Salma Rahme

Licenciatura em Engenharia Biotecnológica

A comprehensive analysis of the role of TTF2/ Lds throughout the cell cycle

Dissertação para obtenção do Grau de Mestre em Bioquímica para a Saúde

Orientador: Sara Carvalhal, Postdoc Researcher, Instituto Gulbenkian de Ciência
Co-orientador: Raquel Oliveira, Principal Investigator, Instituto Gulbenkian de Ciência

Novembro 2020

Salma Rahme

Licenciatura em Engenharia Biotecnológica

A comprehensive analysis of the role of TTF2/ Lds throughout the cell cycle

Dissertação para obtenção do Grau de Mestre em Bioquímica para a Saúde

Orientador: Sara Carvalhal, Postdoc Researcher, Instituto Gulbenkian de Ciência
Co-orientador: Raquel Oliveira, Principal Investigator, Instituto Gulbenkian de Ciência

Júri:

Presidente: Prof. Doutora Maria Teresa Catarino
Arguente: Doutor Leonardo Gaston Guilgur
Vogalais: Doutora Sara Carvalhal

A Faculdade de Ciências e Tecnologia

Novembro 2020

This page has been intentionally left blank.

Copyright

Acknowledgments

I am truly grateful to complete this Master's thesis that meant more than just a degree for me. That would not have been done without the help and the support of many people.

My deep thanks to Portugal and the Portuguese people for their lovely hearts, and their kind words.
To my beautiful Syria, where my heart belongs, I hope one day we will see the sun of justice.
Many thanks to the Global Platform for Syrian students, for saving our future and giving us the chance to learn, and very big thanks to Dr. Helena Barroco for making a difference in this world, thank you because Syrian lives matter to you.
Thank you to the coordinator of the Master, Prof. Teresa Catarino, for her trust and unlimited support, and for being my "elo de ligação".
Thank you to Dr. Sara Carvalhal for her supervision and for helping me during the entire journey, and for being with me whenever I needed her.
Thank you to Dr. Raquel Oliveira for accepting me in her laboratory, and for her feedbacks and professional assistance.
Thank you to all members of the Chromosome dynamics Lab each one of you left a beautiful part in my memory, thank you for the cakes, for the jokes on the WhatsApp group, and for being a family.
Thank you to my family and my friends, I am super lucky to have you in my life, love you all.

Abstract

Coordination of transcription and cell division is critical for life, but how transcription is switched off during mitosis remains unknown. Putative players in this mitotic transcription inactivation (MTI) are the human transcription termination factor 2 (TTF2) and Lodestar (Lds), both considered orthologs genes. Sequencing information is the major premise to support their orthology. In fact, TTF2 and Lds are 39% identical, 56% similar, and belong to the Snf2 helicase-like family. To further corroborate the idea that they are putative orthologs, in addition to their sequence analysis, I reviewed the published literature to discuss their shared functional aspects in particular at transcription and cell division.

TTF2/Lds display similar *in vitro* transcription termination activities supporting being orthologs. Despite termination activity being explored for TTF2 more than Lds, they have dsDNA-dependent ATPase activity, release short transcripts associated with RNA Polymerase II (Pol II) from the DNA template and suppress long transcripts. In addition, TTF2 releases Poll II from the DNA template regardless of the phosphorylation state, and depletion of TTF2 causes retention of Pol II at metaphase chromosomes, implicating TTF2 in MTI.

Concerning their functions on cell division, TTF2 is less described than Lds. Both TTF2/Lds depletion causes erroneous chromosome segregation but the degree and type of errors are very different. Due to the limited published data, it is still premature to conclude if they share ortholog mitotic functions.

From my study of the available literature, despite TTF2 and Lds orthology being substantiated on their transcription functions, it remains unclear whether mitotic defects associated with TTF2/Lds depletion represent different protein functions, one in transcription and another in faithful chromosome segregation, or instead implies MTI mis-regulation can per se affect mitotic fidelity. Further research is needed to dissect this possible dichotomy, namely by the use of live imaging techniques.

Keywords

Transcription termination factor 2 (TTF2); Lodestar (Lds); Orthologs; Transcription termination; Mitosis.

Resumo

A coordenação entre a transcrição e a divisão celular é crucial para a vida. No entanto, é ainda desconhecido como é inativada a transcrição antes da mitose. Pensa-se que as proteínas ortólogas TTF2 e Lds, participam nesta inativação da transcrição mitótica (ITM). De facto, a TTF2 e a Lds são 39% idênticos, 56% semelhantes e ambos pertencem à família das helicases tipo Snf2. Com o objetivo de confirmar esta possível ortologia, revi a literatura publicada nestas proteínas de modo a discutir potenciais funções ortólogas, nomedamente na transcrição e divisão celular.

Tanto a TTF2 como a Lds actuam na terminação de transcrição *in vitro*, apoiando a existência de funções ortólogas. Embora se saiba mais da TTF2 que da Lds, ambas têm atividade ATPase dependente de dsADN, libertam transcritos curtos associados com ARN Polimerase II (Pol II) do ADN e suprimem transcritos longos. A TTF2 liberta também a Poll II independentemente dos seus níveis fosforilação, e a inibição da TTF2 retém a Pol II nos cromossomas mitóticos, sugerindo uma função na ITM.

Relativamente às suas funções na divisão celular, a TTF2 é menos descrita do que a Lds. A depleção da TTF2/ Lds causa segregação incorreta dos cromossomas, mas muito diferentes relativamente ao tipo e serevidade. Dado as escassas publicações no tema, torna-se prematuro concluir se eles compartilham funções mitóticas ortólogas ou não.

Apesar da ortologia da TTF2 e Lds ser fundamentada pelas suas funções na transcrição, pela minha revisão da literatura não ficou claro se os defeitos mitóticos associados à depleção de TTF2 / Lds representam diferentes funções destas proteínas, uma na transcrição e outra na segregação dos cromossomas, ou se estão relacionados com uma desregulação na ITM. São necessários mais estudos para compreender esta possível dicotomia, nomeadamente através do uso de microscopia *in vivo*.

Palavras-chave

Fator 2 de terminação da transcrição (TTF2); Lodestar (Lds); Ortólogos; Inativação da transcrição mitótica.

Table of Content

1 Contents

List of figures

List of tables

Abbreviations

AIDS— Immune Deficiency Syndrome
APC/C— Anaphase Promoting Complex/Cyclosome
ATP— Adenosine triphosphate
CCAPs— CDK9/CCNT1-associated protein complexes
ChIP-seq— Chromatin immunoprecipitation followed by sequencing
CPA— Cleavage and polyadenylation
CPSF— Cleavage and Polyadenylation Specificity Factor
CTD— Carboxyl C-terminal domain
dATP— Deoxyadenosine triphosphate
DNA— Deoxyribonucleic acid
dsDNA— Double-stranded DNA
DSIF— DRB sensitivity-inducing factor
EGFP— Enhanced green fluorescent Protein
GTFs— General transcription factors
HIV-1— Human Immunodeficiency Virus type-1
Lds— Lodestar
Lds-EGFP— EGFP tagged at the C-terminal of Lodestar
LTR— Long terminal repeat
mRNA— messenger RNA
MTA— Mitotic transcriptional activation
MTI— Mitotic transcription inactivation
NEBD— Nuclear envelope breakdown
NELF—Negative elongation factor
NTD— Amino N terminal domain
ORFs— Open reading frames
PAS— Polyadenylation signal
PIC— Pre-initiation complex
Pol II— RNA polymerase II
P-TEFb— Positive transcription elongation factor
RNA— Ribonucleic acid
SC— Sister Chromatids
SMC— Structural maintenance of chromosomes
Tat— Transcriptional transactivator
TFIIB— Transcription factor II B
TFIID— Transcription factor II D
TFIIE— Transcription factor II E
TFIIF— Transcription factor II F
TFIIH— Transcription factor II H
TSS— Transcription Start Site
TTF2— Transcription termination factor 2
Xrn2— 5'-3' Exoribonuclease 2

This page has been intentionally left blank.

1 Introduction

Cell division is the process by which a parent cell divides into two daughter cells which includes the faithful replication and segregation of the genetic material during mitosis. For that, the spatial and temporal coordination between different cellular activities such as transcription and proliferation are crucial to secure cells' survival.

During mitosis, cells undergo a massive structural reorganization of chromatin condensation (reviewed in Shoaib, Nair, and Sørensen 2020), coupled with a dramatic shutdown of transcription, where most genes are transcriptionally inactivated before reactivation in anaphase/telophase (reviewed in Kang et al. 2020). Despite some molecular aspects of Mitotic Transcriptional Inactivation (MTI) start to be elucidated, they remain to be fully understood, as well as the impact of its mis-regulation. Uncovering them will contribute to a deep understanding of the biological and physiological consequences of health and disease that are related to transcription alternations. This global transcription inhibition during mitosis is associated with inhibitory phosphorylation of the different basal and specific transcription factors, and exclusion of RNA polymerases (Gottesfeld and Forbes 1997). However, several other mechanisms involving key factors have been proposed to promote this drastic transcriptional silencing, among which chromosome condensation (Gottesfeld and Forbes 1997; Robellet, Vanoosthuyse, and Bernard 2017). Complexes of the family of the structural maintenance of chromosomes (SMC), condensin (Iwasaki and Noma 2016; Paul, Hochwagen, and Ercan 2019) and cohesin, promote condensation and cohesion between sister chromatids (SC) (Ball and Yokomori 2001), and cooperatively ensure chromosome compaction and SC resolution. *In vivo* and *in vitro* studies, well-described their role in transcriptional regulation (reviewed in Iwasaki et al. 2019) which likely impacts on the silencing of the genome prior to DNA segregation. These SMC complexes regulate chromosome structure by extruding DNA loops (Ganji et al. 2018) and organize the short-range interactions between enhancers and promoters (Iwasaki et al. 2019; W. Li et al. 2015; Rao et al. 2017). The retention of cohesin on the mitotic arms, for instance, results in active RNA polymerase ll (Pol II) and transcription elongation (Perea-Resa et al. 2020). Also, condensin preferentially enriches at the transcriptional termination sites of mitotic actively transcribed genes, which may be relevant to release the mitotic transcripts from DNA (Nakazawa, Arakawa, and Yanagida 2019). In line with the idea of proteins that regulate and contribute to chromosomal architecture and condensation during mitosis, there is Topoisomerase II (Top II), an enzyme that is fundamental for condensation and maintaining the DNA topology, where its decatenation activity disentangles SC intertwines (Hirano and Mitchison 1991; Piskadlo and Oliveira 2017). It has also been suggested that Top II can impact the MTI (Gottesfeld and Forbes 1997). Histone modification has also been correlated to chromosome condensation and segregation (Gurley, Walters, and Tobey 1974; Wei et al. 1999) where phosphorylation at serine 10 in the N-terminal tail of H3 has emerged as a conclusive step in controlling transcription during both mitosis and meiosis in eukaryotic genes (Nowak and Corces 2004). Although considerable studies have been focused on the mechanism of

DNA condensation and segregation, much less is known concerning the mechanism of mitotic transcription silencing. It remains unknown how chromosome compaction can directly interfere with transcription and is inadequate to explain the recent evidence on active transcription in some chromosomal regions (Chan et al. 2012; Palozola, Donahue, et al. 2017), such as the centromeres (Perea-Resa et al. 2020) and bookmarked genes (Palozola, Donahue, et al. 2017).

Equally puzzling is the functional relevance for this transient transcriptional shutdown. A possible reason for such inhibition on transcriptional activity could be the saving of energy, giving cells their own time to be only focused on chromosome segregation. Another suggested reason is that transcription requires open access to the DNA, which could be considered incompatible with chromosome condensation and segregation (Gottesfeld and Forbes 1997; Robellet, Vanoosthuyse, and Bernard 2017).

To uncover both the mechanisms and the functional relevance of MTI, we need a much more detailed analysis of this process. In addition to the chromosome structural factors, novel players might be critical for transcriptional silencing. The ATP-dependent chromatin remodeling complexes could contribute to MIT (reviewed in Chen et al. 2005). It has been found that the phosphorylation of components of the SWI/SNF complex during mitosis inactivates transcription (Sif et al. 1998). Furthermore, in very long genes, premature termination of Pol II has been seen at mitosis and proposed as a way of silencing transcription (reviewed in Gottesfeld and Forbes 1997). Current studies at the Oliveira Lab suggest that a poorly characterized helicase-like *Drosophila* protein called Lodestar encoded by the *lds* gene is a notable candidate that may have crucial roles in mitotic transcription inactivation and chromosome fidelity.

Lds was isolated for the first time from *Drosophila* Kc cell nuclear extracts (D. H. Price, Sluder, and Greenleaf 1987). *In vitro* and *in vivo* studies, showed that it is a factor with roles in transcription termination by Pol II (Xie and Price 1996), and on chromosome stability (Girdham and Glover 1991; Szabad, Máthé, and Puro 1995; Szalontai et al. 2009). Comparative amino acid sequence analysis suggests that the human transcription factor 2 (TTF2) encoded by the *ttf2* gene is its human ortholog. TTF2 plays similar roles in transcription termination and possibly on faithful chromosome segregation (Jiang et al. 2004; M. Liu, Xie, and Price 1998).

Despite the common assumption that orthologs share similar functions, there are several examples of divergence between orthologs functions. In this monograph, I will dissect all the published studies on Lds and TTF2 to perform a parallel analysis on the idea that proteins are putative orthologs and better understand where their functions diverge or not. This integrative study will be firstly focused on the structural analysis of both proteins, followed by the analysis of the *in-vitro* published studies to explore their roles in transcription. I will then focus on published studies in cell division, and finally, I will discuss the available few studies on Lds and TTF2 related to other functions. Different studies have named these transcription termination factors with diverse names and acronymous (for instance in flies, it has been called the *Drosophila* Factor 2 (DmF2) and Lodestar (Lds), while in humans it has

been called the human factor 2 (HuF2) and transcription termination factor 2 (TTF2)); in this monograph, I will always refer to the *Drosophila* protein as Lds and the human as TTF2.

1.1 TTF2/Lds are members of the Snf2 helicases protein family

1.1.1 The helicases and Snf2 family

Chromosomal packing of DNA into a hierarchical chromatin structure serves the regulation and coordination of all DNA processes across the cell cycle including replication, segregation, transcription, recombination, and DNA repair (reviewed in Ferreira, Flaus, and Owen-Hughes 2007). These important cellular processes require direct access to the DNA underlying sequences through nucleosomes remodeling and unwinding of the double-stranded DNA (dsDNA) (reviewed in Kobayashi and Kurumizaka 2019; Racki and Narlikar 2008). Duplex DNA unwinding is catalyzed by a wide range of highly conserved "motor" enzymes known as helicases (Tuteja and Tuteja 2004). More specifically, helicases hydrolyze the adenosine triphosphate (ATP) to translocate along the DNA or RNA backbones and simplify the unwinding of the respective complementary strands by destabilizing the hydrogen bonds between the bases. Sequence analyzes of the helicases from several organisms identified seven to nine conserved fingerprints sequences named "helicase signature motifs": Q, I, Ia, Ib, II, III, IV, V, and VI (Gorbalenya et al. 1988; Tanner et al. 2003; Tuteja and Tuteja 2004). These motifs are clustered in 200–700 amino acids to form the core region of helicases (Tuteja and Tuteja 2004). Q and Ib motifs have been identified in some, but not all helicases. For example, the Q motif is a cluster of nine amino-acids involved in DNA binding but not ATP binding, and it is a characteristic unique motif in the family of DEAD-box family (DEAD or DEAH or DEXH) that are ATP-dependent RNA helicases (reviewed in Ding et al. 2015; Tanner et al. 2003; Tuteja and Tuteja 2004). Helicases are divided into six superfamilies depending on their sharing sequences motifs, as well as, their role in DNA processes (reviewed in Singleton, Dillingham, and Wigley 2007), and even their activity. Some helicases are very active and can separate DNA duplexes at a constant rate, independently of the stability between the base pairs. On the opposite side, there are passive helicases, which move along the polynucleotide chain where base pairs tend to open more easily on the helix terminal (a known process named base fraying) (Ferreira, Amarante, and Weber 2015). Another group of helicases has an intermediate activity because they stop on stable regions that are rich in G/C, but move faster in less stable regions (reviewed in Raj et al. 2019). On the other hand, some proteins within the helicases families are not *bona fide* helicases since they do not unwind DNA, but use ATP hydrolysis to impose another chromatin changes (reviewed in Ryan and Owen-Hughes 2011).

The largest superfamily of Helicases is SF2 which embraces seven conserved signature motifs (I, Ia, II-VI) (Tuteja and Tuteja 2004). Within this superfamily, there are several families among them the Snf2 (see Figure 1.1). Snf2 family proteins share the sequence motifs that are characteristic of the SF2 helicases (Davis, Kunisawa, and Thorner 1992; Gorbalenya and Koonin 1993). However, they lack the usual helicase activity where chromatin remodeling is associated with the dsDNA separation

(Ferreira, Flaus, and Owen-Hughes 2007). Instead, Snf2 family proteins translocate on the dsDNA and utilize the energy of the ATP hydrolysis in order to establish a superhelical torsion in the DNA and/or remodel the nucleosomes (Ryan and Owen-Hughes 2011; Thomä et al. 2005). In other words, their mechanism of action is different from canonical helicases, because Snf2 proteins catalyze dynamic transitions in the structure of chromatin causing topological changes in the DNA instead of unwinding of the double helix (Ferreira, Flaus, and Owen-Hughes 2007). Moreover, they are different when compared with other SF2 members. For instance, the space between their helicase-related motifs III and IV is longer, and the Ia, III, IV, V, and VI motifs have a specific conserved character containing other conserved blocks (Dü Rr et al. 2006; Flaus et al. 2006).

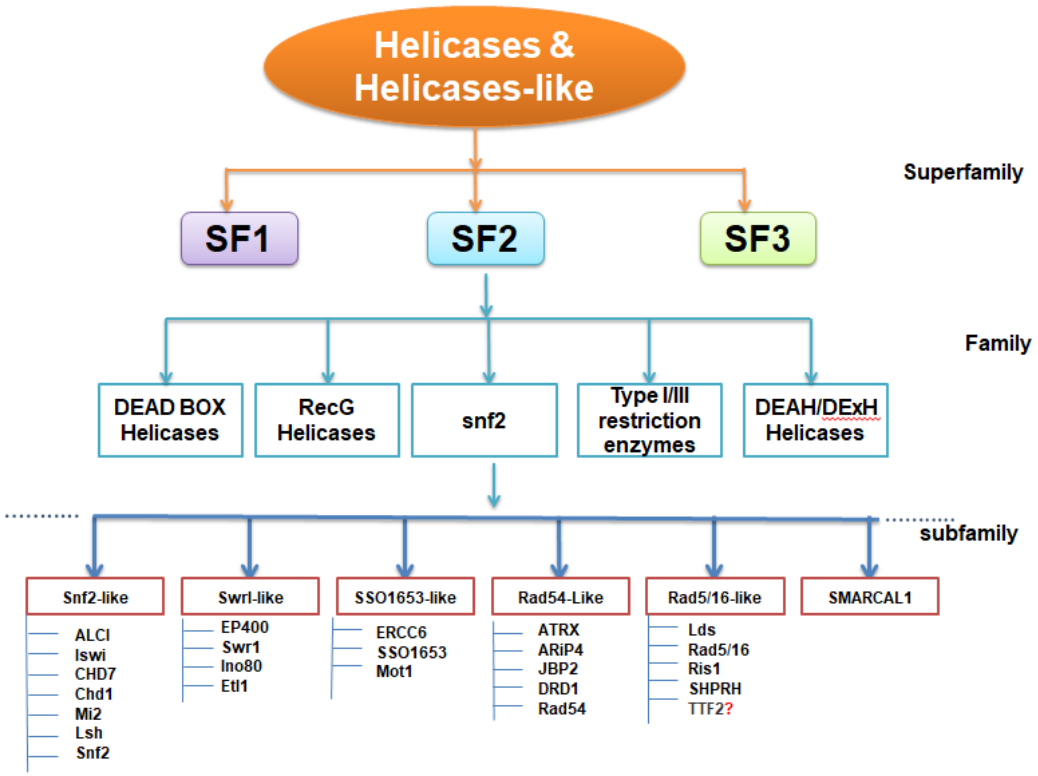

Figure 1.1 Helicases-Like proteins.
Representation of subfamilies within the Snf2 family which belongs to the SF2 superfamily based on conserved structure. Lds is a protein within the Rad5/16-like subfamily (Flaus et al. 2006; Flaus and OwenHughes 2011; Prasad, Lennartsson, and Ekwall 2015; Zhang et al. 2019), but TTF2 is still undefined. Adapted from (Flaus et al. 2006; Prasad, Lennartsson, and Ekwall 2015).

1.1.2 TTF2 and Lds are putative orthologs and belong to the Snf2 family

The ttf2 gene localizes on the first chromosome (1p13.1), has 25 exons (TTF2 transcription termination factor 2 [*Homo sapiens* (human)] - Gene - NCBI 2020), and the predicted molecular weight of its protein is around 130 KDa (ExPASy 2020a). In the fly, the *lds* gene is found on the third chromosome (3R:7), with 5 exons (lds lodestar [*Drosophila melanogaster* (fruit fly)] - Gene - NCBI 2020) and a predicted molecular weight of 118 KDa (ExPASy 2020b).

