THESIS / THÈSE

MASTER IN BIOCHEMISTRY AND MOLECULAR AND CELL BIOLOGY

Role of the kinase CDK-12 in the development of the nematode C. elegans

PONSARD, Pauline

Award date:
2020

Awarding institution:
University of Namur

Faculté des Sciences

ROLE OF THE KINASE CDK-12 IN THE DEVELOPMENT OF THE NEMATODE *C. ELEGANS*

Mémoire présenté pour l'obtention

du grade académique de master 120 en biochimie et biologie moléculaire et cellulaire

Pauline PONSARD

Janvier 2020

Université de Namur
FACULTE DES SCIENCES
Secrétariat du Département de Biologie
Rue de Bruxelles 61 - 5000 NAMUR
Téléphone: + 32(0)81.72.44.18 - Téléfax: + 32(0)81.72.44.20
E-mail: joelle.jonet@unamur.be - http://www.unamur.be

Fonction de la kinase CDK-12 dans le développement du nématode *C. elegans*

PONSARD Pauline

Résumé

La transcription des gènes codant pour des protéines est réalisée par l'ARN polymérase II (Pol II). Cette enzyme est composée de 12 sous-unités nommées Rpb1 à Rpb12 et possède sur sa plus grande sous-unité, Rpb1, un domaine carboxy-terminal (CTD) constitué de répétitions de l'heptapeptide Y_1-S_2-P_3-T_4-S_5-P_6-S_7. Chacun de ces résidus peut être modifié afin de créer différentes combinaisons de séquences permettant le recrutement de facteurs impliqués dans la maturation des ARNm. Les modifications les plus abondantes sont la phosphorylation de la sérine en position 2 (CTD-Ser2) et 5 (CTD-Ser5) par les kinases dépendantes des cyclines, CDK7 (qui phosphoryle la CTD-Ser5) et CDK12/CDK9 (qui phosphorylent la CTD-Ser2). Bien que CDK9 soit essentielle pour l'embryogenèse, CDK12 ne l'est pas et semble être importante uniquement pour la transcription d'un groupe de gènes impliqués dans le développement post-embryonnaire chez l'organisme modèle *Caenorhabditis elegans.* En effet, dans le laboratoire, de précédentes études ont démontré que l'inhibition de CDK-12 chez *C. elegans* entraîne un arrêt développemental au stade post-embryonnaire L1.

Le but de ce mémoire est de poursuivre l'étude de CDK-12, premièrement en vérifiant que le phénotype induit par son inactivation résulte bien d'un déficit de phosphorylation des sérines en position 2 du CTD et deuxièmement en étudiant la régulation de CDK-12 *in vivo.* Nous avions donc pour but de générer deux souches transgéniques, exprimant soit un CTD où les sérines en position 2 sont remplacées par un résidu non-phosphorylable, soit CDK-12 taggée afin d'étudier sa phosphorylation et ses interactants éventuels par spectrométrie de masse. Dans les deux cas, plusieurs variations de la méthode CRISPR/Cas9 ont été utilisées sans succès. Une souche transgénique contrôle ciblant un autre gène a été obtenue, indiquant que la méthode est efficace.

Mémoire de master 120 en biochimie et biologie moléculaire et cellulaire

Janvier 2020

Promoteur: D. Hermand

Université de Namur
FACULTE DES SCIENCES
Secrétariat du Département de Biologie
Rue de Bruxelles 61 - 5000 NAMUR
Téléphone: + 32(0)81.72.44.18 - Téléfax: + 32(0)81.72.44.20
E-mail: joelle.jonet@unamur.be - http://www.unamur.be

Role of the kinase CDK-12 in the development of the nematode *C. elegans*

PONSARD Pauline

Summary

The transcription of protein-coding genes is realized by the RNA polymerase II (Pol II). This enzyme is composed of 12 subunits named Rpb1 to Rpb12 and possesses on its largest subunit, Rpb1, a tail-like carboxy-terminal domain (CTD) composed of repetitions of the heptapeptide sequence Y_1-S_2-P_3-T_4-S_5-P_6-S_7. Each of the CTD residues can be post-translationally modified to form different combinations allowing factors involved in mRNA processing and maturation to dock on Pol II. The most abundant modifications are the phosphorylation of the serine in position 2 (CTD-Ser2) and 5 (CTD-Ser5) by the Cyclin-Dependent Kinases, CDK7 (which phosphorylates the CTD-Ser5) and CDK9/CDK12 (which phosphorylate the CTD-Ser2). Although CDK9 is essential for embryogenesis, CDK12 is not and is rather required for the expression of a subset of genes like genes involved in postembryonic development in the model organism *Caenorhabditis elegans.* Indeed, previous work from the lab showed, that the inactivation of CDK-12 in *C. elegans* results in an early post-embryonic developmental arrest at the L1 stage.

The goal of this master thesis is to pursue the study of CDK-12 by checking that the phenotype induced by its inactivation is indeed consequent to a deficit of CTD-Ser2 phosphorylation and second, by studying how CDK-12 is regulated *in vivo.* We thus aimed to create two *C. elegans* mutant strains, one expressing a CTD where the CTD-Ser2 are replaced by a non-phosphorylable residue and another expressing a tagged CDK-12 to study its phosphorylation status and putative partners by mass spectrometry. To generate these two strains, we used several variations of the CRISPR/Cas9 method, so far with no success. A transgenic control strain targeting another gene was obtained, which indicates that the method is efficient.

Mémoire de master 120 en biochimie et biologie moléculaire et cellulaire

Janvier 2020

Promoteur: D. Hermand

Remerciements

Je souhaiterais tout d'abord remercier mon promoteur, Damien Hermand, pour m'avoir permis de réaliser mon mémoire dans son laboratoire, pour son encadrement, ses précieux conseils ainsi que sa disponibilité.

Je voudrais ensuite remercier mon encadrante, Fanélie. Merci pour tout ce que tu m'as appris, pour le temps passé à mes côtés, pour la relecture attentive de mon mémoire mais aussi pour ton humour bien particulier auquel je l'avoue je n'étais pas toujours très réceptive. Je te souhaite plein de succès dans ta future carrière.

Merci ensuite à toi Valérie ou tatie Valoch, pour tes nombreuses astuces, pour tes cahiers de labo dont j'ai tant eu besoin et pour m'avoir remonté le moral dans les moments difficiles. Merci aussi pour ces fous rires, ces moments de debrief de nos émissions télé préférées ainsi que pour toutes les nouvelles expressions de la langue française que tu m'as apprises.

Merci à toi FX, pour ces discussions sur nos projets communs et les relectures de design. Bonne continuation dans ton projet de thèse (j'espère que tu te mettras à regarder DALS en mon absence pour pouvoir en discuter avec Val de peur qu'elle ne s'ennuie).

Je n'oublie pas tous les gens que j'ai rencontrés au cours de mes 5 années d'étude à l'UNAMUR, que ce soit en Biomed ou Biologie. Merci à vous tous, pour ces moments de réflexion, de travail mais aussi de détente et de rigolade.

Je remercie également ma famille, tout particulièrement ma maman, pour m'avoir toujours soutenue (pendant ce mémoire mais aussi lors des blocus), pour m'avoir poussée à réaliser mes rêves et pour avoir toujours cru en moi.

Enfin, je remercie les membres du jury, Catherine Lambert de Rouvroit, Benjamin Le Calvé, Emilien Nicolas et Julie Storder pour la lecture de ce mémoire.

Table of content

Introduction

1. Transcription

1.1. The central dogma of molecular biology

In 1958, Francis Crick argued that the hereditary information contained in DNA is first transformed into RNA and then into proteins. That is what we call nowadays "the central dogma of molecular biology" (Shapiro, 2009). The first step, the transcription of DNA into RNA, is performed into the nucleus by enzymes named RNA polymerases. The RNA is then exported into the cytoplasm where it can play its role or be translated into proteins by ribosomes (Clancy and Brown, 2008).

