 33384–33395.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Strahl, B.D., and Briggs, S.D. (2021). The SAGA continues: The rise of cis- and trans-histone crosstalk pathways. Biochim. Biophys. Acta BBA - Gene Regul. Mech. *1864*, 194600.

Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. Nat. Struct. Mol. Biol. *20*, 267–273.

Sun, X.-M., Bowman, A., Priestman, M., Bertaux, F., Martinez-Segura, A., Tang, W., Whilding, C., Dormann, D., Shahrezaei, V., and Marguerat, S. (2020). Size-Dependent Increase in RNA Polymerase II Initiation Rates Mediates Gene Expression Scaling with Cell Size. Curr. Biol. *30*, 1217-1230.e7.

Suter, D.M., Molina, N., Gatfield, D., Schneider, K., Schibler, U., and Naef, F. (2011). Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. Science.

Swaffer, M.P., Kim, J., Chandler-Brown, D., Langhinrichs, M., Marinov, G.K., Greenleaf, W.J., Kundaje, A., Schmoller, K.M., and Skotheim, J.M. (2021a). Transcriptional and chromatin-based partitioning mechanisms uncouple protein scaling from cell size. Mol. Cell *81*, 4861-4875.e7.

Swaffer, M.P., Marinov, G.K., Zheng, H., Jones, A.W., Greenwood, J., Kundaje, A., Snijders, A.P., Greenleaf, W.J., Reyes-Lamothe, R., and Skotheim, J.M. (2021b). RNA polymerase II dynamics and mRNA stability feedback determine mRNA scaling with cell size. 2021.09.20.461005.

Taheri-Araghi, S., Bradde, S., Sauls, J.T., Hill, N.S., Levin, P.A., Paulsson, J., Vergassola, M., and Jun, S. (2015). Cell-Size Control and Homeostasis in Bacteria. Curr. Biol. *25*, 385–391.

Takayama, Y., and Takahashi, K. (2007). Differential regulation of repeated histone genes during the fission yeast cell cycle. Nucleic Acids Res. *35*, 3223–3237.

Takemata, N., and Ohta, K. (2017). Role of non-coding RNA transcription around gene regulatory elements in transcription factor recruitment. RNA Biol. *14*, 1–5.

Talia, S.D., Skotheim, J.M., Bean, J.M., Siggia, E.D., and Cross, F.R. (2007). The effects of molecular noise and size control on variability in the budding yeast cell cycle. Nature *448*, 947–951.

Taylor, B.C., and Young, N.L. (2021). Combinations of histone post-translational modifications. Biochem. J. *478*, 511–532.

Thodberg, M., Thieffry, A., Bornholdt, J., Boyd, M., Holmberg, C., Azad, A., Workman, C.T., Chen, Y., Ekwall, K., Nielsen, O., et al. (2019). Comprehensive profiling of the fission yeast transcription start site activity during stress and media response. Nucleic Acids Res. *47*, 1671–1691.

Thompson, D.W. (1942). On growth and form, 2nd ed (Oxford, England: Macmillan, Cambridge University Pre).

Timmers, H.Th.M. (2021). SAGA and TFIID: Friends of TBP drifting apart. Biochim. Biophys. Acta BBA - Gene Regul. Mech. *1864*, 194604.

Tirosh, I., and Barkai, N. (2008). Two strategies for gene regulation by promoter nucleosomes. Genome Res. *18*, 1084–1091.

Tjian, R., and Maniatis, T. (1994). Transcriptional activation: A complex puzzle with few easy pieces. Cell *77*, 5–8.

Tolstorukov, M.Y., Sansam, C.G., Lu, P., Koellhoffer, E.C., Helming, K.C., Alver, B.H., Tillman, E.J., Evans, J.A., Wilson, B.G., Park, P.J., et al. (2013). Swi/Snf chromatin remodeling/tumor suppressor complex establishes nucleosome occupancy at target promoters. Proc. Natl. Acad. Sci. *110*, 10165–10170.