The similarities in their comparative amino acid sequence analysis, reinforced by their published biochemical activities indicate that TTF2 and Lds are putative orthologs (DRSC - DRSC Integrative Ortholog Prediction Tool 2020; M. Liu, Xie, and Price 1998). The first evidence as orthologs came from the Price Lab, where sequencing information showed that the TTF2 in humans and Lds in *Drosophila* are similar proteins (M. Liu, Xie, and Price 1998). Based on TTF2 and Lds amino acid sequences, these proteins are classified as Snf2 family proteins (Flaus et al. 2006; Girdham and Glover 1991; M. Liu, Xie, and Price 1998) (see Figure1.3). Both have the seven characteristic helicase motifs located on the C-terminal and do not exhibit helicase activity (M. Liu, Xie, and Price 1998). In addition, they contain other seven conserved motifs that belong exclusively to the Snf2 family (Pfam: Protein: TTF2_DROME (P34739) 2020; Pfam: Protein: TTF2_HUMAN (Q9UNY4) 2020), with 31% identity and 53% similarity to these family members (M. Liu, Xie, and Price 1998) (see Figure 1.2 and Table 1.1). Also, both proteins have a sequence motif of the family of helicases' DEAH box, localized at positions 603 - 606 and 737 - 740, respectively. In addition, TTF2 and Lds share other three conserved motifs that do not belong to Snf2 family, but instead belong exclusively to TTF2 and Lds sequences (M. Liu, Xie, and Price 1998) (see Figure 1.2). Further, they have a predicted nuclear localization signal that can be detected between the two helicase motifs I and Ia (M. Liu, Xie, and Price 1998) (discussed in more detail in section 1.2.1).

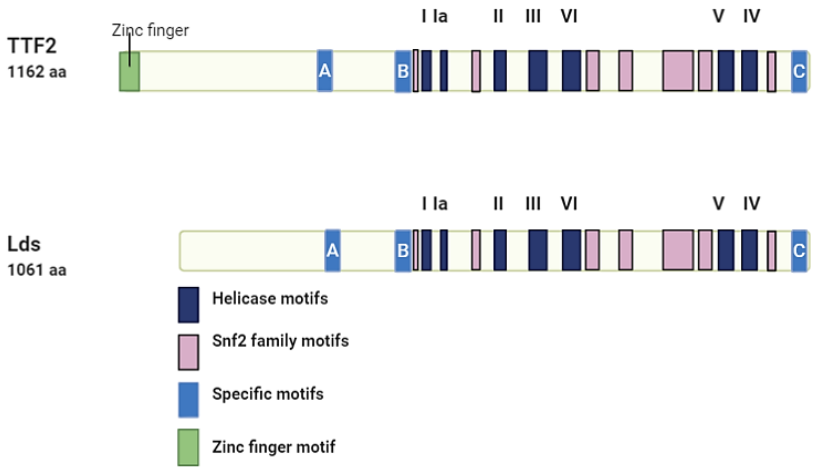

Figure 1.2 Shared motifs between TTF2 and Lds.
Seven helicase motifs (dark blue) are commonly shared between TTF2 and Lds as well as seven motifs that belong to the Snf2 family (light purple), and three specific motifs among each other (light blue). In green is represented a zinc finger motif. Adapted from (M. Liu, Xie, and Price 1998) and UniProt.

Lds is classified as a member of the Rad5/16-like subfamily (see Figure 1.1) (Flaus et al. 2006; Flaus and OwenHughes 2011; Prasad, Lennartsson, and Ekwall 2015; Zhang et al. 2019). This classification is based on the sequence of the helicase like region, and the variety of the length of the spacing between the helicase motifs within this subfamily and the major conserved blocks (Flaus et al. 2006). There are some similarities between Lds and Rad54, a well-characterized protein that belongs to Rad54-like subfamily. Rad54 has a region that forms the protrusion 2 which is a part of the protein that interacts with DNA. Protrusion 2 includes two α-helices (α 17 and α 18) and a short stretch of amino acids. Between the (α 17 and α 18) helices there is an interconnecting region (Szalontai et al. 2009). Lds protein shows the presence of protrusion 2 with both helices (α 17 and α 18) conserved, but the interconnecting region is longer than the one in Rad54 (Szalontai et al. 2009).

When comparing humans' full-length protein TTF2 and the *Drosophila melanogaster* protein Lds, using Blastp, they are 39% identical and 56% similar and only show 11% of the Gaps (no similarity) (see Figure 7.1) (NCBI Blast:sp|Q9UNY4|TTF2_HUMAN Transcription termination... 2020). Their conserved helicase ATP binding domain has 46% identity and 63% similarity (see Figure 7.2) (NCBI Blast:sp|Q9UNY4|583-786 2020), also their highly conserved Helicase C-terminal domains are 52% identical and 74% similarity (see Figure 7.3) (NCBI Blast:sp|Q9UNY4|995-1157 2020). TTF2 sequence exhibits an alignment with the GRF zinc finger sequence (Pfam: Protein: TTF2_HUMAN (Q9UNY4) 2020), but this sequence is not present on the structure of Lds (Pfam: Protein: TTF2_HUMAN (Q9UNY4) 2020) (see Figure 1.3). Moreover, the NH2 terminal domain of TTF2/Lds seems to be unstructured while the “helicase” domains are highly structured (M. Liu, Xie, and Price 1998).

While Lds has been classified by different studies, it remains undefined to which particular subfamily TTF2 belongs. However due to the similarity of the helicase-like region between TTF2 and Lds, I considered TTF2 as a member of the most commonly accepted sub-family for Lds, the Rad5/16-like subfamily (see Figure 1.1).

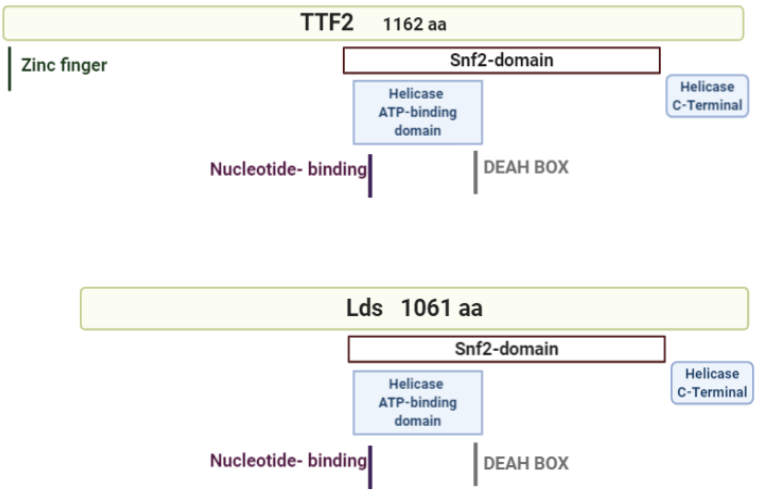

Figure 1.3 An alignment between TTF2 and Lds protein.
The full-length sequences of both proteins show an alignment of Snf2 domain, helicase ATP binding domain, and helicase C-terminal domain. Further to an alignment of the DEAH box motif and Nucleotide-binding motif between both sequences.

Table 1.1 TTF2 and Lds domains identified using UniProt with ID numbers (Q9UNY4) for TTF2 and (P34739) for Lds.

	TTF2		Lds	
Molecular weight	130 KDa		118 kDa	
Domain	**Position (aa)**	**Length (aa)**	**Position(aa)**	**Length (aa)**
Full length		1162		1061
Nucleotide-binding	596 – 603	8	465 – 472	8
Helicase ATP-binding	583 – 786	204	452 – 652	201
Snf2 domain	573-933		442-804	
Helicase C-terminal	995 – 1157	163	891 – 1056	166
DEAH box motif	737 – 740	4	603 – 606	4
Zinc finger motif	5 – 39	35	Not present	

1.2 The localization of TTF2/ Lds is a cell cycle-dependent

To date, two published studies assessed TTF2 and Lds localization using immunofluorescence on fixed cells/tissues (Girdham and Glover 1991; Jiang et al. 2004). However, their localizations using live-imaging techniques were never addressed. Raquel Oliveira's Lab is currently studying both localizations using live imaging approaches.

From imaging studies using fixed samples, the localization of TTF2 and Lds appears almost similar, and cell cycle-regulated. During S and G2 phases, both proteins are mainly in the cytoplasm. Shortly upon mitotic entry, in prophase, both translocate rapidly into the nucleus. At the end of telophase, TTF2/Lds are no longer in the nucleus. Despite these similarities, the localization of TTF2 and Lds at the metaphase plate is different.

At early prophase, the staining of TTF2 can be detected in the cytoplasm and the nucleus, and it organizes in a ring-like shape that accumulates around the nuclear membrane. The formed ring disappears when the nuclear membrane starts to disrupt at mid or late prophase. In metaphase, TTF2 is excluded from the metaphase plate and remains excluded during anaphase and telophase (Figure 1.4 A). In addition, as cells progress through mitosis, the fluorescence intensity of TTF2 decreases significantly (Jiang et al. 2004; Jiang and Price 2004). In comparison, Lds in metaphase localizes at the chromatin and adopts a diamond shape - according to the Glover´s observation – where Lds staining axis is aligned with the axis of the mitotic spindle, yet Lds distribution does not extend to the poles of the spindle, indicating that the spindle envelope gives a boundary to the distribution of the Lds. When sister chromatids are fully separated at the end of anaphase, Lds localizes over the chromatin and only after is cleared from the region between the two daughter nuclei (Girdham and Glover 1991) (Figure.1.4 B).

Experiments from the Oliveira Lab assessed the localization of Lds *in vivo*, using CRISPR/Cas9 technology to add an EGFP at the C-terminal endogenous locus. Live imaging of Lds-EGFP in

Drosophila embryos gives similar results to the fixed imaging performed by Glover and collaborators and shows that Lds is on mitotic chromosomes but we could not observe the diamond diffusion described previously. This can be due to a potential artifact caused by the analysis of fixed tissues. Considering that TTF2 and Lds have a similar cytoplasmic distribution in interphase, but their nuclear distribution is different during mitosis, as Lds is above mitotic chromosomes while TTF2 is excluded, this difference might also be because of an artifact or because each protein has a different function from the another.

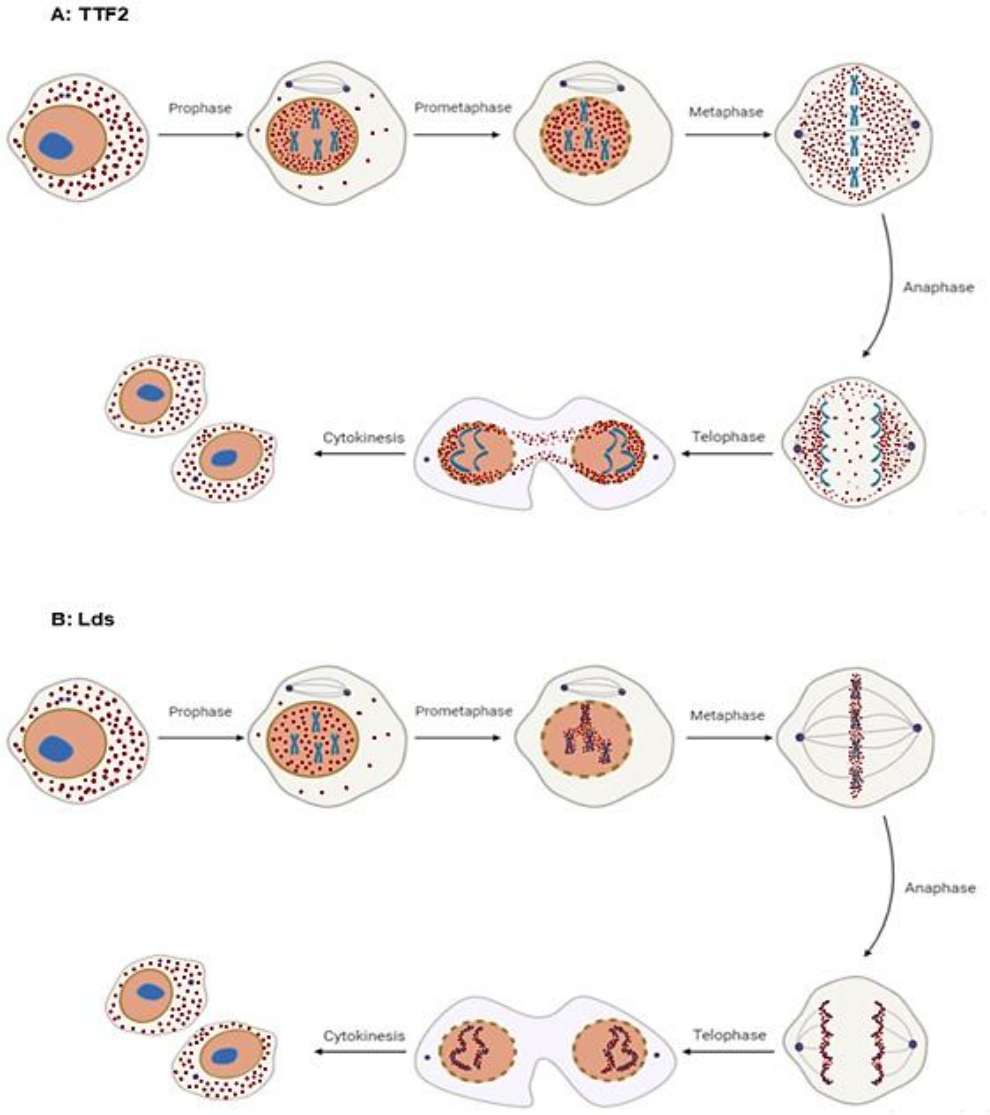

Figure 1.4 The localization of TTF2 and Lds during the cell cycle.
A: TTF2 (red dots) is at the cytoplasm during interphase. At prophase it rapidly translocates to the nucleus. Then, it is excluded from the metaphase plate until telophase. B: Lds (red dots) is in the cytoplasm in interphase. At prophase translocates to the nucleus, where remained associated with the mitotic chromosomes until telophase. Adapted from (Girdham and Glover 1991; Jiang et al. 2004).

1.2.1 TTF2/Lds regulation during the cell cycle

Jiang and colleagues studied the levels of RNA and protein of TTF2 from Hela cells across the cell cycle (Jiang et al. 2004). While total TTF2 RNA levels are stable throughout the cell cycle, western blot analysis shows that TTF2 protein is lower at G1 and its highest amount is detected during mitosis. As cells progress into mitosis, TTF2 protein levels increase gradually. This suggests that the TTF2 protein levels are controlled either by a translational mechanism or protein turnover, but not due to mRNA regulation (Jiang et al. 2004). It remains unknown whether Lds follows a similar trend in cell cycle regulation of protein levels, as no available information exists on Lds mRNA and protein levels throughout the cell cycle.

A computational prediction for TTF2/Lds localization inside the cell (PSORT II Prediction 2020) shows that the probability of TTF2 to be in the nucleus is quite high by 73.9 %, and to be in the cytoplasm is less predicted by 13.0 %. While the probability of Lds localizing at the nuclear or at cytoplasm is equally by 26.1 % (PSORT II server Lds 2020; PSORT II server TTF2 2020). Therefore, it is possible that their localization is regulated differently from each other during the cell cycle.

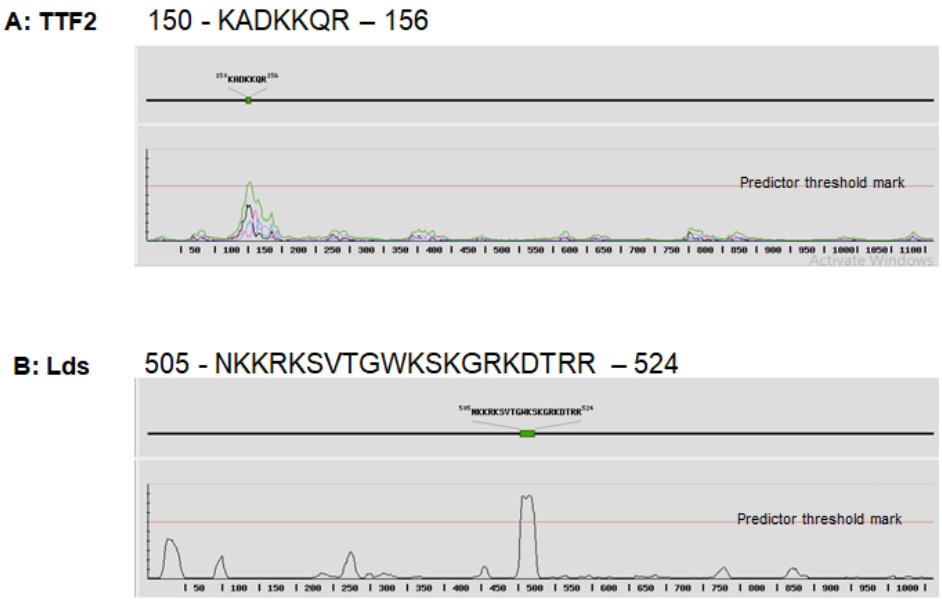

Figure 1.5 Predicted nuclear localization signal for TTF2/Lds.
The predicted nuclear localization signal is shown using the NLStradamus tool. A) Weakly NLS predicted at TTF2 region 150 – 156 (KADKKQR). B) Lds NLS is strongly predicted at region 505 – 524 (NKKRKSVTGWKSKGRKDTRR).

In addition to protein levels, TTF2 may be regulated by the modulation of the ability to access DNA. From the localization studies described in the previous section, it is known that both Lds and TTF2 only enter the nucleus upon nuclear envelope breakdown (NEBD). However, both have a nuclear localization signal (NLS) (Girdham and Glover 1991; M. Liu, Xie, and Price 1998) (see Figure 1.5)

implying that their access to chromatin may be actively regulated, with most protein present in the cytoplasm. While Lds has the NLS signal at 505 – NKKRKSVTGWKSKGRKDTRR – 524 (Figure 1.5 B), TTF2 has it at 150 – KADKKQR – 156, although the NLS prediction does not reach the threshold for TTF2 (Figure 1.5 A). The NLS is known to coordinate the spatial-temporal localization of several proteins between the cytoplasm and the nucleus (reviewed in Chen et al. 2015; Freitas and Cunha 2009; Harreman et al. 2004; Kaffman and O'Shea 1999). Nonetheless, TTF2/ Lds localize at the cytoplasm during interphase which suggests that their NLS must be repressed by an unknown mechanism, to prevent the translocation of TTF2/ Lds to the nucleus in interphase. A possible mechanism could be adjusting the accessibility and affinity of TTF2/ Lds signal recognition by their receptor, due to phosphorylation or dephosphorylation. For instance, one of the known mechanisms of regulatory localization is a phosphorylation site near the NLS sequence which modulates the binding affinity and interaction of the cargo protein with the nuclear transport receptor named importin α (reviewed in Harreman et al. 2004). In the *S. cerevisiae* transcription factor Swi6p, phosphorylation at a site proximate to an NLS reduces the affinity of NLS binding to importin α, that decreases the accumulation of Swi6p in the nucleus (Harreman et al. 2004). On the other hand, some other transcription factors cannot be imported into the nucleus without being phosphorylated, for example, the unphosphorylated Signal Transducer and Activator of Transcription (STAT) proteins are localized in the cytoplasm (reviewed in Freitas and Cunha 2009). However, it remains unclear whether this is the case for Lds and TTF2, namely, if they are phosphorylated or unphosphorylated in different cell cycle phases to control their localization between the cytoplasm and the nucleus. Furthermore, this repression could be due to masking NLS from being activated by either a protein-protein interaction which in some cases can do so (Kachaner et al. 2017), or by changes in the charge or conformation of the cargo protein which also can hide the NLS (reviewed in Poon and Jans 2005). According to (Jiang et al. 2004) *in vitro* findings, a higher TTF2 concentration than the normal increases efficiency of removing transcription elongation complexes from the templates. Therefore, we can speculate TTF2 and Lds retention in the cytoplasm must be regulated to prevent inhibition of transcription until mitosis onset. Thus, the rapid relocation of TTF2/ Lds at NEBD to the DNA could be a way to increase TTF2/ Lds concentrations to the required levels to ensure the removal of mitotic transcription elongation complexes.

1.3 Conclusion

TTF2 and Lds belong to the Snf2 helicases-like family. Lds can be classified within the Rad5/16-like subfamily, while it is yet uncertain if TTF2 goes within the same subfamily. Based on the full-length sequence analysis, TTF2 and Lds are putative orthologs, where they are 39% identical and 56% similar. Both proteins share seven characteristic helicase motifs (mostly located on the C-terminal) and seven Snf2 motifs located on the N-terminal domain, and other three specific conserved motifs among each other. Both have conserved nucleotide-binding and helicase ATP binding domains, besides a sequence motif of the family of helicases' DEAH box. TTF2 sequence shows an alignment with the GRF zinc finger sequence which is absent from Lds. The localization of TTF2/ Lds is cell cycle-dependent. While these two proteins have a similar cytoplasmic distribution in interphase, TTF2

is excluded from the metaphase plate and remains excluded during anaphase and telophase while Lds localizes over the chromatin from NEBD until telophase. The regulation of TTF2/Lds during cell cycle remains unknown, as well as the regulation of its predicted NLS sequence.