1.2. The different RNA polymerases

As mentioned, transcription is performed by RNA polymerases. While eubacteria and archaea possess only a single RNA polymerase, eukaryotes have at least three, the RNA polymerase I (Pol I), the RNA polymerase II (Pol II) and the RNA polymerase III (Pol III) (Cramer *et al*, 2008). Pol I is a complex of 14 subunits responsible for the transcription of large ribosomal RNA precursors. These RNAs will be processed into the 28S, 18S and 5.8S ribosomal RNAs (Cramer *et al*, 2008; Cooper, 2000). Pol II is a complex of 12 subunits (Rpb1 to Rbp12) that transcribes coding sequences into pre-mRNAs. It is also responsible for the transcription of most non-coding RNAs (ncRNAs). These ncRNAs have a lot of functions such as the regulation of the gene expression during healthy or pathological development, the translation regulation or even the pre-mRNA processing (Cramer *et al*, 2008*;* Cooper, 2000, Mattick and Makunin, 2006). Pol III is composed of 17 subunits. This RNA polymerase transcribes the transfer RNAs and short ncRNAs (Cramer *et al*, 2008; Cooper, 2000). Two other RNA polymerases, the RNA polymerase IV (Pol IV) and the RNA polymerase V (Pol V) can be found in plants. Pol IV seems to be implicated in the transcription of small interfering RNAs and Pol V in gene silencing (Zhou and Law, 2015).

1.3. The RNA polymerase II CTD

Pol II differs from Pol I or Pol III by the presence of a carboxy-terminal domain (CTD) on its largest subunit, Rpb1 also called AMA-1 in *Caenorhabditis elegans* (Bowman and Kelly, 2014). The CTD is composed of repetitions of the heptapeptide sequence Y_1-S_2-P_3-T_4-S_5-P_6-S_7 and the number of repetitions depends on the complexity of the organism, with 26 repetitions for the yeast *Saccharomyces cerevisiae*, 42 for *C. elegans* and 52 for mammals (Corden, 1990). The CTD is not required for the catalytic activity of the polymerase as an *in vitro* study showed that Pol II has similar transcriptional activity with or without the CTD (Kim and Dahmus, 1989). However, *in vivo*, the deletion of the CTD is lethal in yeast, *Drosophila melanogaster*, and mammalian cells even if it can be truncated (Corden, 2013). Indeed, in yeast, the CTD can be cut to leave only 10 repeats without affecting the viability (Nonet *et al*, 1987) and mammalian cells expressing a CTD with 35 repetitions can be maintained in culture despite showing lower cell viability (Chapman *et al*, 2005). All those studies thus suggested that the CTD is not required for the catalytic activity of the polymerase but is still important for several cellular functions. That could be explained by the fact that the CTD does not play a role in the

Table 1: CDKs orthologs in various species. The principal CDKs phosphorylating the CTD serines with their orthologs in several species and their CTD target(s) (Malumbres, 2014).

H. sapiens	*C. elegans*	*S. pombe*	*S. cerevisiae*	*D. melanogaster*	CTD target(s)
CDK7	CDK-7	Mcs6	Kin28	Cdk7	CTD-Ser5; CTD-Ser7
CDK8	CDK-8	Cdk8	Srb10	Cdk8	CTD-Ser2;CTD-Ser5
CDK9	CDK-9	Cdk9	Bur1	Cdk9	CTD-Ser2; CTD-Ser5
CDK12/CDK13	CDK-12	Lsk1	Ctk1	Cdk12	CTD-Ser2

transcription process itself but rather in the processing of the pre-mRNA by recruiting proteins close to the RNA polymerase II. Indeed, each of the CTD tyrosine, threonine, and serine residues can be phosphorylated or glycosylated and the proline residues can be isomerized. In addition, it exists non-consensus repeats containing lysines and arginines which can be modified by acetylation, methylation or ubiquitination (Ranuncolo *et al*, 2012; Zaborowska *et al*, 2016). Thanks to all these post-translational modifications, the CTD is thought to act as a platform for proteins docking on Pol II allowing the coupling of mRNA transcription, maturation, and processing (Zaborowska *et al*, 2016).

The most studied modifications are the phosphorylations, particularly the phosphorylation of the serine in position 2 (CTD-Ser2) and in position 5 (CTD-Ser5) because they are the most abundant. Indeed, immunopurification and mass-spectrometry analyses of CTD peptides showed that the phosphorylated CTD-Ser2 (CTD-Ser2P) and CTD-Ser5 (CTD-Ser5P) represent 75% of the total phosphorylated amino acids while the phosphorylated threonine 4 represents only 15% and the phosphorylated tyrosine 1 and serine 7 represent less than 5% (Schüller *et al*, 2016).

1.4. The CTD kinases

The kinases phosphorylating the CTD serines belong to the Cyclin-Dependent Kinase (CDK) family. CDKs are serine/threonine kinases whose activity requires the binding of a regulatory subunit, a cyclin. The CDKs can be divided into two groups, CKDs that can bind multiple cyclins and are involved in the regulation of the cell cycle and CDKs that can bind only one cyclin and are involved in the regulation of transcription (Malumbres, 2014). There are four principal CDKs phosphorylating the CTD serines and involved in the transcription cycle (Table 1). CDK7 is part of the TFIIH transcription factor and is responsible with its cyclin H for the phosphorylation of the CTD-Ser5 and CTD-Ser7 during the initiation of the transcription cycle (Malumbres, 2014). CDK8 is part of the Mediator complex and seems to phosphorylate the CTD-Ser2 and CTD-Ser5 *in vitro*. *In vivo*, CDK8 seems to phosphorylate the CTD-Ser5 and to positively regulate the transcription even if its function is highly debated (Zaborowska *et al*, 2016; Tsutsui *et al*, 2011). CDK9 associates with its cyclin T and forms the elongation factor P-TEFb. CDK9 phosphorylates the CTD-Ser2 and CTD-Ser5 allowing complex involved in capping and splicing to dock on Pol II (Malumbres, 2014). CDK12 is the last kinase involved in the phosphorylation of the CTD. It is implicated with its cyclin K in the phosphorylation of the bulk of CTD-Ser2 in order to recruit complexes necessary for the cleavage and polyadenylation of the pre-mRNA (Malumbres, 2014). This last kinase is particularly interesting because, for a long time, the scientific community thought that CDK9 was the genuine and only CTD-Ser2 kinase (Bartkowiak and Greenleaf, 2011). However, in 2010, Bartkowiak and co-workers proved that CDK12 is a *bona fide* CTD-Ser2 kinase. Indeed, they showed by Western blot that treating cells with dsRNAs targeting exons of CDK12 dramatically reduces the phosphorylation of CTD-Ser2 but has no impact on the phosphorylation of CTD-Ser5 (Bartkowiak *et al*, 2010).

The other residues of the CTD can also be phosphorylated. Indeed, CDK9 can also phosphorylate the CTD-Thr4 (Hsin *et al*, 2011) and the CTD-Tyr1 can be phosphorylated by the c-Abl kinase (Baskaran *et al*, 1993) but the biological relevance of these marks remains largely untested.