Toselli-Mollereau, E., Robellet, X., Fauque, L., Lemaire, S., Schiklenk, C., Klein, C., Hocquet, C., Legros, P., N'Guyen, L., Mouillard, L., et al. (2016). Nucleosome eviction in mitosis assists condensin loading and chromosome condensation. EMBO J. *35*, 1565–1581.

Trotta, V., Calboli, F.C., Ziosi, M., and Cavicchi, S. (2007). Fitness variation in response to artificial selection for reduced cell area, cell number and wing area in natural populations of Drosophila melanogaster. BMC Evol. Biol. *7*, S10.

Tsompana, M., and Buck, M.J. (2014). Chromatin accessibility: a window into the genome. Epigenetics Chromatin *7*, 1–16.

Turner, J.J., Ewald, J.C., and Skotheim, J.M. (2012). Cell Size Control in Yeast. Curr. Biol. *22*, R350–R359.

Umeda, M., Tsunekawa, C., Senmatsu, S., Asada, R., Abe, T., Ohta, K., Hoffman, C.S., and Hirota, K. (2018). Histone Chaperone Asf1 Is Required for the Establishment of Repressive Chromatin in Schizosaccharomyces pombe fbp1 Gene Repression. Mol. Cell. Biol.

Vargas-Garcia, C.A., Ghusinga, K.R., and Singh, A. (2018). Cell size control and gene expression homeostasis in single-cells. Curr. Opin. Syst. Biol. *8*, 109–116.

Varsano, G., Wang, Y., and Wu, M. (2017). Probing Mammalian Cell Size Homeostasis by Channel-Assisted Cell Reshaping. Cell Rep. *20*, 397–410.

Venkatesh, S., and Workman, J.L. (2015). Histone exchange, chromatin structure and the regulation of transcription. Nat. Rev. Mol. Cell Biol. *16*, 178–189.

Vera, D.L., Madzima, T.F., Labonne, J.D., Alam, M.P., Hoffman, G.G., Girimurugan, S.B., Zhang, J., McGinnis, K.M., Dennis, J.H., and Bass, H.W. (2014). Differential Nuclease

Sensitivity Profiling of Chromatin Reveals Biochemical Footprints Coupled to Gene Expression and Functional DNA Elements in Maize. Plant Cell *26*, 3883–3893.

Vermeij, G.J. (2016). Gigantism and Its Implications for the History of Life. PLOS ONE *11*, e0146092.

Vettese-Dadey, M., Grant, P.A., Hebbes, T.R., Crane- Robinson, C., Allis, C.D., and Workman, J.L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. EMBO J. *15*, 2508–2518.

Wang, H., Carey, L.B., Cai, Y., Wijnen, H., and Futcher, B. (2009). Recruitment of Cln3 Cyclin to Promoters Controls Cell Cycle Entry via Histone Deacetylase and Other Targets. PLOS Biol. *7*, e1000189.

Warfield, L., Ramachandran, S., Baptista, T., Devys, D., Tora, L., and Hahn, S. (2017). Transcription of Nearly All Yeast RNA Polymerase II-Transcribed Genes Is Dependent on Transcription Factor TFIID. Mol. Cell *68*, 118-129.e5.

Waterborg, J.H. (2011). Dynamics of histone acetylation in vivo. A function for acetylation turnover? Biochem. Cell Biol.

Weber, C.M., Ramachandran, S., and Henikoff, S. (2014). Nucleosomes Are Context-Specific, H2A.Z-Modulated Barriers to RNA Polymerase. Mol. Cell *53*, 819–830.

Weiner, A., Hughes, A., Yassour, M., Rando, O.J., and Friedman, N. (2010). High-resolution nucleosome mapping reveals transcription-dependent promoter packaging. Genome Res. *20*, 90–100.

Wilkins, M.H.F., Zubay, G., and Wilson, H.R. (1959). X-ray diffraction studies of the molecular structure of nucleohistone and chromosomes. J. Mol. Biol. *1*, 179-IN10.