Sequencing information is a major premise to support the idea of TTF2 and Lds are orthologs. In the next sections, I will review the literature to discuss their shared functional aspects namely in transcription and cell division (section 2&3, respectively). While they behave similarly *in vitro*, displaying similar *in vitro* transcription termination activities (M. Liu, Xie, and Price 1998) their mutual contribution in maintaining faithful chromosomes segregation is less conclusive.

2 The role of TTF2/ Lds in transcription

2.1 An overview of transcription

In eukaryotic organisms, gene expression is a highly complex process involving nuclear and cytoplasmic steps. It encompasses transcription, generating the messenger RNA (mRNA) which demands further steps as 5′-end capping, splicing, 3′-end cleavage, and polyadenylation. Finally, exporting the mRNA to the cytoplasm, where protein translation occurs. The mRNA will eventually decay, depending on its stability (reviewed in Das, Sarkar, and Das 2017; Maniatis and Reed 2002; Maquat 2004; Rambout, Dequiedt, and Maquat 2018).

Transcription is an extremely coordinated and regulated process (Shandilya and Roberts 2012). It involves the copying of a sequence of the deoxyribonucleic acid (DNA), that is used as a template to produce ribonucleic acid (RNA), by several RNA polymerases enzymes Pols I, II, and III that are responsible for synthesis of rRNA, mRNA, and tRNA with small ribosomal RNA, respectively (Cramer 2019; Deutschman 2005). RNA polymerases do not need a primer to initiate transcription, and they construct the RNA strand in the 5' to 3' direction (reviewed in Mercadante and Mohiuddin 2020). In this monograph, will focus on Pol II, since it is the polymerase responsible for the transcription of all protein-coding genes in the nucleus (reviewed in Calvo and Garc 2012; Fuda, Ardehali, and Lis 2009).

The cycle of transcription involves three stages: initiation, elongation, and termination (Deutschman 2005; Saunders, Core, and Lis 2006), and multiple rounds of these phases take place in a cyclic pattern (Shandilya and Roberts 2012).

The starting of transcription demands unwinding of the DNA double helix, partially, to reveal the nucleobases on each DNA strand, and form the transcription bubble (Alberts et al. 2002c; Clark, Pazdernik, and McGehee 2019). The initiation of the transcription requires the reorganization of the promoter region (situated at the beginning of a gene), to allow RNA polymerase access (Fuda, Ardehali, and Lis 2009). Within the promoter complex, there are DNA sequence repeats of adenine and thymine named the TATA box, that is located upstream of the transcription start site (Reviewed in Mercadante and Mohiuddin 2020; Tong Ihn Lee and Young 2000; Umarov and Solovyev 2017), and this box assists in DNA melting into separate strands since TA binding is weaker than GC base pairs (Clark, Pazdernik, and McGehee 2019).
However, the eukaryotic RNA polymerase II (Pol II) is unable to initiate transcription alone, and hence, needs auxiliary factors (Nikolov and Burley 1997; Woychik and Hampsey 2002). Thus, transcription of protein coding-gene demands the assembly of a large multiprotein–DNA complex (Nikolov and Burley 1997) named a pre-initiation complex (PIC). This complex consists of the DNA template, Pol II, and five general transcription factors (GTFs). The formation of this complex starts by the binding of a subunit of Transcription factor II D (TFIID) to the TATA box and then recruits the

Transcription factor II B (TFIIB), the unphosphorylated complex of Pol II with Transcription factor II F (TFIIF), Transcription factor II E (TFIIE), and Transcription factor II H (TFIIH). After promoter recognition and transcription initiation, the carboxyl-terminal domain (CTD) of the largest RNA Polymerase II subunit is phosphorylated to facilitate promoter clearance and entering into the elongation phase (Woychik and Hampsey 2002).

The efficiency of transcription can be boosted by short sequence binding regions, called enhancers, which can be located far or close from their target gene, either upstream or downstream. Enhancers are not necessary for transcription initiation, but increase transcription by binding to transcription factors (TFs) (Reviewed in Pennacchio et al. 2013). TFs bind to the sequence-specific DNA region where transcription will occur (Lambert et al. 2018), in addition to both enhancer and promoter (Tippens, Vihervaara, and Lis 2018). Multiple transcription factors are bound cooperatively to individual enhancers (Panne 2008) and regulate transcription through physical contacts that involve the creation of a DNA loop (Krivega and Dean 2012). Furthermore, TFs are generally classified into two main classes: the one that controls the initiation and the other that controls elongation. Exceptionally, some transcription factors may contribute to monitoring both transcription initiation and elongation (Lee and Young 2013).

Even though initiation is a switch “on” for Pol II (David H. Price 2008), producing the mRNA requires a higher-level of regulation during elongation. Transcriptional elongation can be divided into two stages, the early elongation, followed by the productive elongation. Upon initiation, Pol II starts adding the nucleotides at the 3´ ends of the RNA nascent transcript, repeatedly. The rate of elongation by RNA Pol II is not constant throughout a gene. For instance, in metazoans, Pol II initially transcribes between 20–60 nucleotides of the transcript (Kwak and Lis 2013) and then pauses (Ardehali et al. 2009). It is therefore common that Pol II initiates transcription at the 5′ end of a gene without building a full-length of mRNA (Guo and Price 2013). Genome-wide studies in human and *Drosophila* proved that many genes have a poised Pol II around 50 nucleotides downstream of the transcription start site (TSS) (Guenther et al. 2007; Muse et al. 2007; Zeitlinger et al. 2007). Furthermore, the majority of engaged Pol II in human cells are promoter-proximal paused Pol II that is considered as a reservoir able to respond to activation signals (Guo and Price 2013; Guo, Turek, and Price 2014), and only under activation circumstances, Pol II is released to produce full-length transcripts (Hargreaves, Horng, and Medzhitov 2009). The paused Pol II necessitates some time to resume elongation, and there is a dwell time range at specific pause sites. Therefore, if the polymerase neither leaves the pause site nor initiate termination during that specific time window, it is named arrested (Reviewed in Guo and Price 2013). However, the escape from pause into productive elongation is a very highly regulated process, dependent on the kinase activity of the positive elongation factor b (P- TEFb), which mediates the pause release by the CTD phosphorylation and contrary to the effects of the negative elongation factors (Marshall et al. 1996; Peng et al. 1998; Price 2000).

Maintenance of paused Pol II is regulated by numerous factors, including the negative elongation factor (NELF), DRB sensitivity-inducing factor (DSIF) (Kwak and Lis 2013; Wada et al. 1998), and Gdown1 (reviewed in Cheng et al. 2012; Guo, Turek, and Price 2014). Gdown1 is a substoichiometric

subunit of Pol II that can reduce the efficiency of initiation due to competition with TFIIF (Hu et al. 2006). In contrast, at the time of initiation, Gdown1 is stabilizing Pol II in a paused conformation (Cheng et al. 2012).

2.1.1 Termination of transcription

Transcription termination is the last step in the transcription cycle. This irreversible step involves the release of nascent transcripts and Pol II from the DNA template (Reviewed in Kuehner, Pearson, and Moore 2011; Lemay and Bachand 2015).

The transcription termination pathways are complicated and not very well understood, and diverse in both yeast and metazoans. Transcription termination by Pol II does not happen at a conserved position or specific distance from the 3′-end of the gene (reviewed in Richard and Manley 2009). It can take place upstream or downstream of genes and within open reading frames (ORFs) (reviewed in Kuehner, Pearson, and Moore 2011). Transcription termination most frequently occurs at the end of the gene, as it is in most of the protein-coding genes. Pol II termination downstream of the 3′ end necessitates a functional Polyadenylation signal (PAS) and it is always combined with mature 3′ end processing (Proudfoot, Furger, and Dye 2002). Additionally, transcription termination may happen prematurely, even before reaching the 3′ ends of genes, called premature transcription termination (reviewed in Eaton and West 2020; Guo and Price 2013). Additionally, transcription termination can be at the promoter-proximal region, where in the absence of P-TEFb, a large number of polymerases terminate their activity near the promoter in a known process named abortive elongation (Guo and Price 2013; N F Marshall and Price 1992).

Termination of transcription at the end of the gene is the most studied one (reviewed in Eaton and West 2020). Most protein-coding genes have a highly conserved PAS. This means a highly conserved 3′ end of mRNA to guide Pol II termination. A PAS includes a 5′-AAUAAA-3′ motif, followed by a G/U-rich sequence towards the 3′ end. Once Pol II transcribes the PAS (reviewed in Shandilya and Roberts 2012), the PAS will be recognized by a complex called cleavage and polyadenylation (CPA). CPA combines the assembly of different factors such as cleavage and polyadenylation specificity factor (CPSF), cleavage stimulatory factor, and the cleavage factors I and II sub-complexes (reviewed in Eaton and West 2020). The association between CPA and PAS results in pre-mRNA cleavage, followed by polyadenylation of the upstream mRNA (reviewed in Eaton and West 2020; Shandilya and Roberts 2012). The PAS-dependent mechanism can be explained by two different models: the first model is the allosteric/anti-terminator which proposes that the transcription of PAS induces the dissociation of elongation factors from Pol II, or it causes a conformational change within Pol II that leads to the termination (reviewed in Eaton and West 2020); the second model is the Torpedo Model. This process relies on the exoribonuclease activity of Xrn2 (in human and Rnt1 in budding yeast) (reviewed in Eaton and West 2020; Lemay and Bachand 2015; Woychik and Hampsey 2002), where their depletion leads to a termination defect (Kim et al. 2004; West, Gromak, and Proudfoot 2004). In this model, PAS cleavage provides an unprotected RNA 5′ end recognized by the 5′→3′ exonuclease. This enzyme degrades the Pol II-associated RNA promoting the termination

by the release of Pol II from the DNA template (reviewed in Eaton and West 2020; West, Gromak, and Proudfoot 2004).

However, a unified model that combines both models has also been suggested (reviewed in Eaton and West 2020). There are some evidences that indicate that Xrn2 has a general Pol II termination activity that is not limited to the PAS-dependent mechanism (reviewed in Eaton and West 2020) and may therefore also operate in other transcription termination events. Indeed, Xrn2 and Rnt1 terminate *in vitro* Pol II activity in the absence of a PAS sequence in the transcript (Park, Kang, and Kim 2015).

Until now, the mechanism of promoter-proximal and premature termination are not adequately explained, maybe because Pol II is unstable and exposed to many destabilizing factors in these regions (reviewed in Eaton and West 2020). The selection of each specific pathway may be determined by the combination between transcription termination factors, the distinct signals on the transcripts, and the particular phosphorylation patterns of the CTD of the largest subunit of Pol II (RPB1) (Kuehner, Pearson, and Moore 2011; Porrua and Libri 2015; Shandilya and Roberts 2012). CTD consists of heptad repeats (Tyr1, Ser2, Pro3, Thr4, Ser5, Pro 6, Ser7) (Schüller et al. 2016), and acts as a scaffold for the recruitment of termination complexes (Kuehner, Pearson, and Moore 2011). The phosphorylation state of the CTD may participate in promoting an alternative mechanism of termination. Genome-wide analysis revealed that phosphorylated status of the CTD Pol II changes its affinity for binding proteins (Kuehner, Pearson, and Moore 2011). It has been noticed at the 5′ end of genes that the ratio of Ser5/Ser2-phosphorylated CTD is higher than it is at the 3′ end, because at initiation time the CTD is most phosphorylated on Ser5 (Kuehner, Pearson, and Moore 2011). Once Pol II progresses through elongation the ratio changes and the phosphorylation on Ser2 increases. When Pol II proceeds to the 3′ end, the Ser2 phosphorylation will be most prevalent (Kuehner, Pearson, and Moore 2011). The Ser2-phosphorylated CTD has an essential role in both 3′ processing and transcription termination, where the accumulation of Ser2 phosphorylation at the CTD notifies the termination machinery for effective cleavage and polyadenylation of pre-mRNAs (Shandilya and Roberts 2012).

Transcription termination in non-coding RNAs that are transcribed by Pol II, for example in the small nuclear RNA (snRNA), has a different pathway from transcription termination in protein-coding genes. This pathway known as sen1-dependent termination was discovered in the yeast *Saccharomyces cerevisiae*, and it requires the actions of a helicase (RNA and DNA helicase Sen) with other key factors, where Sen1 may disrupt the Pol II active site and dissociates polymerase complexes in DNA-dependent ATPase activity (Kuehner, Pearson, and Moore 2011).

After termination, the phosphorylated Pol II is recycled into its unphosphorylated form, allowing the GTFs and Pol II to launch the second round of transcription (reviewed in Woychik and Hampsey 2002). Pol II recycling is important in order to avoid transcription interference between neighboring genes (Rondón et al. 2009), and it prevents the generation of antisense RNAs that can interfere with produced pre-RNA and cause aberrant gene expression (Richard and Manley 2009).

In the next sections, I will discuss the described functions of TTF2/Lds in transcription termination.

2.2 *In vitro* studies support that TTF2/ Lds display transcription termination activity

The transcriptional functional role of TTF2/Lds is widely explored *in vitro*, but limited to a single study *in vivo*, focused on TTF2. The first isolation of Lds protein was from *Drosophila* Kc cells nuclear extracts as one of several factors that affect the transcription of the promoter-containing DNA templates. Lds was recognized for its ability to repress the generation of long Pol II transcripts (Price, Sluder, and Greenleaf 1987). Then Lds was found to terminate the transcription in an ATP-dependent fashion without having a helicase activity (Xie and Price 1996, 1997, 1998). In 1998, gene sequence allowed the identification of Lds human putative homolog, TTF2. TTF2 also terminates *in vitro* transcription in an ATP-dependent manner (Liu, Xie, and Price 1998). The termination activity of TTF2 was investigated on Pol II and Pol I (Hara et al. 1999) and then it's effect on splicing. (Leonard et al. 2003). After that, Price and colleagues showed the role of TTF2 on MTI *in vivo* (Jiang et al. 2004; Jiang and Price 2004). In 2012, two interesting studies were published, one supporting further the role of TTF2 on premature termination (Brannan et al. 2012); while the other showed TTF2 as a part of a complex that affects the HIV-1 production. And only later, the regulation of TTF2 by Gdown1 was addressed (Guo, Turek, and Price 2014).

In the next sections, the role of TTF2/ Lds on transcription termination and their ATPase activities will be discussed in detail.

2.2.1 The transcription termination activities of TTF2/ Lds

In vitro assays showed that TTF2/ Lds release the nascent transcripts – with a length range between 10-75 nucleotides (nt) – from the early elongation complexes, and this release activity from the DNA template is proportional to increasing amount of TTF2/ Lds in the assay (Jiang et al. 2004; M. Liu, Xie, and Price 1998; Xie and Price 1996, 1997). In a comparative study, TTF2 can release nascent transcripts (20nt) with equally high efficiency from *Drosophila's* early elongation complex. However, releasing of human transcripts by Lds is less efficient than TTF2 (M. Liu, Xie, and Price 1998). Despite these differences, TTF2 and Lds are clearly involved *in vitro* premature transcription termination (Brannan et al. 2012; Jiang et al. 2004; Xie and Price 1996). Interestingly, Lds also releases nascent transcripts shorter than 10nt indicating a possible role of it on abortive elongation (Xie and Price 1996), but TTF2 activity on such a short transcript has never been tested. Concerning long transcripts, while it has been mentioned that Lds releases transcripts up to 112nt (data not shown (Xie and Price 1997)), TTF2 can release transcripts up to 300nt (Jiang et al. 2004). This indicates that TTF2/ Lds can release elongation complexes that are functionally subjected to the transition into productive elongation. However, regardless of its releasing transcripts activity, an increasing amount of Lds suppresses gradually the appearance of long transcripts from a

transcription reaction with the promoter-containing DNA template, suggesting it inhibits transcription (Xie and Price 1996).

Transcription termination studies are broader explored for TTF2 than Lds. TTF2 has a transcription release activity on both Pol I and Pol II (Hara et al. 1999; Jiang et al. 2004), but how Lds affects any RNA polymerase has never been stated. TTF2 disrupts the ternary complex of RNA Pol II, as well as Pol I, that stalled at a thymine dimer and releases both polymerases from the template (Hara et al. 1999). Furthermore, the footprinting strategy illuminates equal results where TTF2 releases RNA Pol II and Pol I stalled at T<>T (Hara et al. 1999). In addition, the release activity by TTF2 was further tested on the Pol II ternary complex stalled on the template due to “nucleotide starvation” (Hara et al. 1999). Using pulse-chase assay, it was shown that in absence of TTF2, almost half of the short transcripts are converted to produce full-length Pol II transcripts (330nt) after supplementing with nucleotides. This means the ‘’nucleotide starvation’’ did not affect the ternary complex of these transcripts. In contrast, in presence of TTF2, the majority of the transcripts remained truncated (Hara et al. 1999). Curiously, TTF2 could not release *E. coli* RNA polymerase stalled by nucleotide starvation, suggesting that release of a stalled RNA polymerase is a specific property of the human TTF2 and Pol II (Hara et al. 1999). In particular to Pol II, it has been shown, using HeLa nuclear extracts, that an increased amount of recombinant TTF2, in the presence of ATP, causes the release of Pol II completely, as well as the early transcripts (25 nt) from the DNA template regardless of the phosphorylation state of Pol II (Jiang et al. 2004). This same study approached the functional role of TTF2 using siRNA techniques in HeLa cells. Knockdown of TTF2 causes Pol II retention in mitotic chromosomes, with a remarkable increase in its signal above and around the condensed chromosomes. Given the transition between initiation and productive elongation, Pol II can be detected by specific phosphorylation on its CTD, this study detected the presence of an elongated Pol II which is marked by phosphorylation of serine 2 (Ser2) (reviewed in Bowman and Kelly 2014; Zhou, Li, and Price 2012) . This study supports that TTF2 absence promotes the retention of the RNA Pol II in its productive elongation state (Jiang et al. 2004; Jiang and Price 2004). Interestingly, the retention of Pol II CTD Ser2 phosphorylation at the mitotic chromosomes was no longer seen in a rescue experiment with a GFP-TTF2 expression vector (Jiang and Price 2004). A parallel role of Lds *in vivo* has never been explored. Nevertheless, preliminary data from our lab suggests that Lds is a key factor for the release of nascent transcripts from mitotic chromosomes, where Lds RNAi delays the time of the nascent transcript exclusion (unpublished observations, Catarina Carmo).

Interestingly, a study from 2012 links TTF2 with the 5'-3' exonuclease Xrn2, an RNA processing factor required in transcription termination by Pol II through the ‘‘torpedo’’ mechanism (Sansó et al. 2016). The co-immunoprecipitation results showed that TTF2, Edc3, Dcp1a, and Dcp2 co-immunoprecipitate with Xrn2 (Brannan et al. 2012). In fact, TTF2 was the most abundant protein identified in the pull down. TTF2 interaction with Dcp1a was been previously described in a two-hybrid assay in yeast (Leonard et al. 2003) further supporting TTF2 interaction with the decapping regulatory subunit. Using ChIP-seq in unsyncronised HeLa cells, TTF2, Dcp1a, and Xrn2 were found to be near the TSSs at 5' ends and at genes with 5' peaks of paused Pol II (Brannan et al. 2012). Combined knockdown of Xrn2

and TTF2 causes a significant increase in the density of Pol II at promoter-distal positions, both upstream and downstream of the TSS, enhancing the read-through of transcription; and a decrease in the relative density of Pol II in promoter-proximal positions. Knockdown of Xrn2 or TTF2 alone gave similar results with a reduced redistribution of Pol II. Depletion of decapping factor (Dcp1a) together with TTF2 and Xrn2 mimics the effect of activators that induce transcriptional elongation. These last results resemble what happens upon depletion of the elongation factor Spt5 (Brannan et al. 2012), a component of DRB sensitivity-inducing factor (DSIF) complex that is involved in stabilizing paused Pol II (Wada et al. 1998). Consequently, it was speculated that TTF2 interacts and cooperates with Xrn2 to terminate transcription at the promoter-proximal regions, possibly by the ‘‘torpedo’’ mechanism. TTF2 and decapping factors may work as an entry point for Xrn2 functions by exposing a 5' monophosphate end on the nascent RNA, which provides a suitable phosphorylated substrate for the Xrn2 that will prevent productive elongation (Brannan et al. 2012).

2.2.2 The dsDNA-dependent ATP activity of TTF2/ Lds

TTF2/ Lds require ATP hydrolysis for the transcript release activity and transcription termination (Liu, Xie, and Price 1998; Xie and Price 1997, 1998). These activities are stimulated by dsDNA (Liu, Xie, and Price 1998; Xie and Price 1997), and suppressed in the presence of ssDNA. The ssDNA not only failed to activate the ATP-dependent release activity of TTF2/ Lds (Liu, Xie, and Price 1998; Xie and Price 1998), but also inhibited Lds activity even with presence of an equal amount of dsDNA (Xie and Price 1998). In a comparative study, the ATP hydrolysis activity of TTF2 was slightly higher than Lds activity in the absence of dsDNA (Liu, Xie, and Price 1998). This small difference in the activity between them has never been further explored, so we cannot exclude that this difference arises from technical limitations (e.g. the presence of contaminating DNA) (Liu, Xie, and Price 1998).