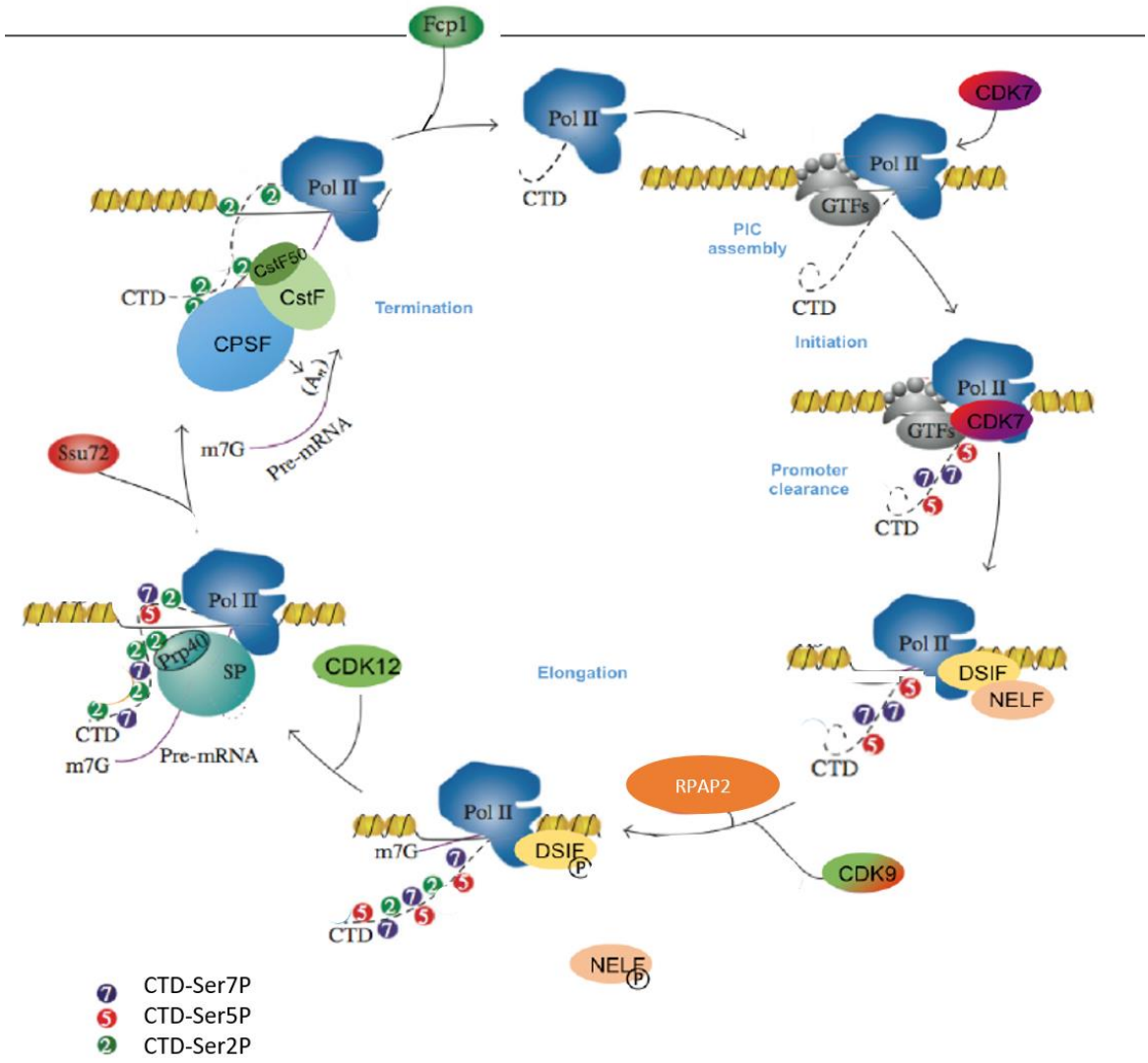


Figure 1: Pol II transcription cycle with the CTD code. The non-phosphorylated CTD Pol II with general transcription factors (GTFs) and the Mediator complex form the Pre-Initiation Complex (PIC). Conformational changes of the PIC allow the initiation step to begin, but at this stage, the polymerase is still linked to the promoter. The phosphorylation of the CTD-Ser5 and CTD-Ser7 by the subunit CDK7 of the transcription factor TFIIH release Pol II, leading to the elongation step. However, this early elongation is stopped by Pol II pausing, removed by the phosphorylation of DSIF and NELF by the subunit CDK9 of P-TEFb. CDK9 also phosphorylates the CTD-Ser5 allowing the addition of the cap (m7G). During the elongation step, CDK9 phosphorylates the CTD-Ser2 and RPAP2 begins to dephosphorylate the CTD-Ser5. This allows to recruit the spliceosome (SP). The last step of the transcription cycle is the termination where Ssu72 dephosphorylates the CTD-Ser7 and CTD-Ser5 and where CDK12 phosphorylates the CTD-Ser2. This allows to recruit the cleavage and polyadenylation machinery (CPSF and CstF). When the cycle is finished, Fcp1 dephosphorylate the CTD-Ser2 and the polymerase can begin a new cycle (adapted from Zhang *et al*, 2012a).

1.5. The CTD phosphatases

The phosphorylation of the CTD serines is very dynamic with phosphorylation and dephosphorylation during the transcription cycle. Indeed, the CTD-Ser7 is phosphorylated by CDK7 at the beginning of the transcription cycle and is dephosphorylated by a phosphatase called Ssu72 in the 3' region of the transcript unit. It is a necessary event for transcription termination (Zhang *et al*, 2012b). The CTD-Ser5P is important to recruit the capping machinery. It is dephosphorylated by RPAP2 in the 5' region and by Ssu72 in the 3' region (Egloff *et al*, 2012; Krishnamurthy *et al*, 2004). The phosphate on the CTD-Ser2 is removed by Fcp1 near the termination site (Cho *et al*, 2001).

1.6. The RNA polymerase II transcription cycle

The transcription is a six-step cycle comprising the Pre-Initiation Complex (PIC) assembly, the open complex formation, the initiation, the promoter clearance, the elongation, and the termination (Figures 1 and 2) (Svejstrup, 2004).

The PIC is composed of the general transcription factors (TFIIB, TFIID, TFIIE, TFIIH, and TFIIF), the Mediator complex and the non-phosphorylated CTD Pol II. This complex assembles at the promoter region. The promoter region is composed of several elements such as the TATA, the CAAT or the GC box (Karp, 2010). The first step is the binding of TFIID to the TATA box (25bp upstream the transcription start site) thanks to its TATA Binding Protein (TBP). TFIIB next interacts with TBP and recruits TFIIF accompanied by Pol II. TFIIH and TFIIE next join the complex. TFIIE regulates TFIIH that possesses a helicase activity to open the DNA, an ATPase activity and a kinase activity mediated by CDK7 (Orphanides *et al*, 1996). The Mediator complex is also part of the Pre-Initiation Complex. It serves as a link between the PIC and non-general transcription factors to "transmit" information from these factors to the polymerase (Björklund and Gustafsson, 2005). The next step is the open complex formation performed by the action of TFIIH which opens the DNA. This leads to conformational changes of the PIC to form an open complex ready to transcribe (Cheung and Cramer, 2012; Kumar *et al*, 1998). The initiation step can thus start and the polymerase synthesizes the first nucleotides. However, at this stage, Pol II is still linked to the Pre-Initiation Complex and to release it, CDK7 phosphorylates the CTD-Ser5 and CTD-Ser7. It dissociates the Mediator from the PIC and releases the polymerase leading to the elongation step of the transcription (Wong *et al*, 2014). Nevertheless, this early elongation is stopped by Pol II pausing due to both the intrinsic capacity of the polymerase for pausing and to the action of the regulatory factors DSIF (DRB-Sensitivity-Inducing-Factor) and NELF (Negative Transcription Elongation Factor) (this last complex is absent in some species of nematodes, *S. cerevisiae*, and plants) (Yamaguchi *et al*, 2013). This pausing allows the regulation of the transcription. Indeed, it is a rate-limiting step that can impact the mRNA abundance because the polymerase can even continue the transcription cycle or "fail" from the gene to stop the transcription. It has also other functions as helping the addition of the cap to the pre-mRNA or the coupling of the translation with the transcription in bacteria (Mayer *et al*, 2017). Pol II pausing is relieved by the phosphorylation of DSIF and NELF by the CDK9 kinase as part of the factor P-TEFb. Those phosphorylations dissociate NELF from the DNA and change the conformation of DSIF which becomes a positive elongation factor (Bowman and Kelly, 2014; Yamaguchi *et al*, 2013). CDK9 is also responsible, at least partly, for the phosphorylation of the CTD-Ser5 allowing the recruitment of the capping enzymes (Zaborowska *et al*, 2016). Indeed, a mutant yeast expressing alanine