Wilson, B.G., and Roberts, C.W.M. (2011). SWI/SNF nucleosome remodellers and cancer. Nat. Rev. Cancer *11*, 481–492.

Wilusz, C.J., and Wilusz, J. (2004). Bringing the role of mRNA decay in the control of gene expression into focus. Trends Genet. *20*, 491–497.

Wood, E., and Nurse, P. (2013). Pom1 and cell size homeostasis in fission yeast. Cell Cycle *12*, 3417–3425.

Wood, E., and Nurse, P. (2015). Sizing up to Divide: Mitotic Cell-Size Control in Fission Yeast. Annu. Rev. Cell Dev. Biol. *31*, 11–29.

Woodcock, C.L., Safer, J.P., and Stanchfield, J.E. (1976). Structural repeating units in chromatin. I. Evidence for their general occurrence. Exp. Cell Res. *97*, 101–110.

Workman, J.L. (2006). Nucleosome displacement in transcription. Genes Dev. *20*, 2009–2017.

Workman, J.L., Abmayr, S.M., Cromlish, W.A., and Roeder, R.G. (1988). Transcriptional regulation by the immediate early protein of pseudorabies virus during in vitro nucleosome assembly. Cell *55*, 211–219.

Wu, C.-Y., Rolfe, P.A., Gifford, D.K., and Fink, G.R. (2010). Control of Transcription by Cell Size. PLOS Biol. *8*, e1000523.

Xi, Y., Yao, J., Chen, R., Li, W., and He, X. (2011). Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. Genome Res. *21*, 718–724.

Yaakov, G., Jonas, F., and Barkai, N. (2021). Measurement of histone replacement dynamics with genetically encoded exchange timers in yeast. Nat. Biotechnol. *39*, 1434–1443.

Yague-Sanz, C., Vázquez, E., Sánchez, M., Antequera, F., and Hermand, D. (2017). A conserved role of the RSC chromatin remodeler in the establishment of nucleosome-depleted regions. Curr. Genet. *63*, 187–193.

Yuan, G.-C., Liu, Y.-J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., and Rando, O.J. (2005). Genome-Scale Identification of Nucleosome Positions in S. cerevisiae. Science *309*, 626–630.

Zaret, K. (1999). Micrococcal Nuclease Analysis of Chromatin Structure. Curr. Protoc. Mol. Biol. *45*, 21.1.1-21.1.17.

Zenklusen, D., Larson, D.R., and Singer, R.H. (2008). Single-RNA counting reveals alternative modes of gene expression in yeast. Nat. Struct. Mol. Biol. *15*, 1263–1271.

Zhang, H., and Reese, J.C. (2007). Exposing the core promoter is sufficient to activate transcription and alter coactivator requirement at RNR3. Proc. Natl. Acad. Sci. *104*, 8833–8838.

Zhang, Q., Oh, D.-H., DiTusa, S.F., RamanaRao, M.V., Baisakh, N., Dassanayake, M., and Smith, A.P. (2018). Rice nucleosome patterns undergo remodeling coincident with stress-induced gene expression. BMC Genomics *19*, 97.

Zhang, T., Zhang, W., and Jiang, J. (2015). Genome-Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants. Plant Physiol. *168*, 1406–1416.

Zhang, Y., Moqtaderi, Z., Rattner, B.P., Euskirchen, G., Snyder, M., Kadonaga, J.T., Liu, X.S., and Struhl, K. (2009). Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. Nat. Struct. Mol. Biol. *16*, 847–852.

Zhou, J., Wang, X., He, K., Charron, J.-B.F., Elling, A.A., and Deng, X.W. (2010). Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in Arabidopsis reveals correlation between multiple histone marks and gene expression. Plant Mol. Biol. *72*, 585–595.

Zhurinsky, J., Leonhard, K., Watt, S., Marguerat, S., Bähler, J., and Nurse, P. (2010). A Coordinated Global Control over Cellular Transcription. Curr. Biol. *20*, 2010–2015.