Contrary to termination activity studies, as described in the previous section, the dsDNA-dependent ATPase activity is more explored for Lds than TTF2. An example of this is the nucleotide specificity, which has never been addressed for TTF2 (Xie and Price 1997). Only ATP and deoxyadenosine triphosphate (dATP) supported the Lds transcript release activity, while any of GTP, CTP, UTP, dGTP, dCTP, or dTTP did not. Also, this Lds activity does not required any protein cofactors (Xie and Price 1997). Further, a gel mobility shift assay shows an interaction between Lds and dsDNA, where Lds forms a stable complex with the dsDNA in the absence of ATP, and that ATP destabilizes this complex (Xie and Price 1998). This suggests that binding of Lds to dsDNA is a primary condition for the release activity (Xie and Price 1998). Also, the topology of DNA seems to affect the release activities, where, *in vitro*, supercoiled DNA has a slightly higher efficiency on the release activity comparing with linearized DNA (Xie and Price 1997, 1998).

Giving that Lds/TTF2 are effective transcription termination factors, their activities must be strictly controlled to prevent unscheduled transcription termination. Their regulation is poorly explored and it is discussed for TTF2 in the next section.

2.3 Regulation of TTF2 activities

While regulation of Lds transcription activities remain unexplored, the regulation of TTF2 is currently limited to a single functional link between Gdown1 and TTF2 (Guo, Turek, and Price 2014).

Gdown1 is an additional Pol II subunit. In this section, whenever Pol II is in a complex with Gdown1 it will be referred as Pol II(G). *In vitro*, Gdown1 suppresses the effect of activators on Pol II(G) working as direct negative regulator, and therefore transcription initiation of Pol II(G) requires a mediator (Hu et al. 2006; Li and Price 2012). Gdown1 *in vitro* stabilizes the promoter-proximal paused and prevents the stimulation of elongation by TFIIF (Cheng et al. 2012; Guo and Price 2013). Therefore, it is likely that Gdown1 inhibits the binding of TFIIF to Pol II at the level of initiation. When TTF2 *in vitro* termination assays were performed in the presence of Gdown1, the transcription release activity known for TTF2 reduces remarkably (Cheng et al. 2012; Guo, Turek, and Price 2014). This indicates that Gdown1 is blocking TTF2 activity, which cannot be reverted by adding more TTF2 protein amount to the assay (Guo, Turek, and Price 2014).

The N-terminal domain of Gdown1 has a conserved motif (LPDKG) that can be found in humans, *Drosophila,* and among other species, with the exception of *Xenopus,* that contain a similar peptide (LPDRG) (Guo, Turek, and Price 2014). Alignment analysis showed that TTF2 and putative orthologs of many species, including *Drosophila melanogaster,* carried the same conserved motif (LPDKG). This conserved motif across the different species suggests a link between Gdown1 and TTF2 function, possible at their regulation. The deletion of this motif (Gdown1 Δ LPDKG deletion) reduces its ability to inhibit both TFIIF and TTF2 activities (Guo, Turek, and Price 2014). Given the above, the LPDKG motif participates in the inhibition of TTF2. Moreover, this mutation reduces Pol II(G) pool. These results suggest TTF2 transcription termination dependency on Gdown1. In fact, *in vitro*, phosphorylation of Gdown1 on Ser-270 minimizes its ability to block TTF2 (Guo, Turek, and Price 2014).

Although the N-terminal domain of Gdown1 is primarily required to inhibit TTF2 activity, the C-terminal domain on TTF2 has a very slight inhibition effect. However, the full length of Gdown1 has the strongest effect to inhibit TTF2-mediated termination, and none of the domains alone is sufficient to completely do so (Guo, Turek, and Price 2014). This highlights the link between Pol II, Gdown1, and TTF2 that would be interesting to explore further, namely if they interact with each other and if their activity is dependent on each other.

2.4 An additional role of TTF2 in transcription regulation

2.4.1 Splicing and TTF2- CDC5L protein interaction

Splicing is a fundamental procedure within gene expression of eukaryotic cells to produce mature mRNA, where introns (the non-coding region of the genes) must be removed, and the flanking exons

(the coding region of the gene) ligated, resulting in open reading frames that translate into proteins (reviewed in Effenberger, Urabe, and Jurica 2017). From the chemical point of view, splicing requires two transesterification reactions that are catalyzed by the spliceosome complex (Effenberger, Urabe, and Jurica 2017; Will and Lührmann 2011).

It was identified in a yeast two-hybrid system and validated by co-immunoprecipitation analyses an interaction between TTF2 and the cell division cycle 5-like protein (Cdc5L) (Leonard et al. 2003). Cdc5L is a component of the spliceosome complex, a highly conserved protein in eukaryotes from yeast to humans (Wan and Huang 2014), as a crucial pre-mRNA splicing factor (Burns et al. 1999; Grote et al. 2010). *In vitro* immunodepletion of Cdc5L from HeLa nuclear extracts represses the formation of pre-mRNA splicing, but does not block the spliceosome assembly. It has been reported as an oncogene and is highly expressed in some cancers (Martin et al. 2014; Sadikovic et al. 2010). Cdc5L is also known to be a positive regulator of G2/M in mammalian cells (Lei et al. 2000). Overexpression of Cdc5L in mammalian cells leads to a shorter G2 phase, and depletion of Cdc5L delays mitotic entry (Boudrez et al. 2000; Huang et al. 2003). Further, Cdc5L is essential for mitotic progression. Cdc5L depletion produces mitotic arrest, chromosome misalignments while maintaining an active spindle assembly checkpoint which causes mitotic catastrophe (Mu et al. 2014). Truncated TTF2 polypeptide that overlaps with the CDC5L-binding region can also, *in vitro,* inhibit pre-mRNA splicing by disrupting spliceosome assembly (Leonard et al. 2003). This polypeptide (1-118 aa) does not contain any of the conserved Snf2 motifs, and does not show an alignment with Lds protein. From this study, we can conclude that TTF2 disrupts the assembly of the spliceosome complexes and therefore play an additional role in splicing inhibition (Leonard et al. 2003). Interestingly, TTF2 and Cdc5L were both detected in purified spliceosome complexes using gel filtration and chromatography techniques (Leonard et al. 2003). Although this result suggests TTF2 may be a stable component of the human spliceosome complex, TTF2 has not been identified by mass spectrometry in the splicing complexes (Leonard et al. 2003). Therefore, if the TTF2 is a stable component of the spliceosome complex, it may be present in low abundance.

2.4.2 TTF2 function on HIV replication

The human immunodeficiency virus type-1 (HIV-1) belongs to the lentivirus family and is classified as highly pathogenic causing the acquired immune deficiency syndrome (AIDS). The HIV injects its genetic material, a single-stranded RNA, into the cell. After, it uses the viral reverse transcriptase enzyme to convert RNA into DNA in the cytoplasm of the host. The newly viral DNA is integrated into the host's nucleus (reviewed in Craigie and Bushman 2012). Transcription of the integrated provirus is hugely dependent on Pol II of the host and considered crucial for HIV replication (reviewed in Liu et al. 2014). Although the host Pol II complex initiates transcription successfully, producing short viral transcription, these short transcripts are unable to support the viral replication (reviewed in Ott, Geyer, and Zhou 2011), and Pol II fails to move forward on the viral template because of the action of the elongation inhibition factors discussed previously (reviewed in Ramakrishnan et al. 2012). To overcome this physiological mechanism, HIV-1 produces, immediately after infection, a short completely spliced mRNAs encoding a viral essential regulatory protein called the transcriptional

transactivator (Tat) (Karn and Stoltzfus 2012; Ott, Geyer, and Zhou 2011). Tat is responsible for the activation of the transcriptional elongation from the viral long terminal repeat (LTR) (Jones 1997) and boosts Pol II elongation efficiency; thus producing full length of viral transcripts (Kao et al. 1987). However, Tat cannot act alone. P-TEFb interacts with the transactivation response (TAR) element, a conserved structure of RNA stem-loop, to start productive elongation (Ott, Geyer, and Zhou 2011; Ramakrishnan et al. 2012).

P-TEFb itself is composed of catalytic subunits are the regulatory cyclin T (CCNT) subunit that has three forms (CCNT1, CCNT2a or CCNT2b) and the cyclin-dependent kinase 9 (CDK9) (Ott, Geyer, and Zhou 2011; Ramakrishnan et al. 2012). In the case of HIV-1 elongation, Tat can only bind to CCNT1-containing P-TEFb complexes (reviewed in Bisgrove et al. 2007; Ott, Geyer, and Zhou 2011; Tahirov et al. 2010).

A study performed in HeLa cell nuclei extracts to identify CDK9/CCNT1-associated proteins (CCAPs) found 12 distinct multi-protein complexes, among them a complex that included CDK9/CCNT1/TTF2. Depletion of TTF2 causes an enhancement of Tat activation, similar to activation levels than another identified CCAP protein, FBXO11-which plays a role in the phosphorylation-dependent ubiquitination pathway to control NF-kB activity. Co-depletion of TTF2 and FBXO11 caused a combination of CCAP complex depletions that showed the greatest level of Tat enhancement as well as effect on the reactivation of a latent HIV-1 provirus in cells that normally have P-TEFb access limited and therefore the production of HIV-1 limited. Despite the biological complexity and heterogeneity in the catalytic core of P-TEFb, including the presence of TTF2 as CCAPs, these findings suggest TTF2 is relevant for HIV-1 replication. However, the exact nature and contribution of TTF2 in HIV-1 production is far from being understood, although it is already known to not be dependent on NF-kB activity.

2.5 Conclusion

In this section I dissected all published articles with respect to TTF2 and Lds transcriptional function. TTF2, as Lds, releases the transcripts associated with Pol II from the template regardless of their length and in an ATP dependent manner (Jiang et al. 2004; M. Liu, Xie, and Price 1998; Xie and Price 1996, 1997). In addition, TTF2 releases Pol II from the DNA template regardless of the phosphorylation state of Pol II and TTF2 depletion promotes the retention of the RNA Pol II on mitotic chromosomes (Jiang and Price 2004). While this suggests a link between MTI and TTF2, if the same occurs for Lds remains unknown. Despite the weak association with MTI, it is clear that both TTF2 and Lds participate in transcription termination. A possible model involved in transcription termination could be via the ‘’torpedo’’ mechanism. TTF2 cooperates with the exonuclease Xrn2 to terminate premature transcription and preventing the bidirectional elongation of Pol II by the ‘‘torpedo’’ mechanism, in association with decapping factors that might mediate the decapping of 5' monophosphate end of the nascent transcripts. In addition, this mechanism would also include a regulatory step where TTF2 and Gdown1 compete for Pol II binding.

Overall, TTF2 and Lds share a function in transcription termination which supports the idea of being canonical orthologs. In the next section we will discuss their functions in cell division. Nevertheless, it is important to say that TTF2 function has been related with DNA-protein interaction beyond transcription termination (section 2.4), suggesting to be a factor important in transcription regulation. Whether these functions are also conserved in *Drosophila* remains unknown.

3 The role of TTF2/ Lds in cell division

3.1 Overview of cell division and chromosome segregation defects

The cell cycle is the passaging of a cell through a sequence of phases and stages that ultimately coordinate the genome replication and the division of each replicated copy into two identical daughter cells. The cell cycle is sub-divided into two main phases, the interphase and mitotic phase (M phase) (reviewed in Coffman 2004).

In interphase, the cell will be prepared for division through three different stages. As the name suggests, the first gap phase (G1 phase) is the initial stage that occurs after the cell has been formed. At this phase, the needed metabolic changes for cell growth and DNA synthesis preparation take place. Cells manufacture proteins and duplicate their cellular organelles such as the mitochondria and ribosomes. If a cell is committed to dividing, it will go through the G1/S checkpoint and enter the following phase. At the synthesis phase (S phase), the cell replicates its genetic material within the nucleus, and centrosomes are also duplicated (reviewed in Cooper 2000). In some mammalian cells, S phase takes around 10–12 hours, half the time of the cell cycle (Alberts et al. 2002a). While in other organisms it is faster as, for example, *Drosophila* syncytial embryos replicate their entire genome within 3.4 min (reviewed in Seller and O'Farrell 2018; Shermoen, McCleland, and O'Farrell 2010). At the second gap phase (G2 phase), which happens after S phase, the cell prepares for division with more metabolic changes to build what cells need from the cytoplasmic materials in the upcoming phases. Before progression to the next phase, the cell must cross the second checkpoint G2/M. Both gap phases ensure that all conditions are gathered accurately before commitment to replication (S phase) and cell division, through the key checkpoints G1/S and G2/M, respectively (Alberts et al. 2002a).

The mitotic phase (M) comprises two main processes: the nuclear division (karyokinesis) which is also known as mitosis, and cytokinesis. During mitosis, cells rely on very precise mechanisms to avoid segregation defects and therefore ensure chromosome stability at the newly formed daughter cells. Mitosis is also subdivided into different phases. It starts with prophase where the replicated chromosomes condense. At the same time, in the cytoplasm, the mitotic spindle starts to assemble by microtubules emanating from each centrosome. While centrosomes move to opposite poles of the cell, the nuclear envelope breaks down. Then, at prometaphase, microtubules extend in length and will attach to a platform of structural proteins called a kinetochore located in the centromeric region of each sister chromatid (SC). This promotes an orientation of the chromosomes in a way that each centromere of SC will be attached to the opposite spindle pole (reviewed in Alberts et al 2008). In metaphase, chromosomes condense into their distinctive X-shaped and will align at the equatorial plane of the spindle forming the metaphase plate. From the time of replication until this moment, the SC are kept together by the cohesin complex. Cohesin is a tripartite ring composed of the structural maintenance of chromosomes proteins 1 and 3, (SMC1 and SMC3), linked together by a non-SMC

subunit named Scc1/Rad21 (Ivanov and Nasmyth 2005). The cohesion between the SC generates a dynamic tension that counterforces the spindle pulling force to prevent premature SC separation (reviewed in Mirkovic and Oliveira 2017; Oliveira et al. 2010; Tanaka et al. 2000). Because of that and before moving to anaphase, the cells must cross through a third checkpoint, named the M checkpoint or the spindle assembly checkpoint (SAC). SAC ensures the fidelity of the SC segregation through regulating the cohesin cleavage and delaying the anaphase onset by direct inhibition to the Anaphase Promoting Complex/Cyclosome (APC/C) (reviewed in Mirkovic and Oliveira 2017). When all the chromatid pairs have been aligned at the metaphase plate, SAC will be satisfied and allows anaphase onset. During anaphase, the APC/C is activated and causes indirect activation for separase, cleaving centromeric cohesin, where the active APC/C mediates the proteolysis of securin, a separase inhibitor (Alexandru et al 2001). When cohesin is cleaved, the centromeric cohesion between SC is lost. Hence, the SC separate equally and each chromatid is pulled towards the opposite pole. Once each set of chromosomes arrives at its pole and starts to decondense, at telophase, the nuclear envelope reassembles. At same time while nuclear envelope completes reformation, the mitotic spindle disassembles, and the contractile ring starts to appear (Alberts et al 2008; Herdiana 2013). Upon completion of M phase, cytokinesis occurs, where the contractile ring previously assembled and composed of actin and myosin II filaments, contracts at the equator of the cell to form the cleavage furrow. This furrow progresses inwards to physical dissociation of the cytoplasmic components into the two daughter cells (Alberts et al. 2002b).

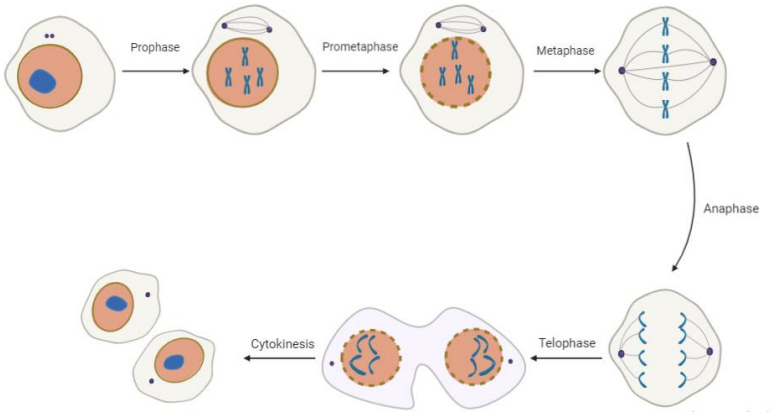

Figure 3.1 The mitotic phase during the cell cycle
Mitosis starts by prophase where chromosomes condense and the spindle starts to appear. At prometaphase, the mitotic spindle is attached to each of the SC. In metaphase, chromosomes condense into X-shaped and align at the metaphase plate. At anaphase, each chromatid of SC is pulled to the opposite pole. Finally, in telophase, chromosomes decondense and the nuclear envelop reforms. Upon completing mitosis, cytokinesis occurs.

If the process of mitosis goes wrong, it can lead to erroneous segregation of chromosomes, and if the newly generated cells survive, will contain an abnormal karyotype. Such mis-segregation of a whole chromosome or a fragment of a chromosome is known as aneuploidy, a hallmark of cancer cells and other genetic diseases such as Down syndrome, and the mosaic variegated aneuploidy (MVA) (Orr,

Godek, and Compton 2015). Various mitotic impaired mechanisms are described to cause these segregation errors and ultimately aneuploidy. These can be generated by the loss of cohesin, SAC imperfections, and kinetochore-microtubule attachment errors or multipolar spindles, which are formed upon over duplication of centrosomes (Orr, Godek, and Compton 2015). Multinucleated or tetraploid cells can be originated because of insufficient cytokinesis or it can be because of slippage which happens when cells are arrested in mitosis but later are able to exit without SAC satisfaction. Additionally, telomere dysfunction gives rise to dicentric chromosomes, which have two centromeres, with a high probability of dividing into one daughter cell or being broken at the dissociation time (Orr, Godek, and Compton 2015). Live and fixed imaging techniques allow the visualization of different types of mitotic errors such as anaphase DNA bridges. DNA bridges can arise from DNA replication problems that remain unresolved in mitosis such as incomplete replication, unresolved DNA catenations, and chromosome structural defects (reviewed in Bakhoum et al. 2017; Orr, Godek, and Compton 2015).The bridges resulting in breaks and non-disjunction in the chromosomes, causing structural aneuploidy (Orr, Godek, and Compton 2015). All in brief, will affect the new proliferation stages of these cells causing stress and loss of genome integrity.

3.2 TTF2/Lds are involved in maintaining faithful chromosomes segregation

3.2.1 Different mutant alleles of Lds cause mitotic segregation defects

In addition to the *in vitro* studies described above (see section 2), Lds function in *Drosophila* has also been addressed through mutagenesis studies on the *lds* gene, although these studies are not extensive. Girdham and Glover were the first to look into Lds function *in vivo*, using 4 mutant alleles of the *lds* gene (Girdham and Glover 1991). Homozygous Lds mutant females laid embryos that died at the early timing of embryogenesis and never hatched, and those mutations were excluded from the paternal rescue. These mutated Lds embryos have a heterogeneous phenotype, suffering from various defects in the nuclear division and the structure of the cell layers. Some Lds embryos cannot develop beyond the syncytial stages and have an uneven distribution of nuclei at the surface of the syncytial. Further, Lds mutations cause chromatin bridges during mitosis. In addition, Lds embryos from hemizygous females have a similar defective phenotype of chromatin bridging. Analyzing the mRNA level of Lds indicates that the transcripts are more predominant in tissues associated with high mitotic activity (ovaries and early embryos), while its level is reduced in males, and not found in the females lacking ovaries (Girdham and Glover 1991). Thus, the Lds transcript was proposed to be an ovary specific of the adult female. For these reasons, Lds is considered a complete maternal effect gene and a protein with a role in the cell division where those Lds mutations are loss of function mutations that cause segregation defects (Girdham and Glover 1991). Overall, the common phenotype characteristic from these embryos is a failure to separate chromatin at anaphase with tangled DNA, bridges, and fragmented DNA (Girdham and Glover 1991), suggesting problems in DNA condensation. Also, Oliveira Lab (A. Boavida, unpublished) results of Lds RNAi, from the

Drosophila wing disc, exhibit various segregation defects such as nondisjunction, lagging, and bridges. However, the cause of the observed mitotic defects remains unclear.

3.2.1 Lds mutation *Horka ^{D}* causes segregation defects

Two other studies in *Drosophila melanogaster* highlighted a specific point mutation in the *lds* gene called *Horka^{D}* as a dominant female-sterile (Fs) mutant (Szabad et al. 1989). *Horka^{D}* is originated through a single-nucleotide change $G^{2424}\rightarrow A$, which results in the replacement of Ala^{777} by Thr in the Lds protein (Szalontai et al. 2009).