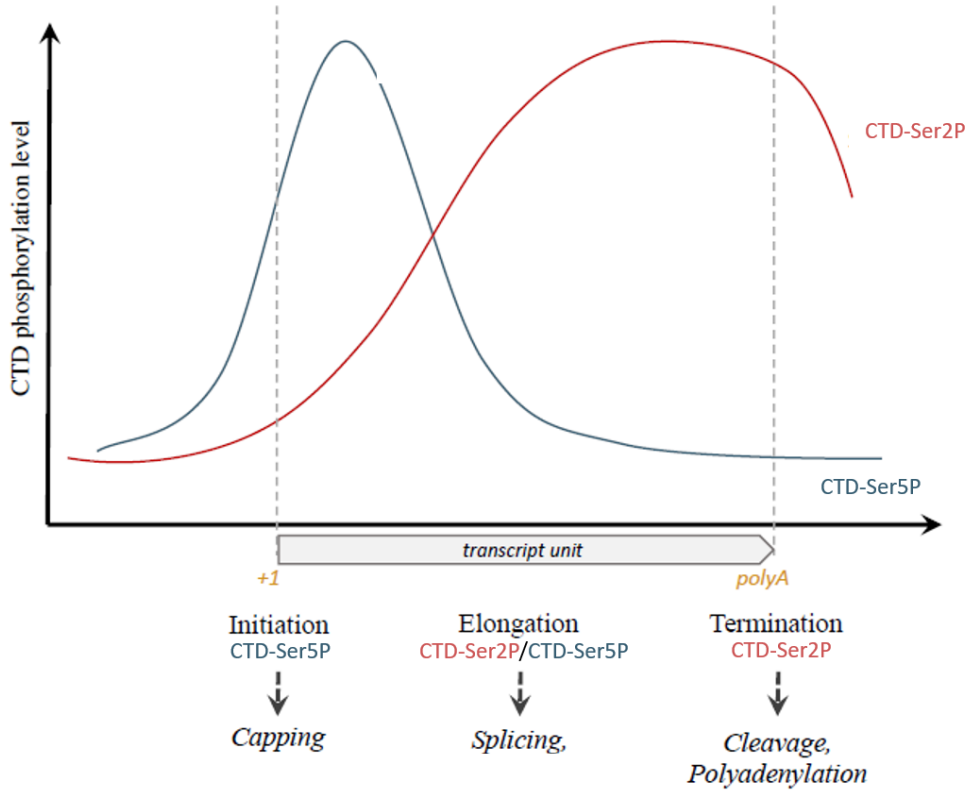


Figure 2: CTD-Ser2 and CTD-Ser5 phosphorylation during the transcription cycle. The CTD-Ser5P is present during the initiation and elongation step and plays a role to recruit the capping machinery and the spliceosome. The CTD-Ser2P is present during the elongation and termination step and plays a role in the recruitment of the spliceosome and the cleavage and polyadenylation factors.

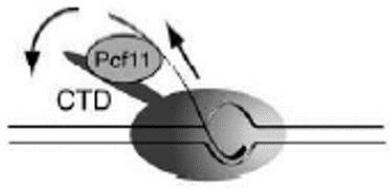


Figure 3: Principle of the allosteric model. As the cleavage and polyadenylation factor Pcf11 is linked to the CTD-Ser2P and to the nascent transcript, when the CTD changes its conformation at the poly(A)-signal, Pcf11 follows it, generating some pressure on the nascent RNA. This disrupts the complex formed between the nascent RNA and the template DNA leading to the disassembly of the elongation complex (Zhang *et al*, 2005).

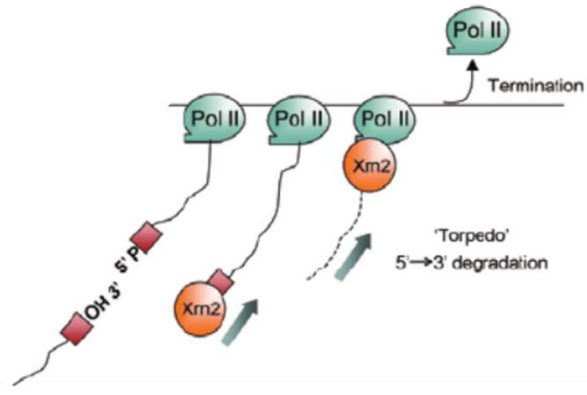


Figure 4: Principle of the torpedo model. The cleavage of the pre-mRNA generates a uncapped 5' PO_4 extremity (5'P) which can serve as a substrate for the enzyme XRN2. XRN2 thus degrades the left-over RNA, reaches Pol II and dissociate the elongation complex (adapted from Kornblihtt, 2004).

instead of serine 5 (CTD-S5A) is lethal, but by fusing the capping enzyme MceI to Pol II expressing a CTD-S5A, the viability is restored suggesting that the only essential function of the CTD-Ser5P is to recruit the capping machinery, at least in yeast (Schwer and Shuman, 2011). After the action of P-TEFb, Pol II is able to transcribe the gene, it is the elongation step. During this step, CDK9 will phosphorylate the CTD-Ser2 and RPAP2 will remove the phosphate on the CTD-Ser5. It means that during this step, the rate of CTD-Ser2P increases while the CTD-Ser5P decreases. This presence at the same time of CTD-Ser5P and CTD-Ser2P enables the spliceosome and CDK12 to join the pre-mRNA and Pol II (Bowman and Kelly, 2014; David *et al*, 2011). At the 3' end of the gene, Pol II enters in the last step of the transcription cycle, the termination. During this step, Ssu72 is working and dephosphorylate the CTD-Ser7 and CTD-Ser5 while CDK12 phosphorylates the CTD-Ser2 (Bowman and Kelly, 2014; Krishnamurthy *et al*, 2004; Zhang *et al*, 2012b). This presence of CTD-Ser2P allows the recruitment of the cleavage and polyadenylation complex to the pre-mRNA (Lunde *et al*, 2010). The CTD-Ser2P is next removed by Fcp1 and the dephosphorylated Pol II can begin a new cycle (Cho *et al*, 2001).

There are two alternative models for transcription termination and release of Pol II from the DNA (Briand, 2015). The allosteric model (Figure 3) suggests that the detachment of Pol II is mediated by conformational changes and by interaction with termination factors like Pcf11 (Briand, 2015). Indeed, Pcf11, a cleavage and polyadenylation factor, is able to disassemble the Pol II elongation complex by linking the CTD-Ser2P and the nascent RNA forming a bridge between them. When the polymerase reaches the poly(A)-signal, the CTD changes its conformation, Pcf11 accompanies the CTD, generating some pressure on the nascent RNA. This pressure disrupts the duplex formed between the transcript and the DNA template leading to the dissociation of the elongation complex (Zhang *et al*, 2005). In the second model, the torpedo model (Figure 4), the cleavage of the pre-mRNA is required. This cleavage exposes an uncapped 5' PO_4 extremity which can serve as the substrate for the exonuclease XRN2. XRN2 thus degrades the left-over RNA, moves towards the polymerase and disassembles the elongation complex (Briand, 2015; Luo *et al*, 2006). This model is supported by studies showing that the exonuclease Rat1 (yeast homolog of XRN2) is located at the end of protein-coding genes, or that cells deleted for Rat1 show a termination defect (Kim *et al*, 2004). However, this model has been challenged by reports that termination can occur without pre-mRNA cleavage. Indeed, in yeast, a mutant of Pcf11 has cleavage defect but the termination is not affected (Luo *et al*, 2006). Consequently, a third model is proposed, the hybrid allosteric-torpedo model because neither the Pol II conformational modifications nor the RNA degradation by Rat1 is sufficient to induce the termination. This new model encompasses the two others and is based on the discovery that Rat1 promotes the recruitment of 3' end processing factors like Pcf11 at the poly(A)-site. In this hybrid model, Rat1 is part of the cleavage and polyadenylation complex and this complex is responsible for the pre-mRNA cleavage, the left-over RNA degradation and the Pol II conformational changes causing the release of the elongation complex (Briand, 2015; Luo *et al*, 2006).

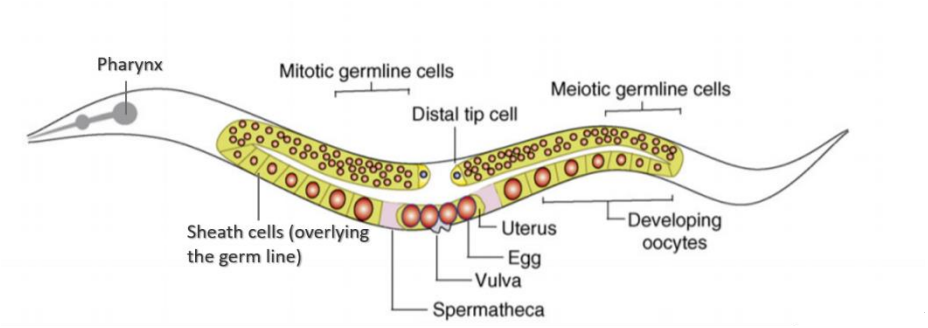


Figure 5: Hermaphrodite reproductive system. The oocytes arise by meiosis from a syncytim of undifferentiated germline cells. They are fertilized by passing through the spermatheca and the newly formed eggs develop into the uterus before being laid by the egg-laying apparatus (adapted from Mukhopadhyay and Tissenbaum, 2007).

2. The nematode *C. elegans*

2.1. Origins of *C. elegans*

C. elegans was introduced in genetic laboratories by Sydney Brenner in 1963. At that time, the DNA structure had just been discovered and nothing was known about how the genes controlled the development of a multicellular organism. Brenner was thus looking for an animal with which he could answer the question "How genes might specify the complex structure found in higher organisms" (Brenner, 1974). To solve this problem, he wanted to study the nervous system of an animal with a quite simple neuronal system and handy to manipulate. After some investigations, he chose the nematode *C. elegans* (Brenner, 2003; Couillault and Kurz, 2010).

C. elegans was first described by Maupas in 1900 (Nigon and Félix, 2018) and first used in laboratories by Dougherty and Calhoun in 1957 (Sterken *et al*, 2015). In those days, there were two *C. elegans* strains. One strain, Bergerac was isolated in Lyon by Nigon and the other one, Bristol, was isolated near Bristol by Staniland in 1956 (Ferris and Hieb, 2015). The two strains are very different. The Bergerac strain exhibits a high spontaneous mutation frequency due to the presence of a high copy number of the *TcI* transposon and is also temperature-sensitive. Indeed, they become sterile when cultivated at temperatures above 18°C (Riddle *et al*, 1997). This explains why most studies using nematodes are made with the Bristol strain and why Dougherty and Calhoun imported this strain in their laboratory in 1957. Concurrently, Brenner unsuccessfully tried to isolate *C. elegans* from his own garden. He thus requested the nematode from Dougherty who sent it to him. Brenner called this strain N2 because it was his second attempt to establish a *C. elegans* line (Sterken *et al*, 2015).

Today the Bristol N2 strain is used as the *wild-type* strain in thousands of laboratories.

2.2. *C. elegans* generalities

C. elegans is a transparent nematode, 1 mm in length when adult, that lives in rotting fruits, rotting plants and in the compost where he feeds on bacteria (Frézal and Félix, 2015). It exists mainly as a hermaphrodite organism although male arises at low frequency (<0.2%) (Corsi *et al*, 2015). It is an eutelic organism meaning that it has a fixed number of cells, namely 959 somatic cells for the hermaphrodite worm and 1031 for the male (Wood, 1988). Of these 959 cells (or 1031), 302 are neurons that make more than 7000 chemical synapses and gap junction connections. It explains why *C. elegans* is widely used in neurobiology (Corsi *et al*, 2015). The remaining cells constitute the epidermis, the muscles, the digestive system, and the reproductive system. The epidermis, composed of syncytial cells, forms the outer epithelial layer and secretes a protective layer composed of collagen, lipids, and glycoproteins called cuticle. The cuticle determines the shape of the body, protects the nematode and provides anchoring points for mononucleated muscle cells forming the body-wall muscles, responsible for animal movements. In addition to the body-wall muscles, *C. elegans* has pharyngeal muscles, vulval muscles, and anal muscles (Corsi *et al*, 2015). The pharyngeal and anal muscles accompanied by neurons form the pharynx and the anus, which are part of the digestive system with the mouth and the intestine (Corsi *et al*, 2015). The vulval muscles are part of the reproductive system, which is different between the hermaphrodite and the male organisms. The hermaphrodite reproductive system is divided into three parts (Figure 5), the somatic gonad, the germline cells, and the egg-laying apparatus. The somatic gonad, located alongside the

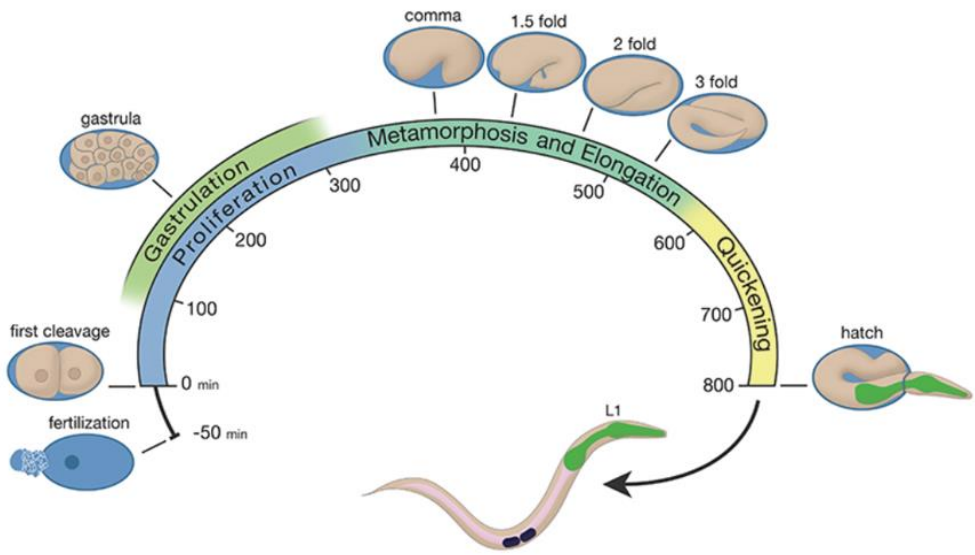


Figure 6: *C. elegans* embryonic development. The embryonic development is composed of six steps, the proliferation, the gastrulation, the metamorphosis, the elongation, the quickening and the hatching (Wormatlas).