Horka^{D}, contrary to Lds loss of function alleles, is implicated in both paternal and maternal effects (Szabad, Máthé, and Puro 1995; Szalontai et al. 2009). Similar to Lds loss of function alleles, homozygous *Horka^{D}* mutation is associated with mitotic defects (Szabad et al. 1989; Szabad, Máthé, and Puro 1995; Szalontai et al. 2009). Embryos with these mutations are characterized by chromosomal arrest through the initial cleavage division, together with lagging chromosomes (Szabad, Máthé, and Puro 1995) and DNA bridges (Szalontai et al. 2009). Moreover, in the eggs, *Horka^{D}* mutation presents defective spindle assembly and triggers the centrosomes displacement from their spindle (Szalontai et al. 2009), as well as, chromosome nondisjunction during oogenesis and early embryogenesis (Szabad, Máthé, and Puro 1995; Szalontai et al. 2009).

In the case of heterozygous *Horka^{D}* female flies, they laid normal egg numbers as the wild type, but more than 90% of these eggs unable to divide normally (Szalontai et al. 2009). Contrary to the previous study where the maternal contribution to the Lds mutation phenotype was equated (Girdham and Glover 1991). *Horka^{D}* sperm also shows abnormalities during the mitotic division, including abnormal condensation, metaphase disarrangement, lagging, and chromosome nondisjunction at telophase. Around 10% of *Horka^{D}* sperm onions (during particular stage called onion-stage spermatid) have a significant changed diameter (either larger or smaller) when compared to regular spermatid nuclei (Szabad, Máthé, and Puro 1995). This chromosomal instability (CIN) during spermatogenesis causes the diplo//haplo mosaic phenotype in embryos including gynandromorphs (Szabad, Máthé, and Puro 1995; Szalontai et al. 2009). In light with this, when wild type females are mated by the sperm from *Horka ^{D}* mutated males, the offspring faced chromosome loss mainly in the paternally derived X chromosome and carried diplo//haplo mosaics, including gynandromorphs ($X^{Y}\underline{X}//X^{Y}O$) (Szabad, Máthé, and Puro 1995). In addition, *Horka ^{D}* drives the loss of the autosomal chromosomes (second, third, and fourth) independently from X chromosome loss. Also, *Horka ^{D}* causes the loss of the XY chromosome simultaneously, but not Y chromosome alone (Szabad, Máthé, and Puro 1995).

In summary, *Horka^{D}* mutation causes defects in males as well as in homozygous (*Horka^{D}/Horka^{D}*) and heterozygous (*Horka^{D}* /+) females. Also, seems that having this point mutation, *Horka^{D}*, is worse than losing the protein, because the null allele rescues the sterility despite it does not rescue the mitotic defects. These conclusions where inferred from a study that address the function of Lds using *horka^{pseudorevertant} (Horka^{rvP}). Horka^{rvP}* are pseudorevertant alleles were induced through *Horka^{D}* P-element-mediated mutagenesis (Szalontai et al. 2009). Interestingly when wild type females fertilized by *Horka^{rvP}* mutated male, only three alleles out of nine *Horka^{rvP}* mutations produce some mosaics

features of *Horka*^{D}, with a very low percentage of $X^{Y}\underline{X}//X^{Y}O$ phenotype amongst the offspring. The mosaics phenotype did not appear in the other six *Horka*^{rvP} *,* and larvae produced by heterozygous females, in the majority of the eggs, were capable of hatching (Szalontai et al. 2009). Among the three mutant alleles that drive mosaic phenotype, they further study one of them, named *Horka*^{rvP2}. Analyzing the phenotype of the *Horka*^{rvP2} clarifies that *horka*^{rvP2} is a null allele that rescues the sterility of the females to a certain degree but is unable to rescue the segregation defects. In particular, they found that *Horka*$^{rvP2\ /-}$ females produce normal eggs that as fertilize as in the wild type. From these fertilized eggs, larvae were disabled and not able to hatch at all. Similar to *Horka*^{D} defects in the eggs and the embryos, *Horka*^{rvP2} causes segregation defects, abnormal spindle, and centrosome disorders (Szalontai et al. 2009).

In summary, *Horka*^{D} is involved in paternal and maternal effects, and causes mitotic defects during oogenesis, spermatogenesis, and early embryos. *Horka*^{rvP2} is a null allele that rescues the sterility of the females but is unable to rescue the segregation defects.

3.2.2 TTF2 depletion causes segregation defects

Addressing the role of TTF2 in cell division specifically during mitosis is even less explored when comparing with the studies performed in Lds, and it is limited to a single study. The knockdown of TTF2 causes a failure in mitotic transcription termination and promotes the retention of Pol II in a productive elongation state (Jiang et al. 2004; Jiang and Price 2004). These defects could be attributed to TTF2, in a rescue experiment with ectopic GFP-TTF2 expression vector active Pol II was no longer retained on mitotic chromosomes but rather Pol II signal was excluded from chromosomes (discussed previously in section 2.) (Jiang and Price 2004). Interestingly, TTF2 depletion causes a rise in the percentage of the multinucleated phenotype by 9%, but did not drive cell death (Jiang et al. 2004). The reason for such mitotic defect caused by TTF2 depletion remains unclear and is clearly very different for the type of errors observed in the Lds studies. Nevertheless, we can speculate that these errors are caused by unappropriated mitotic transcription inactivation, and the failure in MTI probably lead to some mild chromosome structural defects.

Thereby, TTF2 can be a key player in MTI due to its role in repressing transcription elongation complexes from DNA binding during mitosis, and terminating the transcription activity of Pol II (Jiang and Price 2004). Yet, how and to what extent this function contributes to mitotic fidelity remains unclear.

3.3 Conclusions

Overall, while the role of TTF2 is largely unexplored with a single experiment showing a slight increase in the percentage of multinucleated cells after TTF2 depletion (Jiang et al. 2004), Lds role in the maintenance of genome integrity is well supported by evidence (Girdham and Glover 1991; Szabad, Máthé, and Puro 1995; Szalontai et al. 2009). Lds's loss of function mutations cause DNA tangling and bridges (Girdham and Glover 1991). *Horka*^{D} is a point mutation in the *lds* gene with both maternal and paternal effects. In the female, *Horka*^{D} causes sterility because of tangled and

disjunction DNA during oogenesis (Szabad et al. 1989; Szalontai et al. 2009). In the male, *Horka*^{D} causes CIN in the sperm which minimizes the fertility and produce embryos with a mosaic phenotype that suffer from the bridge, lagging (Szabad, Máthé, and Puro 1995).

However, the differences in the observed types of mitotic defects caused by the absence of TTF2 or Lds proteins do not support the idea of both being orthologs regarding their cell division function.

4 Discussion and future prospective

Coordination between transcription silencing and chromosome segregation must be secured for faithful mitosis. Mitosis is characterized by dramatic changes in chromatin architecture where chromosomes condense to secure the faithful separation of the sister chromatids, and most of the transcription machinery is dramatically shut down before it is reactivated in anaphase/telophase at the time of nuclear envelope reformation (reviewed in Kadauke and Blobel 2013; Kang et al. 2020; Palozola, Liu, et al. 2017; Timmers and Verrijzer 2017). However, the details of the MTI machinery are largely unknown, and how the transition from active to silent transcription at the beginning of mitosis is achieved still needs to be fully illustrated, as well as the impact of an improper MTI.

The Oliveira Lab is currently interested in TTF2 function because TTF2 and Lds are recognized as factors involved in the inhibition of transcription (as discussed in section 2) and the loss or mutation of Lds causes highly defective DNA segregation in various *Drosophila* tissues (as discussed in section3). TTF2 and Lds are considered putative orthologues based on their sequencing analysis and their biochemical properties (M. Liu, Xie, and Price 1998). Analyzing their sequencing shows that they are 39% identical and 56% similar. TTF2 sequence is longer and contains a zinc finger motif of GRF-type that is not present in Lds. This motif is most probably involved in DNA binding.

A question is raised of whether TTF2/ Lds is a novel player of the MTI and whether or not the mitotic defects caused by TTF2/ Lds mutations are due to their role in MTI or other distinct functions in mitosis.

To better support this hypothesis, in this monograph, I deeply dissected all published studies of TTF2 and Lds, together with on-going unpublished results on Lds from the Oliveira Lab. For that, I subdivided this monograph into three main parts: section 1 is focused on sequencing data of both proteins and their localization studies; section 2 focuses on the main functions that have been so far published for TTF2/Lds in transcription, and section 3 is dedicated to their functions related with chromosome segregation. In this section I will combine all the studies to address how the main hypothesis that connects this study is supported by the published literature. Nonetheless, it is important to consider that published works on these proteins are limited (a total of seven papers on Lds and eight on TTF2, despite the first publication being in 1987). Moreover, studies on Lds or TTF2 cannot always be compared, since some aspects of the function and roles of TTF2/Lds have been explored solely for TTF2 or Lds.

4.1 TTF2/Lds is a key player in MTI

As stated above, the mechanisms that lead to MTI are largely unknown. MTI could simply result from changes in the phosphorylation levels of general and basic transcription factors, such as TFIID and TFIIB, which prevent transcription initiation (Akoulitchev and Reinberg 1998; Gottesfeld and Forbes 1997; Liang et al. 2015; Long et al. 1998). Post-translational modifications of specific transcription

factors have been suggested to contribute to MTI by reducing their DNA-binding affinity during mitosis (Raccaud and Suter 2018). Another mechanism that has been proposed is that change in the chromosome structure and compaction contribute to MTI. It has been assumed in this aspect that chromosome condensation eliminates the transcriptional machinery and leads to transcription termination (Palozola, Liu, et al. 2017; Timmers and Verrijzer 2017) by building physical barriers (reviewed in Antonin and Neumann 2016). Multiple players as cohesin, condensin, and topoisomerase II contribute to chromosome architecture and condensation during mitosis, and this process has been argued as an essential participant in the spatial repression of mitotic transcription (Gottesfeld and Forbes 1997; Robellet, Vanoosthuyse, and Bernard 2017). Despite that, even with fully condensed chromosomes, some regions, such as centromeres, have their transcription ongoing during mitosis (Perea-Resa et al. 2020). Moreover, the presence of gene bookmarking is now being established. This bookmarking has been seen as a mechanism related to the reprogramming of gene expression and cell identity and involves the retention of transcription factors at promoter elements during mitosis (Kadauke et al. 2012; Kadauke and Blobel 2013; Palozola, Donahue, et al. 2017; Sarge and Park-Sarge 2005). The mechanism of how these specific chromosomal regions can override and expel drivers of MTI is still unknown. However, the ongoing transcription in these regions demonstrates that the previously proposed mechanisms of MTI are somehow more generic and would fail to predict the more recent observations. In fact, MTI requires not only repression of the initiation of new transcription events but also removal of the engaged Pol II machinery in a controlled way within the mitotic chromosomes. There are different models describing what is happening. The first one is that, transcription is suddenly interrupted and aborted, wherever Pol II is and even before reaching the end of the gene (O'Farrell 1992; Shermoen and O'Farrell 1991) by using dedicated machinery of the canonical mechanisms that drive transcription abortion. The other model suggests that genes with transcriptionally engaged Pol II that did not finish their transcription will continue transcription until reaching the end (Liang et al. 2015). This will happen at early prophase, under the control of Mitotic Transcriptional Activation (MTA) which will speed up their transcription to be finished; only after that Pol II will be cleared from the mitotic chromosomes at late prophase (Liang et al. 2015). Then MTI will lead to a global impairment of transcription re-initiation (Liang et al. 2015).

However, it is possible that in order to ensure MTI upon mitotic entry, transcription will not be allowed by the action of particular transcription termination factors that lead to aborted transcription (abortive and/or premature) directly at sites of transcription.

Hypothetically, TTF2/Lds might be one of the critical key players of MTI, which may be supported by the following data. Their localization as a first evidence. During interphase, TTF2/Lds are both localized in the cytoplasm. At the beginning of mitosis there is a rapid redistribution of TTF2/Lds into the nucleus (Girdham and Glover 1991; Jiang et al. 2004). This re-localization coincides with the moment of MTI. In line with their re-localization, the TTF2 protein level rises in mitotic cells compared with cycling and G1 cells (Jiang et al. 2004).

A second evidence is the *in vitro* proved termination activity of TTF2/Lds proteins (M. Liu, Xie, and Price 1998; Xie and Price 1996, 1997). There is compelling evidence that both proteins are involved

in premature transcription termination because they release the nascent transcripts from the early elongation complex (Jiang et al. 2004; Xie and Price 1996). The role of TTF2 in premature transcription termination is further supported by TTF2 - Xrn2 interaction and localization near TSSs to limit the bidirectional movement of Pol II (Brannan et al. 2012). Xrn2 has a known role on premature transcription termination (Contreras, Benkirane, and Kiernan 2013; Wagschal et al. 2012). Thus, TTF2 and Xrn2 might aid each other to insure MTI. Moreover, TTF2 can release Pol II (Hara et al. 1999; Jiang et al. 2004), as well as Pol I (Hara et al. 1999), and TTF2 releases the elongated transcripts regardless of the length and the phosphorylation state of Pol II (Jiang et al. 2004). Also, TTF2 depletion causes the retention of the Ser2 Pol II on mitotic chromosomes in metaphase suggesting its involvement in MTI (Jiang et al. 2004).

Taking the premise of the temporal importance of MTI in the cell cycle, we would expect that when major regulators of MTI are affected, we will have segregation defects due to the mis-regulation of influential members of mitotic fidelity and genome architecture. In line with this, cohesin retention on the chromosomes causes ectopic mitotic transcription outside the centromeric region, mis-localization of Aurora B and Shugoshin, and segregation defects (Perea-Resa et al. 2020). Similarly, mitotic defects are observed upon TTF2/Lds dysfunction (Girdham and Glover 1991; Jiang et al. 2004; Szabad, Máthé, and Puro 1995; Szalontai et al. 2009). It is well proven that mutations in *lds* gene drive a massive chromosome instability and cause segregation defects (Szabad, Máthé, and Puro 1995; Szalontai et al. 2009). Coupled with this, TTF2 depleted HeLa cells have shown a slight rise in multinucleation (by 9%) (Jiang et al. 2004). Upon TTF2 siRNA treatment, cells showed Ser2 Pol II on mitotic chromosomes maintaining active Pol II and nascent transcripts on mitotic chromosomes (Jiang et al. 2004; Jiang and Price 2004). This raises the hypothesis that failures in MTI may compromise mitotic fidelity and affect cell identity.

However, whether or not mitotic defects are associated with defects in MTI, and whether the TTF2/Lds roles in transcription termination are the cause of segregation errors or instead specific functions of these proteins in mitosis, remain to be clarified and require further experiments. The evidence so far is scarce and limited to *in vitro* studies and fixed imaging. The Oliveira Lab is using real-time live imaging to disentangle these open questions. Ongoing work from the Lab has found functional interactions between Lds and DNA structural proteins important for mitotic fidelity namely, Top II, cohesin, and condensin (Oliveira, Boavida, Carmo, Coelho, unpublished data). Further unpublished work from C. Carmo, using an *in vivo* transcription reporter system in *Drosophila* embryos that allows visualization of nascent transcripts with high temporal and spatial resolution, found that Lds is an obligatory factor for effective release of nascent transcripts upon mitotic entry, as Lds RNAi embryos have a delay in the time of the nascent transcript exclusion. Lds RNAi in *Drosophila* embryos and wing disc shows many segregation defects such as nondisjunction, lagging, and bridges which is usually seen in embryos where Top II and condensin are affected (Carmo and Boavida, unpublished data). Also, Lds has been identified in a genetic screen to determine proteins able to rescue mitotic defects caused by premature loss of cohesin (Oliveira, unpublished data). Previous studies from Oliveira Lab have identified specific conditions capable of restoring defects

associated with cohesion loss, in a screen for genes whose depletion modulates *Drosophila* wing development when SC cohesion is impaired. Cohesion deficiency was induced by knockdown of San which results in an intermediate adult wing phenotype that is sensitive to phenotypic modulation (Silva et al. 2018). Lds is classified as a suppressor gene, where combining San RNAi with Lds RNAi restores mitotic fidelity and produces weak wing defects. These results may suggest a third possibility that Lds works in MTI and plays additional roles in mitosis, possibly related to the structure of the chromosomes. To better understand the role of both functions, further experiments are necessary, namely to find a way to suffice one of the functions while the other remains unaffected (e.g using temporal resolution and injection of Lds protein or using mutated versions, as the point mutation Horka^{D}*).* Horka^{D} represents an advantage in this dissection because it is conserved in the TTF2 sequence based on the sequencing analysis (see the alignment results 7.1), and for a comparative study it would be better to do a conserved point mutation. In parallel, studies on Lds human putative orthologue are missing. These included e.g. evaluating whether findings related with Top II, condensin, or cohesin are conserved. Can the exogenous TTF2 expression in embryos or wing discs rescue Lds depletion phenotypes observed in *Drosophila?* The execution of these experiments could clearly evaluate whether TTF2 and Lds are indeed orthologues, in particular with respect to the mitotic defects observed.

4.2 TTF2/Lds regulation during the cell cycle to promote MTI

As mentioned above, TTF2/Lds localization is one of the major evidences supporting the participation of these proteins in MTI regulation. In particular, both proteins are only found to gain access to chromatin during mitotic stages. TTF2 and Lds cellular localization is almost similar in most phases of the cell cycle except during mitosis. TTF2 is chromatin-bound during the early stages of mitosis but excluded from the metaphase plate and remains excluded during anaphase and telophase (Jiang et al. 2004). In contrast, Lds stays on the mitotic chromosomes through all the phases (Girdham and Glover 1991). Both studies addressed TTF2 and Lds localization using only fix immunofluorescent and not live-imaging techniques. Currently, Oliveira Lab is studying Lds localization *in vivo*. Unpublished live imaging data of Lds-EGFP embryos gives similar results to the fixed imaging performed by Glover and collaborators (Carmo unpublished results). Despite the similarity that Lds is on mitotic chromosomes, Carmo did not observe the diamond diffusion described previously (Girdham and Glover 1991). This diamond diffusion is probably an artifact caused by the fixation of the sample. Addressing the TTF2 localization by live-imaging has not been yet performed, which was originally part of my thesis plan, and therefore it remains unknown whether the exclusion from the mitotic chromosomes from human cells described is a limitation of the technique implemented (Reviewed in Sharp et al. 2020; Stradleigh and Ishida 2015).

The localisation of TTF2/Lds to the chromosomes at the time of NEBD may be related with MTI functions. Re-localization is known to be sufficient to alter some protein functions, however, when we consider the increase of protein levels at mitosis and the published *in vitro* assays, we can speculate to be a combination of factors.

In vitro transcript-release activity of TTF2/Lds is proportional to the increasing amount of protein (Jiang et al. 2004; M. Liu, Xie, and Price 1998; Xie and Price 1996). Therefore, this increase in protein may be highly relevant at the moment of NEBD. We can envision TTF2/Lds at MTI would release early elongated transcripts (Jiang et al. 2004; M. Liu, Xie, and Price 1998; Xie and Price 1997). Also, it is experimentally proven that TTF2 releases Pol II from the DNA template regardless of the phosphorylation state (Jiang et al. 2004) and TTF2 depletion promotes the retention of Pol II in its productive elongation state (Jiang et al. 2004; Jiang and Price 2004), which indicates ongoing transcription after depletion of this protein. Therefore, higher TTF2/Lds amounts upon entering mitosis will strength the MTI.

Despite it not being known how TTF2/Lds access the nucleus and are regulated through the cell cycle, we can speculate that since these proteins have a predicted NLS, their NLS is regulated to prevent these proteins from gaining access into the nucleus before mitosis. A typical regulation of NLS localization is through phosphorylation (Jans 1995; Jans and Hübner 1996), which can modulate the importin binding-affinity to NLS (Reviewed in Poon and Jans 2005). This hypothesis is currently under investigation in the lab for Lds. From my sequencing studies described at section 1, TTF2 NLS is below the threshold. Thus, it is important to check experimentally if deletion of the NLS sequences from TTF/Lds will disrupt their transport to the nucleus; without excluding the possibility that TTF2 and Lds access to the nucleus may be differently regulated.

Alternatively, NLS regulation of TTF2/Lds could be aided or combined by other regulatory mechanisms.

In line with that, during mitosis TTF2 activity might be regulated by Gdown1. Tight binding of Gdown1 represses Pol II function. Gdown1 co-localization with Pol II in nuclei is inversely correlated with active gene transcription in human and fly embryos (Jishage et al. 2018). The current model proposes that Gdown1 inhibits the termination of Pol II by blocking TTF2. During mitosis, Gdown1 is predominately inactivated by phosphorylation which reduces its inhibitory effect against TTF2 (Guo, Turek, and Price 2014). A competition model between the Gdown1 and TTF2 for Pol II binding can explain how TTF2 release activity can be facilitated after Gdown1 phosphorylation (Guo, Turek, and Price 2014). In this condition, Gdown1 is less able to bind to Pol II or to perform the blocking action on TTF2. Therefore, the transcriptionally engaged Pol II complexes, perhaps, are terminated by increasing the level of TTF2 in mitosis and gain access to the nucleus. At mitotic exit, TTF2/Lds redistribute toward the cytoplasm of daughter cells rather than being at the nucleus (Girdham and Glover 1991; Jiang et al. 2004). That corresponds to cells requirement of a quick activation of the transcription machinery, where the presence of both transcription and RNA processing factors is very important for the resumption of transcription in the newly formed daughter nuclei (Prasanth et al. 2003). Whether this is exactly the mechanism it is vaguely supported for TTF2 and completely unexplored for Lds. Testing how Gdown1 could affect Lds is important to uncover the regulation process and confirm the previous findings.