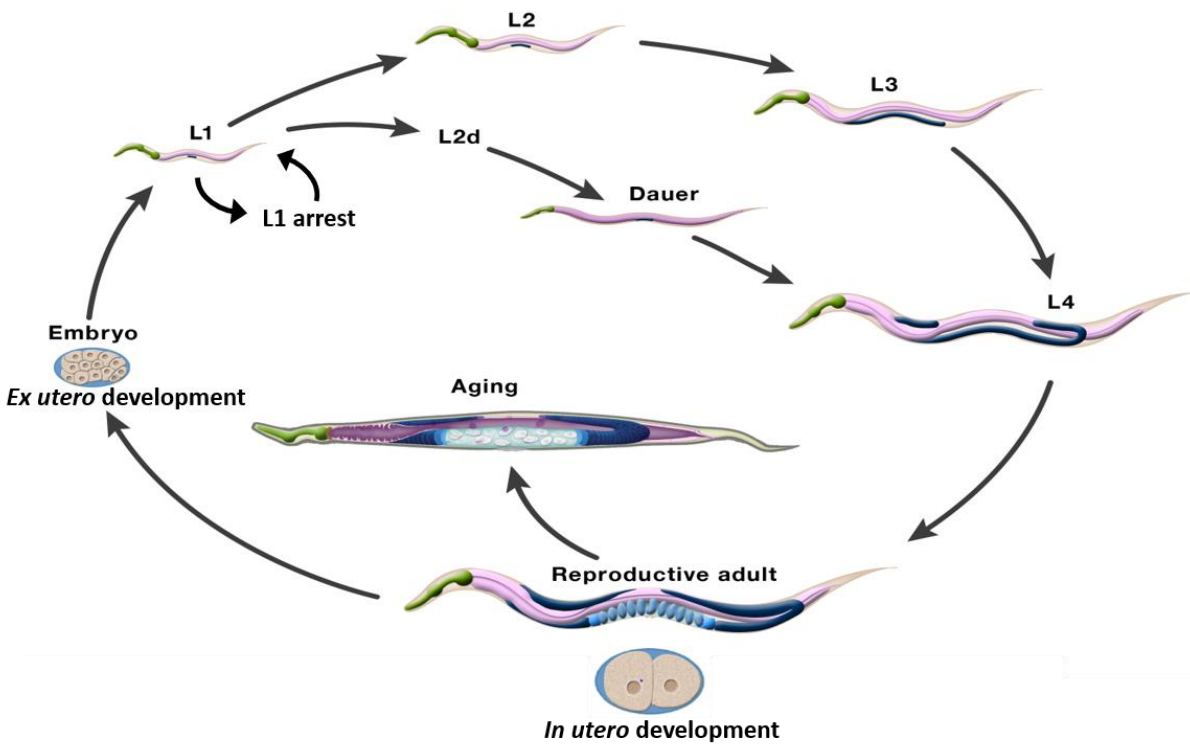


Figure 7: *C. elegans* life cycle. In its normal life cycle, *C. elegans* has four larval stage and one reproductive adult stage. However, under bad conditions such as starvation, high temperature or overcrowding, *C. elegans* has an alternative life cycle. They become dauer worms which are more resistant to hard conditions. It also exits a third possible life cycle called the L1 arrest meaning that worms stop their development in their first larval stage. This happens when L1 larvae hatch in absence of food (adapted from Wormatlas).

intestine, consists of two U-shaped tubes and is composed of the distal tip cell (DTC), which is a single somatic cell constituting a cap over the gonad, the sheath cells overlying the germ cells, the spermatheca containing the sperm and the uterus. The germline cells represent a syncytium of undifferentiated cells located near the DTC and generating the oocytes through meiosis. These oocytes are fertilized by passing through the spermatheca and the newly formed eggs develop into the uterus before being laid by the egg-laying apparatus composed of the vulva, neurons, and muscles. The male reproductive system consists of a single U-shaped gonad that connects to a seminal vesicle stocking the sperm. The *vas deferens* next conducts the sperm to the cloaca in the tail of the worm (Corsi *et al*, 2015; Lints and Hall, 2009).

2.3. *C. elegans* life cycle

2.3.1. Normal life cycle

Right after the fertilization of the oocyte, an impermeable eggshell is formed around the zygote allowing it to develop independently of the adult. Despite this eggshell, the eggs stay inside the uterus and the first phase of the embryogenesis begins (Figure 6). It is the proliferation step where cells undergo stereotyped division. They are all similar in their cytoplasmic contents. In addition, some cells undergo apoptosis and some migrate away from their sister cells. This is called "global cell sorting". At approximately 30 cells, the eggs are laid, the second step of the embryogenesis, the gastrulation, begins and the proliferation keeps on. The gastrulation allows, by "global cell sorting", the formation of functional cell groups and of the ectoderm, endoderm, and mesoderm. The next stage is the morphogenesis where cells become specialized and tissues are formed. The elongation and the quickening next take place. These are the steps where the embryos elongate to look like a worm (comma shape to 3 fold shape) and where muscle movement can be seen. The last step is the hatching (Hall *et al*, 2017). The released larvae (L1) will pass through three other larval stages (L2, L3, and L4), marked by the progressive development of the reproductive system, before becoming reproductive adults (Figure 7). These adults will finally lay eggs until they have utilized all their sperm (approximately 3 days) (Corsi *et al*, 2015).

2.3.2. The dauer stage

Under starvation, high temperature or overcrowded population, the L1 larvae follow an alternative cycle to become dauer worms (Figure 7). Dauer worms are distinct from L2 or L3 larvae. Indeed, they are thinner and possess a cuticle resistant to hard conditions. This cuticle surrounds all the worm including the mouth, preventing it from feeding. Dauer worms can survive for months and can search for food by standing on their tail and attaching to insects and vertebrates which transport them to a suitable environment. This comportment is called nictation. When they find their new environment, they molt and resume their development from the L4 stage (Corsi *et al*, 2015; Lee *et al*, 2012).

2.3.3. The L1 arrest

In *C. elegans*, there is another alternative cycle, the L1 arrest (Figure 7). This arrest is completely different from the dauer stage because the worms' morphology doesn't change. Indeed, nematodes just stop their development in their first larval stage and do not reach the L2 stage (Baugh, 2013). This alternative state occurs when embryos hatch in the absence of food due to a perturbation of the insulin-like IGF-1 pathway. Indeed in the presence of food, nutrients are detected by chemosensory and other neurons, which leads to the production of Insulin-Like

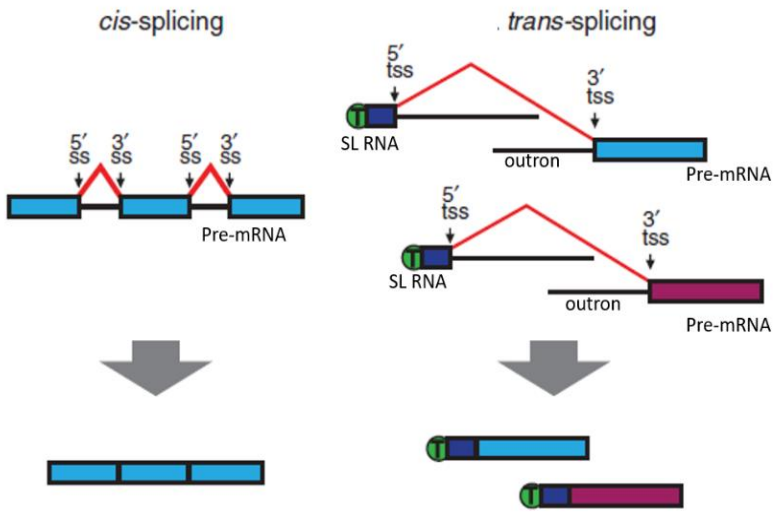


Figure 8: *Cis*-splicing and *trans*-splicing. In *cis*-splicing, the 5' splice site (5'ss) and the 3'splice site (3'ss) are linked by the spliceosome to join two exons (light blue square) from the same pre-mRNA molecule. In *trans*-splicing, the first exon from the pre-mRNA is joined with an exon from a non-coding RNA (the SL RNA). This SL RNA contain a TMG cap (green circle with a T), an exon called the spliced leader (dark blue square), and a 5' splice site (5'tss). The *trans*-splicing thus excises the long outron of the pre-mRNA and joins the spliced leader with the TMG cap to the first exon of the mRNA (adapted from Lasda and Blumenthal, 2011).