All these hypotheses require a deep understanding by both *in vitro* and *in vivo* studies. Nevertheless, we still do not know how Lds protein levels change during the cell cycle and if can be regulated as TTF2. A similar regulation pathway would further support the idea of both proteins being true orthologues.

4.3 New roles of TTF2

Taking into consideration that TTF2 can release not only PolI II but also stalled Pol I from the DNA templates, we cannot exclude that TTF2/Lds termination activity is not limited to MTI, but rather related to general transcription termination with a role at other cell cycle phases (S and G2 phases as a general regulator of gene expression).

In this perspective, although TTF2 has not been detected by immunofluorescence in the nucleus of interphase cells (Jiang et al. 2004); it is possible that in interphase, TTF2 is at nucleus at very low levels or dynamically bound to the DNA. ChIP-seq analysis showed TTF2 localizes near the TSSs at 5' ends and it associates with Xrn2 and mRNA decapping factors (Brannan et al. 2012). These interactions might facilitate a premature transcription termination of Pol II through the ‘‘torpedo’’ mechanism (Brannan et al. 2012). Also, TTF2 can *in vitro* terminate transcripts regardless of the length (Jiang et al. 2004), which implies TTF2 in interphase has a general transcription termination role. This guides the idea that having different layers of regulation across the different cell phases may be due to TTF2/Lds- specific protein interactions.

Moreover, it would be important to clarify whether the *in vitro* findings on Pol I happen *in vivo* and whether TTF2 has indeed a function during interphase. TTF2 disrupts the ternary complex of RNA polymerases that are stalled on thymine dimer, so it might be involved in the preferential DNA repair through the transcription-repair coupling mechanism (Hara et al. 1999). It has been found that the repairing rate of pyrimidine dimers damage, in humans, is faster when the repair pathway is coupled to transcription (reviewed in Douki, von Koschembahr, and Cadet 2017; Hara et al. 1999), Also, the role of Snf2 family proteins on DNA repair is known (Ryan and Owen-Hughes 2011). Therefore, a further detection of TTF2 and Lds role on this possible function would be valuable. If Lds can release RNA complexes, from both *Drosophila* and human, which are stalled upon DNA damage, it might support that TTF2 and Lds are orthologs.

Transcription cannot be seen apart from other RNA processes that can occur in a mutual timeline. Through elongation, it is typical that, the pre-mRNA processing events are initiated immediately at the transcription site, where Pol II couples transcription and pre-mRNA splicing (Hirose and Manley 2000; Proudfoot, Furger, and Dye 2002). Effective processing of nascent pre-mRNA is a prerequisite for mRNA synthesis and exporting mRNA to the cytoplasm (Reed and Hurt 2002). The different hyper-phosphorylation changes that occur on the serine residues within the heptad repeats of the Pol II CTD regulate transcription initiation and elongation and allow the interaction with specific elements such as capping enzyme, splicing factors, and 3′-end cleavage factors (Hirose and Manley 2000; McCracken et al. 1997). According to published literature, TTF2 roles include other essential processes that

contribute to gene expression beyond transcription termination. Protein-protein interactions studies proposed several interacting partners with TTF2, most of them were proteins involved in transcription, DNA binding, and pre-mRNA splicing (Brannan et al. 2012; Leonard et al. 2003). For instance, immunoprecipitation experiments show that TTF2 pulls down other proteins such as Heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) and Heterogeneous nuclear ribonucleoproteins A2/B1 (hnRNP A2/B) (Leonard et al. 2003), which are important proteins for pre-mRNA splicing, mRNA stability, and control of transcription elongation by Pol II (Lemieux et al. 2015). Another example is the Negative cofactor 2 (NC2) protein (Leonard et al. 2003), which is an inhibitor of preinitiation complex assembly (reviewed in Cang and Prelich 2002). Moreover, another identified interaction was between TTF2 and the regulatory protein of the transition G2/M in mammalian cells, Cdc5L, that also is part of the spliceosome complex (Leonard et al. 2003). All of these interactions strongly support a functional role of TTF2 in mRNA splicing, which maybe be important to be addressed in the future. All these interactions, and putative new roles, have been deeply addressed for TTF2; and never tested in Lds. It is relevant to say that TTF2 could have acquired these novel roles beyond the putative roles it shares with Lds.

In addition, TTF2 has been related to viral transcription. TTF2 contributes to HIV replication, where it limits the productive elongation of the virus (reviewed in Fu et al. 2007). A curious fact is that TTF2 is highly expressed in the lymph node which is populated by T and B cells immune cells that prevent pathogen dissemination (Janeway et al. 2001). Given this, it would be interesting to explore TTF2 potential roles in the immune response.

The potential wide range of functions for these proteins needs to be taken into account when addressing their potential role in MTI and mitotic fidelity and its implication in different tissues and cells.

5 Conclusion

Faithful chromosome segregation is fundamental to achieve genome integrity, and the failure to maintain it is the main reason that contributes directly or indirectly to the abnormalities of biological processes that cause several diseases. Also, ensuring the time of initiation and termination of transcription machinery is equally crucial to secure the safety of inherited DNA. Above all, the spatial and temporal coordination between mitotic events and transcription is the primary condition to secure cell identity and survival. Most of the transcription machinery is inactivated during mitosis before a reactivation at the time of mitotic exit. TTF2 and Lds could be crucial players to ensure a spatial transcription inactivation and thereby contribute to mitotic fidelity. Therefore, experimental investigation in TTF2/Lds can clarify the novel role of these proteins in MTI. TTF2 and Lds are considered putative orthologs. If their function is proven to be conversed, namely, in MTI, this will allow us to use *Drosophila* as a model organism for studying the humans' MTI and the consequences of disrupting it. *Drosophila* as a model organism represents several advantages, namely a short life cycle, and easy to manipulate.

In this study, I aimed to dissect all the published studies on TTF2 and Lds to better evaluate what are the functions that corroborate the idea that these proteins are orthologs and participate in MTI. From my review of the literature and the sequencing analysis that I did, they share not only a similar sequence but also one main function – *in vitro* transcription termination. This supports the idea of these proteins to be considered orthologs. Nevertheless, investigation is needed to fill in missing information with respect to their biochemical properties, which sometimes has been solely shown for one protein. At the same time, and more importantly, we need to validate this function by *in vivo* approaches. Concerning their function on cell division and mitotic fidelity, I believe the available published evidence is still not strong enough to state whether TTF2 and Lds indeed share this function. Therefore, and up to now, I do not consider them orthologs on this particular function. Similarly, live imaging experiments such as testing more deeply the type of mitotic defects caused by depletion of TTF2/Lds, and whether the exogenous expression of one of the two proteins can rescue the mitotic defects caused by depletion of the other are necessary to understand TTF2/Lds mitotic functions.

In summary, from all the published studies and on-going experiments at Oliveira lab, I think that TTF2 and Lds are key players of MTI. Besides that, it remains unclear whether their potential role on MTI and impact in chromosome segregation are linked or instead two different functions (transcription termination and mitotic functions). My current vision supports that the two functions are connected to each other.

This page has been intentionally left blank.

6 References

Akoulitchev, Sasha, and Danny Reinberg. 1998. “The Molecular Mechanism of Mitotic Inhibition of TFIIH Is Mediated by Phosphorylation of CDK7.” *Genes and Development* 12(22): 3541–50. www.genesdev.org (October 20, 2020).

Alberts, Bruce et al.

2002a. “An Overview of the Cell Cycle.”

2002b. “Cytokinesis.” https://www.ncbi.nlm.nih.gov/books/NBK26831/ (October 25, 2020).

2002c. “From DNA to RNA.” https://www.ncbi.nlm.nih.gov/books/NBK26887/ (October 12, 2020).

Alexandru, Gabriela et al. 2001. “Phosphorylation of the Cohesin Subunit Scc1 by Polo/Cdc5 Kinase Regulates Sister Chromatid Separation in Yeast.” *Cell* 105(4): 459–72.

Antonin, Wolfram, and Heinz Neumann. 2016. “Chromosome Condensation and Decondensation during Mitosis.” *Current Opinion in Cell Biology* 40: 15–22. https://pubmed.ncbi.nlm.nih.gov/26895139/ (August 29, 2020).

Ardehali, M. Behfar et al. 2009. “Spt6 Enhances the Elongation Rate of RNA Polymerase II in Vivo.” *EMBO Journal* 28(8): 1067–77.

Bakhoum, Samuel F. et al. 2017. “Mitotic DNA Damage Response: At the Crossroads of Structural and Numerical Cancer Chromosome Instabilities.” *Trends in Cancer* 3(3): 225–34. https://europepmc.org/articles/PMC5518619 (November 26, 2020).

Ball, A. R., and K. Yokomori. 2001. “The Structural Maintenance of Chromosomes (SMC) Family of Proteins in Mammals.” *Chromosome Research* 9(2): 85–96. https://link.springer.com/article/10.1023/A:1009287518015 (November 19, 2020).

Bisgrove, Dwayne A., Tokameh Mahmoudi, Peter Henklein, and Eric Verdin. 2007. “Conserved P-TEFb-Interacting Domain of BRD4 Inhibits HIV Transcription.” *Proceedings of the National Academy of Sciences of the United States of America* 104(34): 13690–95. https://pubmed.ncbi.nlm.nih.gov/17690245/ (June 28, 2020).

Boudrez, A. et al. 2000. “NIPP1-Mediated Interaction of Protein Phosphatase-1 with CDC5L, a Regulator of Pre-MRNA Splicing and Mitotic Entry.” *Journal of Biological Chemistry* 275(33): 25411–17.

Bowman, Elizabeth A., and William G. Kelly. 2014. “RNA Polymerase II Transcription Elongation and Pol II CTD Ser2 Phosphorylation: A Tail of Two Kinases.” *Nucleus (United States)* 5(3).

Brannan, Kris et al. 2012. “MRNA Decapping Factors and the Exonuclease Xrn2 Function in Widespread Premature Termination of RNA Polymerase II Transcription.” *Molecular Cell* 46(3): 311–24.

Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, Peter walter. 2008. *M o l e c u l a Br i o l o g y o F*. papers2://publication/uuid/0BF210E2-70B3-4CC9-A859-111EF841687B.

Burns, C. Geoffrey, Ryoma Ohi, Adrian R. Krainer, and Kathleen L. Gould. 1999. “Evidence That Myb-Related CDC5 Proteins Are Required for Pre-MRNA Splicing.” *Proceedings of the National Academy of Sciences of the United States of America* 96(24): 13789–94.

Calvo, Olga, and Alicia Garc. 2012. “RNA Polymerase II Phosphorylation and Gene Expression Regulation.” In *Protein Phosphorylation in Human Health*, InTech. http://dx.doi.org/10.5772/48490 (October 12, 2020).

Cang, Yong, and Gregory Prelich. 2002. “Direct Stimulation of Transcription by Negative Cofactor 2 (NC2) through TATA-Binding Protein (TBP).” *Proceedings of the National Academy of Sciences of the United States of America* 99(20): 12727–32. www.pnas.orgcgidoi10.1073pnas.202236699 (September 23, 2020).

Chan, F. L. et al. 2012. “Active Transcription and Essential Role of RNA Polymerase II at the Centromere during Mitosis.” *Proceedings of the National Academy of Sciences* 109(6): 1979–84. http://www.pnas.org/cgi/doi/10.1073/pnas.1108705109 (October 2, 2020).

Charles A Janeway, Jr, Paul Travers, Mark Walport, and Mark J Shlomchik. 2001. “Survival and Maturation of Lymphocytes in Peripheral Lymphoid Tissues.” https://www.ncbi.nlm.nih.gov/books/NBK27150/ (September 7, 2020).

Chen, Anan, Tara K. Akhshi, Brigitte D. Lavoie, and Andrew Wilde. 2015. “Importin B2 Mediates the Spatio-Temporal Regulation of Anillin through a Noncanonical Nuclear Localization Signal.” *Journal of Biological Chemistry* 290(21): 13500–509.

Chen, Danyang et al. 2005. “Condensed Mitotic Chromatin Is Accessible to Transcription Factors and Chromatin Structural Protiens.” *Journal of Cell Biology* 168(1): 41–54. /pmc/articles/PMC2171683/?report=abstract (October 5, 2020).

Cheng, Bo et al. 2012. “Functional Association of Gdown1 with RNA Polymerase II Poised on Human Genes.” *Molecular Cell* 45(1): 38–50.

Clark, David P., Nanette J. Pazdernik, and Michelle R. McGehee. 2019. “Transcription of Genes.” In *Molecular Biology*, Elsevier, 332–61.

Coffman, James A. 2004. “Cell Cycle Development.” *Developmental Cell* 6(3): 321–27. http://www.cell.com/article/S153458070400067X/fulltext (August 16, 2020).

Contreras, Xavier, Monsef Benkirane, and Rosemary Kiernan. 2013. “Premature Termination of Transcription by RNAP II: The Beginning of the End.” *Transcription* 4(2): 72–76. /pmc/articles/PMC3646057/?report=abstract (November 27, 2020).

Cooper, Geoffrey M. 2000. “The Eukaryotic Cell Cycle.” https://www.ncbi.nlm.nih.gov/books/NBK9876/ (June 25, 2020).

Craigie, Robert, and Frederic D. Bushman. 2012. “HIV DNA Integration.” *Cold Spring Harbor Perspectives in Medicine* 2(7). /pmc/articles/PMC3385939/?report=abstract (October 27, 2020).

Cramer, Patrick. 2019. “Organization and Regulation of Gene Transcription.” *Nature* 573(7772): 45–54.

Das, Subhadeep, Debasish Sarkar, and Biswadip Das. 2017. “The Interplay between Transcription and MRNA Degradation in Saccharomyces Cerevisiae.” *Microbial Cell* 4(7): 212–28.

Davis, J L, R Kunisawa, and J Thorner. 1992. “A Presumptive Helicase (MOT1 Gene Product) Affects Gene Expression and Is Required for Viability in the Yeast Saccharomyces Cerevisiae.” *Molecular and Cellular Biology* 12(4): 1879–92.

Deutschman, Clifford S. 2005. “Transcription.” *Critical Care Medicine* 33(12 SUPPL.): S400–403.

http://journals.lww.com/00003246-200512001-00002 (March 17, 2020).

Ding, Hao et al. 2015. “The Q Motif Is Involved in DNA Binding but Not ATP Binding in ChlR1 Helicase” ed. Maria Spies. *PLOS ONE* 10(10): e0140755. https://dx.plos.org/10.1371/journal.pone.0140755 (August 6, 2020).

Douki, Thierry, Anne von Koschembahr, and Jean Cadet. 2017. “Insight in DNA Repair of UV-Induced Pyrimidine Dimers by Chromatographic Methods.” *Photochemistry and Photobiology* 93(1): 207–15. http://doi.wiley.com/10.1111/php.12685 (November 26, 2020).

“DRSC - DRSC Integrative Ortholog Prediction Tool.” 2020. https://www.flyrnai.org/cgi-bin/DRSC_prot_align.pl?geneid1=45894&geneid2=8458 (August 23, 2020).

Dü Rr, Harald, Andrew Flaus, Tom Owen-Hughes, and Karl-Peter Hopfner. 2006. “Snf2 Family ATPases and DExx Box Helicases: Differences and Unifying Concepts from High-Resolution Crystal Structures.” *Nucleic Acids Research* 34(15): 4160–67. http://www.yeastgenome.org/ (July 20, 2020).

Eaton, Joshua D., and Steven West. 2020. “Termination of Transcription by RNA Polymerase II: BOOM!” *Trends in Genetics* 36(9): 664–75. https://pubmed.ncbi.nlm.nih.gov/32527618/ (October 13, 2020).

Effenberger, Kerstin A., Veronica K. Urabe, and Melissa S. Jurica. 2017. “Modulating Splicing with Small Molecular Inhibitors of the Spliceosome.” *Wiley Interdisciplinary Reviews: RNA* 8(2). /pmc/articles/PMC5253128/?report=abstract (July 31, 2020).

“ExPASy.” 2020a. https://web.expasy.org/cgi-bin/compute_pi/pi_tool1?Q9UNY4@1-1162@average (November 19, 2020).

———. 2020b. https://web.expasy.org/cgi-bin/compute_pi/pi_tool1?P34739@1-1061@average (November 19, 2020).

Ferreira, Helder, Andrew Flaus, and Tom Owen-Hughes. 2007. “Histone Modifications Influence the Action of Snf2 Family Remodelling Enzymes by Different Mechanisms.” *Journal of Molecular Biology* 374(3): 563–79. /pmc/articles/PMC2279226/?report=abstract (August 25, 2020).

Ferreira, Izabela, Tauanne D. Amarante, and Gerald Weber. 2015. “DNA Terminal Base Pairs Have Weaker Hydrogen Bonds Especially for at under Low Salt Concentration.” *Journal of Chemical Physics* 143(17): 175101. http://aip.scitation.org/doi/10.1063/1.4934783 (August 23, 2020).

Flaus, Andrew, David M A Martin, Geoffrey J Barton, and Tom Owen-Hughes. 2006. “Identification of Multiple Distinct Snf2 Subfamilies with Conserved Structural Motifs.”

Flaus, Andrew, and Tom Owen-Hughes. 2011. “Mechanisms for ATP-Dependent Chromatin Remodelling: The Means to the End.” *FEBS Journal* 278(19): 3579–95. http://doi.wiley.com/10.1111/j.1742-4658.2011.08281.x (October 6, 2020).

Freitas, Natalia, and Celso Cunha. 2009. “Mechanisms and Signals for the Nuclear Import of Proteins.” *Current Genomics* 10(8): 550–57.

Fu, Junjiang, Ho-Geun Yoon, Jun Qin, and Jiemin Wong. 2007. “Regulation of P-TEFb Elongation Complex Activity by CDK9 Acetylation.” *Molecular and Cellular Biology* 27(13): 4641–51. http://mcb.asm.org/ (September 7, 2020).

Fuda, Nicholas J., M. Behfar Ardehali, and John T. Lis. 2009. “Defining Mechanisms That Regulate

RNA Polymerase II Transcription in Vivo." *Nature* 461(7261): 186–92.

Ganji, Mahipal et al. 2018. "Real-Time Imaging of DNA Loop Extrusion by Condensin." *Science* 360(6384): 102–5. http://science.sciencemag.org/ (July 15, 2020).

Girdham, C. H., and D. M. Glover. 1991. "Chromosome Tangling and Breakage at Anaphase Result from Mutations in Lodestar, a Drosophila Gene Encoding a Putative Nucleoside Triphosphate-Binding Protein." *Genes and Development* 5(10): 1786–99.

GJ, Huang et al. 2003. "[Screening of New Binding Partners of CIKS with Yeast Two-Hybrid System]." *Sheng wu Gong Cheng xue bao = Chinese Journal of Biotechnology* 19(2): 190–94.

Gorbalenya, Alexander E., and Eugene V. Koonin. 1993. "Helicases: Amino Acid Sequence Comparisons and Structure-Function Relationships." *Current Opinion in Structural Biology* 3(3): 419–29.

Gorbalenya, Alexander E., Eugene V. Koonin, Alexei P. Donchenko, and Vladimir M. Blinov. 1988. "A Novel Superfamily of Nucleoside Triphosphate-Binding Motif Containing Proteins Which Are Probably Involved in Duplex Unwinding in DNA and RNA Replication and Recombination." *FEBS Letters* 235(1–2): 16–24. https://pubmed.ncbi.nlm.nih.gov/2841153/ (December 3, 2020).

Gottesfeld, Joel M., and Douglass J. Forbes. 1997. "Mitotic Repression of the Transcriptional Machinery." *Trends in Biochemical Sciences* 22(6): 197–202. https://pubmed.ncbi.nlm.nih.gov/9204705/ (July 15, 2020).

Grote, Michael et al. 2010. "Molecular Architecture of the Human Prp19/CDC5L Complex." *Molecular and Cellular Biology* 30(9): 2105–19.

Guenther, Matthew G. et al. 2007. "A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells." *Cell* 130(1): 77–88.

Guo, Jiannan, and David H. Price. 2013. "RNA Polymerase II Transcription Elongation Control." *Chemical Reviews* 113(11): 8583–8603.

Guo, Jiannan, Michael E. Turek, and David H. Price. 2014. "Regulation of RNA Polymerase II Termination by Phosphorylation of GDOWN1." *Journal of Biological Chemistry* 289(18): 12657–65.

Gurley, Lawrence R., Ronald A. Walters, and Robert A. Tobey. 1974. "Cell Cycle-Specific Changes in Histone Phosphorylation Associated with Cell Proliferation and Chromosome Condensation." *Journal of Cell Biology* 60(2): 356–864. https://rupress.org/jcb/article-pdf/60/2/356/449026/356.pdf (July 15, 2020).