Peptides (ILPs). These ILPs activate the Insulin-Like Receptor called DAF-2. The pathway downstream of the activation of DAF-2 and mediated by the PI3K leads to the expression of genes involved in worm postembryonic development. By contrast, when no nutrients are available, DAF-2 stays inactive leading to an L1 arrest (Kaplan and Baugh, 2016). L1 arrested larvae are able to survive for 4 weeks due to the expression of genes involved in stress resistance and can resume the normal life cycle when food is available (Baugh, 2013).

2.4. *C. elegans* genome

2.4.1. Generalities

C. elegans is a diploid organism and has six pairs of chromosomes. He has five autosomal chromosomes (I, II, III, IV, and V), and one sexual chromosome (X). As already said, *C. elegans* exists in two sex types. The hermaphrodites possess two X chromosomes (XX) but males possess only one X chromosome (XO). This results from the non-disjunction of the X chromosomes pair during gametogenesis (Corsi *et al*, 2015).

All of the *C. elegans* chromosomes are holocentric meaning that they don't have traditional centromeres. Indeed, the entire length of the chromosome act as the centromere and during the mitosis, the spindle microtubules attach along the whole length of the chromosome. This feature allows the formation of extrachromosomal DNA array which can be inherited by the progeny (Corsi *et al*, 2015). This is relevant during transgenesis as these arrays are not always easily distinguished from genomic insertions.

The particularity of the nematode's genome is the organization of approximatively 15% of genes in operons and the existence of a process called *trans*-splicing (Corsi *et al*, 2015).

2.4.2. Operons and *trans*-splicing

The *C. elegans* genome counts 1250 operons. Operons are clusters of genes under the control of a single promoter. In bacteria, genes in operons are usually part of the same pathway but it is not necessarily the case for nematodes. Indeed, the polycistronic pre-mRNA is co-transcriptionally processed into monocistronic mRNAs. This processing is made by the polyadenylation of the 3' end of the upstream gene and the *trans*-splicing of the 5' end of the downstream gene. The *trans*-splicing is the joining of two RNA exons from two different RNA molecules. This is distinct from normal or *cis*-splicing, which is the joining of two RNA exons from the same RNA molecule (Figure 8) (Blumenthal *et al*, 2018).

In *cis*-splicing (Figure 8), two exons from the same pre-mRNA are joined by the action of the spliceosome, catalyzing two transesterification reactions. The spliceosome consists of small nuclear ribonucleoproteins (snRNPs) and non-snRNP proteins. Each snRNP called U1, U2, U5, and U4/U6 consists of a small nuclear RNA (snRNA) (or two in the case of U4/U6), seven proteins Sm and a variable number of particle-specific proteins. Spliceosome assembly occurs by the ordered interaction of the snRNPs and other factors with the 5' splice site (5'ss), the branch site, the polypyrimidine tract, and the 3' splice site (3'ss) (Will and Lührmann, 2006).

In *trans*-splicing (Figure 8), one exon belongs to the pre-mRNA and the other belongs to a specialized non-coding RNA (SL RNA). This SL RNA contains a 5' trimethylguanosine (TMG) cap, a 22 nucleotide exon called the spliced leader (SL) and an Sm binding site. Consequently, the SL RNA also forms snRNP by getting together with Sm proteins. The difference with *cis*-splicing is that the SL RNA is consumed during the splicing reaction.

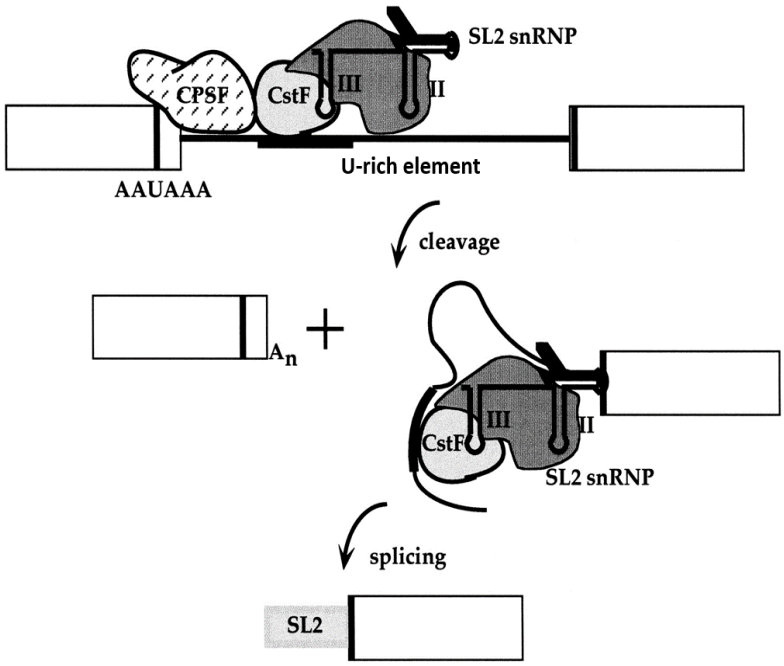


Figure 9: SL2 *trans*-splicing process. The *trans*-splicing of genes in position 2 and over in operons is realized by the SL2 snRNP. This complex interacts with the CstF factor involved in the cleavage and polyadenylation of the pre-mRNA from the upstream gene. It is this specific interaction which allow to recruit the SL2 snRNP and not the SL1 snRNP. In addition, a U-rich element is necessary because the SL2 snRNP binds to this sequence (adapted from Evans *et al*, 2001).

Indeed, the SL RNA contains a 5' splice site allowing the SL to be transferred to the pre-mRNA. The main function of the *trans*-splicing is the protection of the 5' extremity of pre-mRNAs from operons genes. Indeed, eukaryotic pre-mRNAs are capped by the capping machinery but genes within operons are not recognized by the usual capping machinery. Consequently, their pre-mRNAs are protected by the 5' TMG cap donated with the SL. The *trans*-splicing has also other functions such as the removal of outrons (5' non-coding sequence of the pre-mRNA). This outron probably contains AUG codons so if it is not removed, there is a risk that the ribosome begins the translation process by using one of these out-of-frame start codons. Therefore, the *trans*-splicing, by taking off the outron allows being sure that only the correct AUG codon is used (Lasda and Blumenthal, 2011).

C. elegans has two distinct SL RNAs called SL1 RNA and SL2 RNA. They share similar features, they both possess the TMG cap and both associate with Sm proteins. However, the SL1 RNA is used to *trans*-splice single genes and the genes in position 1 within operons while SL2 RNA is specific for the *trans*-splicing of genes located further in operons (position 2 and over) (Lasda and Blumenthal, 2011). Consequently, a mechanism that regulates the selection of the SL2 RNA should exist (Figure 9). Several studies suggest that the *trans*-splicing of the pre-mRNA from the downstream gene is coupled with the polyadenylation of the pre-mRNA from the upstream gene. Indeed, the SL2 snRNP interacts with the 3' end processing factor, CstF (Cleavage stimulatory Factor which binds to a U/GU-rich element) (Evans *et al*, 2001). CstF also interacts with CPSF (Cleavage and Polyadenylation Specificity Factor) - which binds to the Polyadenylation Signal (PAS) AAUAAA - forming a link between CPSF and the SL2 snRNP (Murthy and Manley, 1995). In addition, a U-rich element is necessary for the SL2 *trans*-splicing probably because the SL2 snRNP binds to this U-rich element (Huang *et al*, 2001).