Hara, Ryujiro et al. 1999. "Human Transcription Release Factor 2 Dissociates RNA Polymerases I and II Stalled at a Cyclobutane Thymine Dimer." *Journal of Biological Chemistry* 274(35): 24779–86. http://www.jbc.org/ (August 26, 2020).

Hargreaves, Diana C., Tiffany Horng, and Ruslan Medzhitov. 2009. "Control of Inducible Gene Expression by Signal-Dependent Transcriptional Elongation." *Cell* 138(1): 129–45.

Harreman, Michelle T. et al. 2004. "Regulation of Nuclear Import by Phosphorylation Adjacent to Nuclear Localization Signals." *Journal of Biological Chemistry* 279(20): 20613–21.

Herdiana. 2013. 53 Journal of Chemical Information and Modeling *Biology*.

Hirano, T., and T. J. Mitchison. 1991. "Cell Cycle Control of Higher-Order Chromatin Assembly around

Naked DNA in Vitro." *Journal of Cell Biology* 115(6): 1479–89. https://rupress.org/jcb/article-pdf/115/6/1479/382200/1479.pdf (July 15, 2020).

Hirose, Yutaka, and James L Manley. 2000. *RNA Polymerase II and the Integration of Nuclear Events*. www.genesdev.org (August 29, 2020).

Hu, Xiaopeng, Sohail Malik, Catalin Negroiu, Kyle Hubbard, Chidambaram Natesa Velalar, et al. 2006. 103 *A Mediator-Responsive Form of Metazoan RNA Polymerase II*. www.pnas.orgcgidoi10.1073pnas.0603702103 (August 3, 2020).

Hu, Xiaopeng, Sohail Malik, Costin Catalin Negroiu, Kyle Hubbard, Chidambaram Natesa Velalar, et al. 2006. "A Mediator-Responsive Form of Metozoan RNA Polymerase II." *Proceedings of the National Academy of Sciences of the United States of America* 103(25): 9506–11. https://pubmed.ncbi.nlm.nih.gov/16769904/ (August 3, 2020).

Ivanov, Dmitri, and Kim Nasmyth. 2005. "A Topological Interaction between Cohesin Rings and a Circular Minichromosome." *Cell* 122(6): 849–60. http://www.ncbi.nlm.nih.gov/pubmed/16179255 (April 3, 2020).

Iwasaki, Osamu et al. 2019. "Involvement of Condensin in Cellular Senescence through Gene Regulation and Compartmental Reorganization." *Nature Communications* 10(1): 1–20. https://doi.org/10.1038/s41467-019-13604-5 (October 4, 2020).

Iwasaki, Osamu, and Ken ichi Noma. 2016. "Condensin-Mediated Chromosome Organization in Fission Yeast." *Current Genetics* 62(4): 739–43. /pmc/articles/PMC5184768/?report=abstract (October 2, 2020).

Jans, David A., and Stefan Hübner. 1996. "Regulation of Protein Transport to the Nucleus: Central Role of Phosphorylation." *Physiological Reviews* 76(3): 651–85. https://journals.physiology.org/doi/abs/10.1152/physrev.1996.76.3.651 (September 23, 2020).

Jans, David A. 1995. 311 Biochem. J *The Regulation of Protein Transport to the Nucleus by Phosphorylation*.

Jiang, Yan, Mingyi Liu, Charlotte A. Spencer, and David H. Price. 2004. "Involvement of Transcription Termination Factor 2 in Mitotic Repression of Transcription Elongation." *Molecular Cell* 14(3): 375–86.

Jiang, Yan, and David H Price. 2004. "Rescue of the TTF2 Knockdown Phenotype with an SiRNA-Resistant Replacement Vector." *Cell cycle (Georgetown, Tex.)* 3(9): 1151–53. http://www.ncbi.nlm.nih.gov/pubmed/15467445 (March 2, 2020).

Jishage, Miki et al. 2018. "Architecture of Pol II(G) and Molecular Mechanism of Transcription Regulation by Gdown1." *Nature Structural and Molecular Biology* 25(9): 859–67. https://pubmed.ncbi.nlm.nih.gov/30190596/ (November 26, 2020).

Jones, Katherine A. 1997. "Taking a New TAK on TAT Transactivation." *Genes and Development* 11(20): 2593–99.

Kachaner, David et al. 2017. "Coupling of Polo Kinase Activation to Nuclear Localization by a Bifunctional NLS Is Required during Mitotic Entry." *Nature Communications* 8(1): 1–16. www.nature.com/naturecommunications (October 8, 2020).

Kadauke, Stephan et al. 2012. "Tissue-Specific Mitotic Bookmarking by Hematopoietic Transcription

Factor GATA1." *Cell* 150(4): 725–37. /pmc/articles/PMC3425057/?report=abstract (July 19, 2020).

Kadauke, Stephan, and Gerd A. Blobel. 2013. "Mitotic Bookmarking by Transcription Factors." *Epigenetics and Chromatin* 6(1): 6. /pmc/articles/PMC3621617/?report=abstract (August 30, 2020).

Kaffman, Arie, and Erin K. O'Shea. 1999. "Regulation of Nuclear Localization: A Key to a Door." *Annual Review of Cell and Developmental Biology* 15(1): 291–339.

Kang, Hyeseon et al. 2020. "Dynamic Regulation of Histone Modifications and Long-Range Chromosomal Interactions during Postmitotic Transcriptional Reactivation." http://www.genesdev.org/cgi/doi/10.1101/gad.335794.119. (October 2, 2020).

Kao, Shaw Yi, Andrew F. Calman, Paul A. Luciw, and B. Matija Peterlin. 1987. "Anti-Termination of Transcription within the Long Terminal Repeat of HIV-1 by Tat Gene Product." *Nature* 330(6147): 489–93.

Karn, Jonathan, and C. Martin Stoltzfus. 2012. "Transcriptional and Posttranscriptional Regulation of HIV-1 Gene Expression." *Cold Spring Harbor Perspectives in Medicine* 2(2): a006916. http://perspectivesinmedicine.cshlp.org/ (August 10, 2020).

Kim, Minkyu et al. 2004. "The Yeast Rat1 Exonuclease Promotes Transcription Termination by RNA Polymerase II." *Nature* 432(7016): 517–22. www.nature.com/nature (October 13, 2020).

Kobayashi, Wataru, and Hitoshi Kurumizaka. 2019. "Structural Transition of the Nucleosome during Chromatin Remodeling and Transcription." *Current Opinion in Structural Biology* 59: 107–14.

Krivega, Ivan, and Ann Dean. 2012. "Enhancer and Promoter Interactions-Long Distance Calls." *Current Opinion in Genetics and Development* 22(2): 79–85. https://pubmed.ncbi.nlm.nih.gov/22169023/ (June 24, 2020).

Kuehner, Jason N., Erika L. Pearson, and Claire Moore. 2011. "Unravelling the Means to an End: RNA Polymerase II Transcription Termination." *Nature Reviews Molecular Cell Biology* 12(5): 283–94.

Kwak, Hojoong, and John T. Lis. 2013. "Control of Transcriptional Elongation." *Annual Review of Genetics* 47(1): 483–508.

Lambert, Samuel A. et al. 2018. "The Human Transcription Factors." *Cell*.

"Lds Lodestar [Drosophila Melanogaster (Fruit Fly)] - Gene - NCBI." 2020. https://www.ncbi.nlm.nih.gov/gene/?term=Drosophila+melanogaster+lodestar (November 19, 2020).

Lee, Tong Ihn, and Richard A. Young. 2013. "Transcriptional Regulation and Its Misregulation in Disease." *Cell* 152(6): 1237–51. http://www.cell.com/article/S0092867413002031/fulltext (June 24, 2020).

Lei, X. H., X. Shen, X. Q. Xu, and H. S. Bernstein. 2000. "Human Cdc5, a Regulator of Mitotic Entry, Can Act as a Site-Specific DNA Binding Protein." *Journal of Cell Science* 113(24): 4523–31.

Lemay, Jean François, and François Bachand. 2015. "Fail-Safe Transcription Termination: Because One Is Never Enough." *RNA Biology* 12(9): 927–32. https://www.tandfonline.com/doi/abs/10.1080/15476286.2015.1073433 (July 26, 2020).

Lemieux, Bruno et al. 2015. “A Function for the HnRNP A1/A2 Proteins in Transcription Elongation.” *PLoS ONE* 10(5).

Leonard, Deana, Paul Ajuh, Angus I. Lamond, and Randy J. Legerski. 2003. “HLodestar/HuF2 Interacts with CDC5L and Is Involved in Pre-MRNA Splicing.” *Biochemical and Biophysical Research Communications* 308(4): 793–801.

Li, Tiandao, and David H. Price. 2012. “Gdown1: Making a Link between Mediator and RNA Polymerase II Elongation Control.” *Transcription* 3(4): 177. /pmc/articles/PMC3654766/?report=abstract (October 26, 2020).

Li, Wenbo et al. 2015. “Condensin I and II Complexes License Full Estrogen Receptor α-Dependent Enhancer Activation.” *Molecular Cell* 59(2): 188–202.

Liang, Kaiwei et al. 2015. “Mitotic Transcriptional Activation: Clearance of Actively Engaged Pol II via Transcriptional Elongation Control in Mitosis.” *Molecular Cell* 60(3): 435–45.

Liu, Mingyi, Zhi Xie, and David H. Price. 1998. “A Human RNA Polymerase II Transcription Termination Factor Is a SWI2/SNF2 Family Member.” *Journal of Biological Chemistry* 273(40): 25541–44.

Liu, Rong Diao, Jun Wu, Rui Shao, and Yu Hua Xue. 2014. “Mechanism and Factors That Control HIV-1 Transcription and Latency Activation.” *Journal of Zhejiang University: Science B* 15(5): 455–65.

Long, John J., Anne Leresche, Richard W. Kriwacki, and Joel M. Gottesfeld. 1998. “Repression of TFIIH Transcriptional Activity and TFIIH-Associated Cdk7 Kinase Activity at Mitosis.” *Molecular and Cellular Biology* 18(3): 1467–76. https://pubmed.ncbi.nlm.nih.gov/9488463/ (October 20, 2020).

Maniatis, Tom, and Robin Reed. 2002. “An Extensive Network of Coupling among Gene Expression Machines.” *Nature* 416(6880): 499–506.

Maquat, Lynne E. 2004. “Nonsense-Mediated MRNA Decay: Splicing, Translation and MRNP Dynamics.” *Nature Reviews Molecular Cell Biology* 5(2): 89–99.

Marshall, N F, and D H Price. 1992. “Control of Formation of Two Distinct Classes of RNA Polymerase II Elongation Complexes.” *Molecular and Cellular Biology* 12(5): 2078–90. /pmc/articles/PMC364379/?report=abstract (October 13, 2020).

Marshall, Nick F., Junmin Peng, Zhi Xie, and David H. Price. 1996. “Control of RNA Polymerase II Elongation Potential by a Novel Carboxyl- Terminal Domain Kinase.” *Journal of Biological Chemistry* 271(43): 27176–83.

Martin, Jeffrey W. et al. 2014. “Digital Expression Profiling Identifies RUNX2, CDC5L, MDM2, RECQL4, and CDK4 as Potential Predictive Biomarkers for Neo-Adjuvant Chemotherapy Response in Paediatric Osteosarcoma.” *PLoS ONE* 9(5).

McCracken, Susan et al. 1997. “The C-Terminal Domain of RNA Polymerase II Couples MRNA Processing to Transcription.” *Nature* 385(6614): 357–60. https://pubmed.ncbi.nlm.nih.gov/9002523/ (August 29, 2020).

Mercadante, Anthony A., and Shamim S. Mohiuddin. 2020. StatPearls *Biochemistry, Replication and Transcription*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/pubmed/30986011 (October

12, 2020).

Mirkovic, Mihailo, and Raquel A. Oliveira. 2017. “Centromeric Cohesin: Molecular Glue and Much More.” *Progress in molecular and subcellular biology* 56: 485–513. http://www.ncbi.nlm.nih.gov/pubmed/28840250 (March 17, 2020).

Mu, R. et al. 2014. “Depletion of Pre-MRNA Splicing Factor Cdc5L Inhibits Mitotic Progression and Triggers Mitotic Catastrophe.” *Cell Death and Disease* 5(3): e1151. /pmc/articles/PMC3973201/?report=abstract (October 27, 2020).

Muse, Ginger W. et al. 2007. “RNA Polymerase Is Poised for Activation across the Genome.” *Nature Genetics* 39(12): 1507–11.

Nakazawa, Norihiko, Orie Arakawa, and Mitsuhiro Yanagida. 2019. “Condensin Locates at Transcriptional Termination Sites in Mitosis, Possibly Releasing Mitotic Transcripts.” *Open Biology* 9(10): 190125. https://royalsocietypublishing.org/doi/10.1098/rsob.190125 (October 15, 2020).

“NCBI Blast:Sp|Q9UNY4|583-786.” 2020. https://blast.ncbi.nlm.nih.gov/Blast.cgi (August 24, 2020).

“NCBI Blast:Sp|Q9UNY4|995-1157.” 2020. https://blast.ncbi.nlm.nih.gov/Blast.cgi (August 24, 2020).

“NCBI Blast:Sp|Q9UNY4|TTF2_HUMAN Transcription Termination...” 2020. https://blast.ncbi.nlm.nih.gov/Blast.cgi (August 24, 2020).

Nikolov, D. B., and S. K. Burley. 1997. “RNA Polymerase II Transcription Initiation: A Structural View.” *Proceedings of the National Academy of Sciences of the United States of America* 94(1): 15–22.

Nowak, Scott J., and Victor G. Corces. 2004. “Phosphorylation of Histone H3: A Balancing Act between Chromosome Condensation and Transcriptional Activation.” *Trends in Genetics* 20(4): 214–20.

O’Farrell, Patrick H. 1992. “Big Genes and Little Genes and Deadlines for Transcription.” *Nature* 359(6394): 366–67. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2754300/ (November 26, 2020).

Oliveira, Raquel A. et al. 2010. “Cohesin Cleavage and Cdk Inhibition Trigger Formation of Daughter Nuclei.” *Nature Cell Biology* 12(2): 185–92.

Orr, Bernardo, Kristina M. Godek, and Duane Compton. 2015. “Aneuploidy.” *Current Biology* 25(13): R538–42. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4714037/ (August 20, 2020).

Ott, Melanie, Matthias Geyer, and Qiang Zhou. 2011. “The Control of HIV Transcription: Keeping RNA Polymerase II on Track.” *Cell Host and Microbe* 10(5): 426–35. /pmc/articles/PMC3478145/?report=abstract (August 10, 2020).

Palozola, Katherine C., Greg Donahue, et al. 2017. “Mitotic Transcription and Waves of Gene Reactivation during Mitotic Exit.” *Science* 358(6359): 119–22. http://science.sciencemag.org/ (July 25, 2020).

Palozola, Katherine C., Hong Liu, Dario Nicetto, and Kenneth S. Zaret. 2017. “Low-Level, Global Transcription during Mitosis and Dynamic Gene Reactivation during Mitotic Exit.” *Cold Spring Harbor symposia on quantitative biology* 82: 197–205. http://symposium.cshlp.org/content/82/197.full (October 20, 2020).

Panne, Daniel. 2008. “The Enhanceosome.” *Current Opinion in Structural Biology* 18(2): 236–42.

https://pubmed.ncbi.nlm.nih.gov/18206362/ (June 24, 2020).

Park, Jieun, Myungjin Kang, and Minkyu Kim. 2015. “Unraveling the Mechanistic Features of RNA Polymerase II Termination by the 5’-3’ Exoribonuclease Rat1.” *Nucleic Acids Research* 43(5): 2625–37. /pmc/articles/PMC4357727/?report=abstract (October 13, 2020).

Paul, Matthew Robert, Andreas Hochwagen, and Sevinç Ercan. 2019. “Condensin Action and Compaction.” *Current Genetics* 65(2): 407–15. https://doi.org/10.1007/s00294-018-0899-4 (October 2, 2020).

Peng, J. et al. 1998. “RNA Polymerase II Elongation Control.” In *Cold Spring Harbor Symposia on Quantitative Biology*, Cold Spring Harbor Laboratory Press, 365–70.

Pennacchio, Len A. et al. 2013. “Enhancers: Five Essential Questions.” *Nature Reviews Genetics* 14(4): 288–95. /pmc/articles/PMC4445073/?report=abstract (October 12, 2020).

Perea-Resa, Carlos, Leah Bury, Iain M. Cheeseman, and Michael D. Blower. 2020. “Cohesin Removal Reprograms Gene Expression upon Mitotic Entry.” *Molecular Cell* 78(1): 127-140.e7.

“Pfam: Protein: TTF2_DROME (P34739).” 2020. https://pfam.xfam.org/protein/P34739 (August 23, 2020).

“Pfam: Protein: TTF2_HUMAN (Q9UNY4).” 2020. https://pfam.xfam.org/protein/Q9UNY4 (August 23, 2020).

Piskadlo, Ewa, and Raquel A. Oliveira. 2017. “A Topology-Centric View on Mitotic Chromosome Architecture.” *International Journal of Molecular Sciences* 18(12). https://pubmed.ncbi.nlm.nih.gov/29258269/ (July 15, 2020).

Poon, Ivan K.H., and David A. Jans. 2005. “Regulation of Nuclear Transport: Central Role in Development and Transformation?” *Traffic* 6(3): 173–86. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1600-0854.2005.00268.x (November 1, 2020).

Porrua, Odil, and Domenico Libri. 2015. “Transcription Termination and the Control of the Transcriptome: Why, Where and How to Stop.” *Nature Reviews Molecular Cell Biology* 16(3): 190–202.

Prasad, Punit, Andreas Lennartsson, and Karl Ekwall. 2015. “The Roles of SNF2/SWI2 Nucleosome Remodeling Enzymes in Blood Cell Differentiation and Leukemia.” *BioMed Research International* 2015. https://pubmed.ncbi.nlm.nih.gov/25789315/ (September 8, 2020).

Prasanth, Kannanganattu V., Paula A. Sacco-Bubulya, Supriya G. Prasanth, and David L. Spector. 2003. “Sequential Entry of Components of Gene Expression Machinery into Daughter Nucleiv.” *Molecular Biology of the Cell* 14(3): 1043–57. /pmc/articles/PMC151578/?report=abstract (August 29, 2020).

Price, D. H., A. E. Sluder, and A. L. Greenleaf. 1987. “Fractionation of Transcription Factors for RNA Polymerase II from Drosophila Kc Cell Nuclear Extracts.” *Journal of Biological Chemistry* 262(7): 3244–55. https://europepmc.org/article/med/3818640 (October 28, 2020).

Price, David H. 2000. “P-TEFb, a Cyclin-Dependent Kinase Controlling Elongation by RNA Polymerase II.” *Molecular and Cellular Biology* 20(8): 2629–34.

———. 2008. “Poised Polymerases: On Your Mark...Get Set...Go!” *Molecular Cell* 30(1): 7–10.

Proudfoot, Nick J., Andre Furger, and Michael J. Dye. 2002. “Integrating MRNA Processing with

Transcription." *Cell* 108(4): 501–12.

"PSORT II Prediction." 2020. https://psort.hgc.jp/form2.html (November 22, 2020).

"PSORT II Server Lds." 2020. https://psort.hgc.jp/cgi-bin/runpsort.pl (November 22, 2020).

"PSORT II Server TTF2." 2020. https://psort.hgc.jp/cgi-bin/runpsort.pl (November 22, 2020).

Raccaud, Mahé, and David M. Suter. 2018. "Transcription Factor Retention on Mitotic Chromosomes: Regulatory Mechanisms and Impact on Cell Fate Decisions." *FEBS Letters* 592(6): 878–87. http://doi.wiley.com/10.1002/1873-3468.12828 (October 20, 2020).

Racki, Lisa R., and Geeta J. Narlikar. 2008. "ATP-Dependent Chromatin Remodeling Enzymes: Two Heads Are Not Better, Just Different." *Current Opinion in Genetics and Development* 18(2): 137–44.

Raj, Saurabh et al. 2019. "Mechanistic Characterization of the DEAD-Box RNA Helicase Ded1 from Yeast as Revealed by a Novel Technique Using Single-Molecule Magnetic Tweezers." *Nucleic Acids Research* 47(7): 3699–3710. https://academic.oup.com/nar/article/47/7/3699/5319129 (August 23, 2020).

Ramakrishnan, Rajesh et al. 2012. "Identification of Novel CDK9 and Cyclin T1-Associated Protein Complexes (CCAPs) Whose SiRNA Depletion Enhances HIV-1 Tat Function." *Retrovirology* 9.

Rambout, Xavier, Franck Dequiedt, and Lynne E. Maquat. 2018. "Beyond Transcription: Roles of Transcription Factors in Pre-MRNA Splicing." *Chemical Reviews* 118(8): 4339–64.

Rao, Suhas S.P. et al. 2017. "Cohesin Loss Eliminates All Loop Domains." *Cell* 171(2): 305-320.e24.

Reed, Robin, and Ed Hurt. 2002. "A Conserved MRNA Export Machinery Coupled to Pre-MRNA Splicing." *Cell* 108(4): 523–31. https://pubmed.ncbi.nlm.nih.gov/11909523/ (August 29, 2020).

Richard, Patricia, and James L. Manley. 2009. "Transcription Termination by Nuclear RNA Polymerases." *Genes and Development* 23(11): 1247–69. https://pubmed.ncbi.nlm.nih.gov/19487567/ (October 13, 2020).

Robellet, Xavier, Vincent Vanoosthuyse, and Pascal Bernard. 2017. "The Loading of Condensin in the Context of Chromatin." *Current Genetics* 63(4): 577–89. https://link.springer.com/article/10.1007/s00294-016-0669-0 (October 2, 2020).

Rondón, Ana G., Hannah E. Mischo, Junya Kawauchi, and Nick J. Proudfoot. 2009. "Fail-Safe Transcriptional Termination for Protein-Coding Genes in S. Cerevisiae." *Molecular Cell* 36(1): 88–98.

Ryan, Daniel P., and Tom Owen-Hughes. 2011. "Snf2-Family Proteins: Chromatin Remodellers for Any Occasion." *Current Opinion in Chemical Biology* 15(5): 649–56.

Sadikovic, Bekim et al. 2010. "Expression Analysis of Genes Associated with Human Osteosarcoma Tumors Shows Correlation of RUNX2 Overexpression with Poor Response to Chemotherapy." *BMC Cancer* 10.

Sansó, Miriam et al. 2016. "P-TEFb Regulation of Transcription Termination Factor Xrn2 Revealed by a Chemical Genetic Screen for Cdk9 Substrates." *Genes and Development* 30(1): 117–31. /pmc/articles/PMC4701974/?report=abstract (September 22, 2020).

Sarge, Kevin D., and Ok Kyong Park-Sarge. 2005. "Gene Bookmarking: Keeping the Pages Open." *Trends in Biochemical Sciences* 30(11): 605–10.

http://www.cell.com/article/S0968000405002707/fulltext (July 19, 2020).

Saunders, Abbie, Leighton J. Core, and John T. Lis. 2006. “Breaking Barriers to Transcription Elongation.” *Nature Reviews Molecular Cell Biology* 7(8): 557–67. http://www.nature.com/articles/nrm1981 (March 17, 2020).

Schüller, Roland et al. 2016. “Heptad-Specific Phosphorylation of RNA Polymerase II CTD.” *Molecular Cell* 61(2): 305–14.

Seller, Charles A., and Patrick H. O’Farrell. 2018. “Rif1 Prolongs the Embryonic S Phase at the Drosophila Mid-Blastula Transition.” *PLoS Biology* 16(5). /pmc/articles/PMC5963817/?report=abstract (October 24, 2020).

Shandilya, Jayasha, and Stefan G.E. Roberts. 2012. “The Transcription Cycle in Eukaryotes: From Productive Initiation to RNA Polymerase II Recycling.” *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1819(5): 391–400.

Sharp, Judith A., Carlos Perea-Resa, Wei Wang, and Michael D. Blower. 2020. “Cell Division Requires RNA Eviction from Condensing Chromosomes.” *The Journal of cell biology* 219(11). https://doi.org/10.1083/jcb.201910148 (November 16, 2020).

Shermoen, Antony W., Mark L. McCleland, and Patrick H. O’Farrell. 2010. “Developmental Control of Late Replication and s Phase Length.” *Current Biology* 20(23): 2067–77.

Shermoen, Antony W., and Patrick H. O’Farrell. 1991. “Progression of the Cell Cycle through Mitosis Leads to Abortion of Nascent Transcripts.” *Cell* 67(2): 303–10. /pmc/articles/PMC2755073/?report=abstract (November 26, 2020).

Shoaib, Muhammad, Nidhi Nair, and Claus Storgaard Sørensen. 2020. “Chromatin Landscaping At Mitotic Exit Orchestrates Genome Function.” *Frontiers in Genetics* 11. /pmc/articles/PMC7052122/?report=abstract (October 2, 2020).

Sif, Saïd, P. Todd Stukenberg, Marc W. Kirschner, and Robert E. Kingston. 1998. “Mitotic Inactivation of a Human SWI/SNF Chromatin Remodeling Complex.” *Genes and Development* 12(18): 2842–51. /pmc/articles/PMC317164/?report=abstract (October 5, 2020).

Silva, Rui D. et al. 2018. “Absence of the Spindle Assembly Checkpoint Restores Mitotic Fidelity upon Loss of Sister Chromatid Cohesion.” *Current Biology* 28(17): 2837-2844.e3.

Singleton, Martin R., Mark S. Dillingham, and Dale B. Wigley. 2007. “Structure and Mechanism of Helicases and Nucleic Acid Translocases.” *Annual Review of Biochemistry* 76(1): 23–50.

Stradleigh, Tyler W., and Andrew T. Ishida. 2015. “Fixation Strategies for Retinal Immunohistochemistry.” *Progress in Retinal and Eye Research* 48: 181–202. https://linkinghub.elsevier.com/retrieve/pii/S135094621500018X (November 16, 2020).

Szabad, J et al. 1989. “Isolation and Characterization of Dominant Female Sterile Mutations of Drosophila Melanogaster. II. Mutations on the Second Chromosome.” *Genetics* 122(4): 823–35. http://www.ncbi.nlm.nih.gov/pubmed/2503422 (May 22, 2020).

Szabad, J, E Máthé, and J Puro. 1995. “Horka, a Dominant Mutation of Drosophila, Induces Nondisjunction and, through Paternal Effect, Chromosome Loss and Genetic Mosaics.” *Genetics* 139(4): 1585–99. http://www.ncbi.nlm.nih.gov/pubmed/7789762 (May 14, 2020).

Szalontai, Tamas et al. 2009. “HorkaD, a Chromosome Instability-Causing Mutation in Drosophila, Is

a Dominant-Negative Allele of Lodestar." *Genetics* 181(2): 367–77.

Tahirov, Tahir H. et al. 2010. "Crystal Structure of HIV-1 Tat Complexed with Human P-TEFb." *Nature* 465(7299): 747–51. https://pubmed.ncbi.nlm.nih.gov/20535204/ (June 28, 2020).

Tanaka, Tomoyuki, Jörg Fuchs, Josef Loidl, and Kim Nasmyth. 2000. "Cohesin Ensures Bipolar Attachment of Microtubules to Sister Centromeres and Resists Their Precocious Separation." *Nature Cell Biology* 2(8): 492–99.

Tanner, N. Kyle et al. 2003. "The Q Motif: A Newly Identified Motif in DEAD Box Helicases May Regulate ATP Binding and Hydrolysis." *Molecular Cell* 11(1): 127–38. http://www.cell.com/article/S1097276503000066/fulltext (July 20, 2020).

Thomä, Nicolas H. et al. 2005. "Structure of the SWI2/SNF2 Chromatin-Remodeling Domain of Eukaryotic Rad54." *Nature Structural and Molecular Biology* 12(4): 350–56.

Timmers, H. T.Marc, and C. Peter Verrijzer. 2017. "Mitotic Chromosomes: Not So Silent After All." *Developmental Cell* 43(2): 119–21. https://pubmed.ncbi.nlm.nih.gov/29065303/ (October 20, 2020).

Tippens, Nathaniel D., Anniina Vihervaara, and John T. Lis. 2018. "Enhancer Transcription: What, Where, When, and Why?" *Genes and Development* 32(1): 1–3. /pmc/articles/PMC5828389/?report=abstract (September 10, 2020).

Tong Ihn Lee, and R. A. Young. 2000. "Transcription of Eukaryotic Protein-Coding Genes." *Annual Review of Genetics* 34: 77–137. https://pubmed.ncbi.nlm.nih.gov/11092823/ (October 12, 2020).

"TTF2 - Transcription Termination Factor 2 - Homo Sapiens (Human) - TTF2 Gene & Protein." 2020. https://www.uniprot.org/uniprot/Q9UNY4#sequences (August 23, 2020).

"TTF2 Transcription Termination Factor 2 [Homo Sapiens (Human)] - Gene - NCBI." 2020. https://www.ncbi.nlm.nih.gov/gene/8458 (August 25, 2020).

Tuteja, Narendra, and Renu Tuteja. 2004. "Unraveling DNA Helicases. Motif, Structure, Mechanism and Function." *European Journal of Biochemistry* 271(10): 1849–63. http://doi.wiley.com/10.1111/j.1432-1033.2004.04094.x (July 20, 2020).

Umarov, Ramzan Kh, and Victor V. Solovyev. 2017. "Recognition of Prokaryotic and Eukaryotic Promoters Using Convolutional Deep Learning Neural Networks." *PLoS ONE* 12(2). /pmc/articles/PMC5291440/?report=abstract (October 12, 2020).

Wada, Tadashi et al. 1998. *DSIF, a Novel Transcription Elongation Factor That Regulates RNA Polymerase II Processivity, Is Composed of Human Spt4 and Spt5 Homologs We Report the Identification of a Transcription Elongation Factor from HeLa Cell Nuclear Extracts That Causes Pausing of RNA Polymerase II (Pol II) in Conjunction with The*. www.genesdev.org (August 3, 2020).

Wagschal, Alexandre et al. 2012. "Microprocessor, Setx, Xrn2, and Rrp6 Co-Operate to Induce Premature Termination of Transcription by RNAPII." *Cell* 150(6): 1147–57. /pmc/articles/PMC3595997/?report=abstract (November 27, 2020).

Wan, Li, and Jun Huang. 2014. "The PSO4 Protein Complex Associates with Replication Protein a (RPA) and Modulates the Activation of Ataxia Telangiectasia-Mutated and RAD3-Related (ATR)." *Journal of Biological Chemistry* 289(10): 6619–26.

Wei, Yi et al. 1999. “Phosphorylation of Histone H3 Is Required for Proper Chromosome Condensation and Segregation.” *Cell* 97(1): 99–109.

West, Steven, Natalla Gromak, and Nick J. Proudfoot. 2004. “Human 5′ → 3′ Exonuclease Xm2 Promotes Transcription Termination at Co-Transcriptional Cleavage Sites.” *Nature* 432(7016): 522–25. www.nature.com/nature. (October 13, 2020).

Will, Cindy L., and Reinhard Lührmann. 2011. “Spliceosome Structure and Function.” *Cold Spring Harbor Perspectives in Biology* 3(7): 1–2. /pmc/articles/PMC3119917/?report=abstract (August 2, 2020).

Woychik, Nancy A., and Michael Hampsey. 2002. “The RNA Polymerase II Machinery: Structure Illuminates Function.” *Cell* 108(4): 453–63. http://www.cell.com/article/S0092867402006463/fulltext (July 26, 2020).

Xie, Zhi, and David Price. 1997. “Drosophila Factor 2, an RNA Polymerase II Transcript Release Factor, Has DNA-Dependent ATPase Activity.” *Journal of Biological Chemistry* 272(50): 31902–7.

Xie, Zhi, and David H. Price. 1996. “Purification of an RNA Polymerase II Transcript Release Factor from Drosophila.” *Journal of Biological Chemistry* 271(19): 11043–46. https://pubmed.ncbi.nlm.nih.gov/8626643/ (July 16, 2020).

———. 1998. “Unusual Nucleic Acid Binding Properties of Factor 2, an RNA Polymerase II Transcript Release Factor.” *Journal of Biological Chemistry* 273(6): 3771–77.

Zeitlinger, Julia et al. 2007. “RNA Polymerase Stalling at Developmental Control Genes in the Drosophila Melanogaster Embryo.” *Nature Genetics* 39(12): 1512–16.

Zhang, Dongdong et al. 2019. “Identification and Expression Analysis of Snf2 Family Proteins in Tomato (Solanum Lycopersicum).” *International Journal of Genomics* 2019.

Zhou, Qiang, Tiandao Li, and David H. Price. 2012. “RNA Polymerase II Elongation Control.” *Annual Review of Biochemistry* 81(1): 119–43.

7 Appendices

Score	Expect	Method	Identities	Positives	Gaps
489 bits(1260)	1e-158	Compositional matrix adjust.	303/778(39%)	436/778(56%)	87/778(11%)

```
Query  436   LPDKGQKLIKQIQELEEVLSGLTLSPEQGTNEKSNSQV-PQQSHFTKTTTGPPHLVPPQP  494
             LPDKG +++K+I          TL  E   +E+  S +  QQS+        P L PP+
Sbjct  320   LPDKGSQIMKRID---------TLRRELAMDEQWISALRVQQSNVPAVRVVKPTLNPPR-  369

Query  495   LPRRGTQPVGSLELKSACQVTAGGSSQCYRGHTNQDHVHAVWKITSEAIGQLHRSLESCP  554
                     + +L+     +          Y G       +    +T E++  LH SLE  P
Sbjct  370   -----APSIDTLDWDELSEA-VNEIKPVYTGAQGMATFNNQKALTLESLKDLHVSLEDLP  423

Query  555   GETVVAEDPAGLKVPLLLHQKQALAWLLWRESQKPQGGILADDMGLGKTLTMIALILTQK  614
             G  V+AEDP GLKV L+ HQK ALAW+ WRE + P+GGILADDMGLGKTLTMI+ +L  K
Sbjct  424   GPEVLAEDPVGLKVSLMNHQKHALAWMSWRERKLPRGGILADDMGLGKTLTMISSVLACK  483

Query  615   NQEKKEE---------------KEKSTALTWLSKDDSCDFTSHGTLIICPASLIHHWKNE  659
             N ++  E               K++ +   W SK    D    GTL++CPASL+  W++E
Sbjct  484   NGQEMSEGKDESSDSDSEDDKNKKRKSVTGWKSKGRK-DTRRGGTLVVCPASLLRQWESE  542

Query  660   VEKRVNSNKLRVYLYHGPNRDSRARVLSTYDIVITTYSLVAKEIPTNKQEAEIPGANLNV  719
             VE +V+  KL V ++HG NR+++ + L  YDIV+TTY +VA+E
Sbjct  543   VESKVSRQKLTVCVHHGNNRETKGKYLRDYDIVVTTYQIVARE----------------  585

Query  720   EGTSTPLLRIAWARIILDEAHNVKNPRVQTSIAVCKLQACARWAVTGTPIQNNLLDMYSL  779
               + + +  + W RIILDEAH V+N + Q+S+AVC L+   RWA+TGTPIQN  LD+Y+L
Sbjct  586   HKSLSAVFGVKWRRIILDEAHVVRNHKSQSSLAVCDLRGKYRWALTGTPIQNKELDVYAL  645

Query  780   LKFLRCSPFDEFNLWRSQVDNGSKKGGERLSILTKSLLLRRTKDQLDSTGRPLVILPQRK  839
             LKFLRCSPFD+ + W+  +DN S  G  RL++L KSL+LRRTK QL S G+ L  LP ++
Sbjct  646   LKFLRCSPFDDLHTWKKWIDNKSAGGQNRLNLLMKSLMLRRTKAQLQSDGK-LNSLPNKE  704

Query  840   FQLHHLKLSEDEETVYNVFFARSRSALQSYL-KRHESRGNQSGRSPNN--PFSRVALEFG  896
              +L  + L ++E  VY      SR+    +L +R E   + + RS  N   ++++    G
Sbjct  705   LRLIEISLDKEEMNVYQTVMTYSRTLFAQFLHQRAERETDFNYRSDANKPTYNQIKDPNG  764

Query  897   SEEPRH---SEAADSPRSSTVH-ILSQLLRLRQCCCHLSLLKSALDPMELKGEGL-----  947
             +    H   +  A S +    H IL  LLRLRQ CCH  L+ + LD  E +  G
Sbjct  765   AYYKMHEKFARMAGSKKEVKSHDILVLLLRLRQICCHPGLIDAMLDGEESQTMGDHSSDS  824

Query  948   ---VLSLEEQLSALTLSE-------------------LRDSEPSSTVSLNGTFFKMELFE  985
                + L  QL+ L +++                    L D    +  S N       +F
Sbjct  825   DTPEIDLLAQLNKLAITDTSTDGQQSVANAGDDGPPLLPDEARIAKASKNLLKRSNPVFN  884

Query  986   GMRESTKISSLLAELEAIQRNSASQKSVIVSQWTNMLKVVALHLKKHGLTYATIDGSVNP  1045
               R S+KI+ ++  L+     S+  K+++VSQWT++L ++  HL K G+   +++G++
Sbjct  885   LHRPSSKINMVIQILKTSILKSSDDKAIVVSQWTSVLDILRDHLSKDGVATLSLNGTIPV  944

Query  1046  KQRMDLVEAFN-HSRGPQVMLISLLAGGVGLNLTGGNHLFLLDMHWNPSLEDQACDRIYR  1104
             K R D+V  FN  +   +V+L+SL AGGVGLNL G NHL LLD+HWNP LE QA DRIYR
Sbjct  945   KNRQDIVNEFNDRNNQKRVLLLSLTAGGVGLNLIGANHLLLLDLHWNPQLEAQAQDRIYR  1004

Query  1105  VGQQKDVVIHRFVCEGTVEEKILQLQEKKKDLAKQVLSGSGESVTKLTLADLRVLFGI  1162
             VGQ+K+V+I++F+C  TVE++I  LQ+KK DLA  VL+G+  S +KLT+ DL+ LFG+
Sbjct  1005  VGQKKNVIIYKFMCVDTVEQRIKGLQDKKLDLADGVLTGAKVS-SKLTIDDLKGLFGM  1061
```

Figure 7.1 The alignment results of TTF2 and Lds.
Sequencing analysis shows that the two proteins are 39% identical and 56% similar. TTF2 (ID number is Q9UNY4) is on the top and Lds (ID number is P34739) is on the bottom. The alignment has been performed using the Blastp tool.

```
 Score          Expect  Method                        Identities        Positives         Gaps
 190 bits(483)  2e-66   Compositional matrix adjust.  101/219(46%)     138/219(63%)      33/219(15%)

Query  1    WRESQKPQGGILADDMGLGKTLTMIALILTQKNQEKKEE---------------KEKSTA  45
            WRE + P+GGILADDMGLGKTLTMI+ +L  KN ++  E                K++ +
Sbjct  1    WRERKLPRGGILADDMGLGKTLTMISSVLACKNGQEMSEGKDESSDSDSEDDKNKKRKSV  60

Query  46   LTWLSKDDSCDFTSHGTLIICPASLIHHWKNEVEKRVNSNKLRVYLYHGPNRDSRARVLS  105
              W SK    D    GTL++CPASL+  W++EVE +V+  KL V ++HG NR+++ + L
Sbjct  61   TGWKSKGRK-DTRRGGTLVVCPASLLRQWESEVESKVSRQKLTVCVHHGNNRETKGKYLR  119

Query  106  TYDIVITTYSLVAKEIPTNKQEAEIPGANLNVEGTSTPLLRIAWARIILDEAHNVKNPRV  165
             YDIV+TTY +VA+E    +K  + + G               + W RIILDEAH V+N +
Sbjct  120  DYDIVVTTYQIVARE---HKSLSAVFG--------------VKWRRIILDEAHVVRNHKS  162

Query  166  QTSIAVCKLQACARWAVTGTPIQNNLLDMYSLLKFLRCS  204
            Q+S+AVC L+   RWA+TGTPIQN  LD+Y+LLKFLRCS
Sbjct  163  QSSLAVCDLRGKYRWALTGTPIQNKELDVYALLKFLRCS  201
```

Figure 7.2 An alignment of the helicase ATP binding domain.
TTF2 and Lds conserved ATP binding domain were aligned and showed a 46% identity and 63% similarity.

```
 Score          Expect  Method                        Identities        Positives         Gaps
 160 bits(406)  5e-56   Compositional matrix adjust.  79/152(52%)      113/152(74%)      2/152(1%)

Query  13   SASQKSVIVSQWTNMLKVVALHLKKHGLTYATIDGSVNPKQRMDLVEAFN-HSRGPQVML  71
            S+  K+++VSQWT++L ++  HL K G+   +++G++  K R D+V  FN  +   +V+L
Sbjct  16   SSDDKAIVVSQWTSVLDILRDHLSKDGVATLSLNGTIPVKNRQDIVNEFNDRNNQKRVLL  75

Query  72   ISLLAGGVGLNLTGGNHLFLLDMHWNPSLEDQACDRIYRVGQQKDVVIHRFVCEGTVEEK  131
            +SL AGGVGLNL G NHL LLD+HWNP LE QA DRIYRVGQ+K+V+I++F+C  TVE++
Sbjct  76   LSLTAGGVGLNLIGANHLLLLDLHWNPQLEAQAQDRIYRVGQKKNVIIYKFMCVDTVEQR  135

Query  132  ILQLQEKKKDLAKQVLSGSGESVTKLTLADLR  163
            I  LQ+KK DLA  VL+G+  S +KLT+ DL+
Sbjct  136  IKGLQDKKLDLADGVLTGAKVS-SKLTIDDLK  166
```

Figure 7.3 An alignment of the Helicase C-terminal domains.
TTF2 and Lds conserved C-terminal domain were aligned. Their domain sequences show a 52% identity and 74% similarity.