2.5. *C. elegans* genome engineering

Genetic mutations arise spontaneously at a low level in the process of evolution. In the lab, however, they can be induced by several techniques to create mutant organisms used in genetics to study gene function.

2.5.1. Genome-wide mutagenesis

Historically, mutant organisms were produced using mutagenic agents targeting the whole DNA and thus inducing mutations in the entire genome. These mutagenic agents are diverse chemical products, high-energy radiations or transposons. The transposon-mediated technique consists of disrupting genes by the random insertion of a transposon after the action of the appropriate transposase. Indeed, by crossing one *C. elegans* strain (EG1470) expressing the exogenous transposon *Mos1* with one strain (EG2762) expressing the transposase, a hybrid strain, where the transposon is randomly inserted into the genome, is obtained. This hybrid strain has disrupted genes that can give interesting phenotypes (Kutscher and Shaham, 2014).

2.5.2. Gene-targeted mutagenesis

The previous genome-wide mutagenesis is generally used for forward genetic screens. By contrast, reverse genetic screens are usually performed using gene-targeted mutagenesis. Techniques based on homologous recombination between exogenous DNA and the organism's genome were thus developed in order to achieve mutant organisms with mutations only in a locus of interest. (Kutscher and Shaham, 2014).

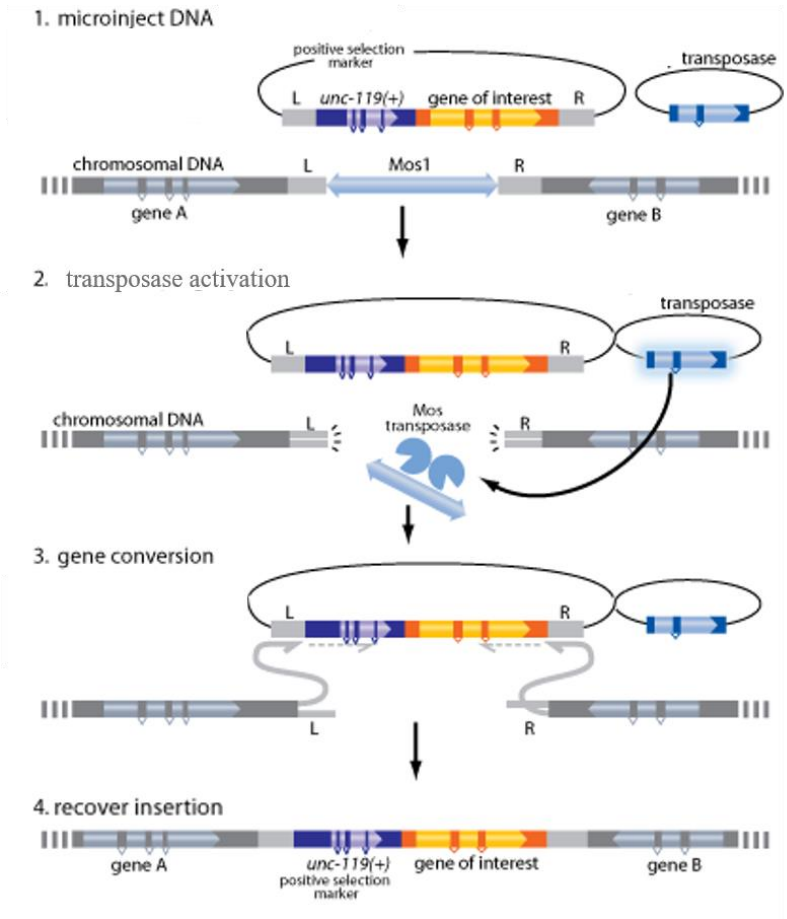


Figure 10: The MosSCI technique. The MosSCI technique is based on the previous insertion of the exogenous *Mos1* transposon in intergenic regions of the genome. With the injection of the appropriate transposase, the transposon is removed leading to a double-strand DNA break. This break is repaired by homologous recombination with an injected repair template containing homologous sequence to the extremity of the break along with a positive selection marker, usually *unc-119(+)* (adapted from Frøkjaer-Jensen *et al*, 2008).

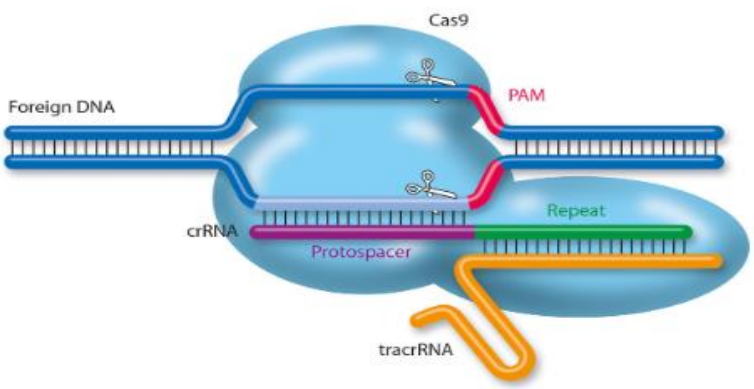


Figure 11: The CRISPR/Cas9 technique. The endonuclease Cas9 identifies its target DNA by binding to the PAM motif and after by hybridization of the crRNA with its complementary DNA. When Cas9 has identified its target, it generates a double-strand DNA break (Max-Planck-Gesellschaft website).

Homologous recombination was widely used in yeast but this method is not very effective in *C. elegans.* Indeed, the injection of exogenous DNA into the nematode generates large extrachromosomal arrays or random integration into the genome (Evans, 2006). However, it was demonstrated that in *C. elegans,* a DNA break generated by the excision of the endogenous *Tc1* transposon can be repaired by homologous recombination. This suggests that recombination functions in nematodes if a DNA double-strand break is previously induced (Plasterk and Groenen, 1992). There are two methods based on the induction of a break into the worm's genome and the reparation of this break by recombination, the transposon-based technique and the CRISPR/Cas9 technique (Kutscher and Shaham, 2014).

The transposon-based technique is based on the previous insertion of *Mos1* transposon near or in the gene of interest. Approximately, 13000 loci were tagged with the transposon by the NemaGENETAG consortium. Consequently, when the desired strain is chosen among the 13000, the injection of the *Mos1* transposase generates a DNA break which is repaired by homologous recombination with an injected repair template. This technique is called MosTIC (*Mos1*-excision Transgene-Instructed gene Conversion) and allows to introduce point mutations, deletions or insertions near or in a locus of interest (Kutscher and Shaham, 2014). Similarly, another technique, the MosSCI (Figure 10) (*Mos1* mediated Single Copy Insertion) is also based on the previous insertion of the *Mos1* transposon, on the action of the transposase and on the reparation of the generated break by homologous recombination. The difference with the MosTIC is that in the MosSCI, the *Mos1* transposon is situated in intergenic regions allowing the incorporation of a stable mutated copy of a gene. Because the insertion does not occur at the endogenous locus, the transgene has to contain the promoter and the terminator of the gene of interest (Frøkjaer-Jensen *et al,* 2008).

To identify mutants for the insertion of the desired mutation, a positive selection marker is added in the repair template. This marker is usually *unc-119(+)* (Figure 10). It is a rescuing marker. Indeed, usually, worms that are used for MosTIC or MosSCI have a mutation in the *unc-119* gene leading to paralysis. Consequently, by placing the rescuing marker in the repair template, it will be incorporated into the genome along with the gene of interest leading to non-paralyzed worms. These non-paralyzed worms normally carry the desired modification (Frøkjaer-Jensen *et al,* 2008).

The other method, the CRISPR/Cas9 technique is also based on the presence of a double-strand DNA break which is repaired by homologous recombination with a donor repair template (Dickinson and Goldstein, 2016). The difference with the transposon-based technique is that the break is not induced by the removal of a transposon but by the action of the endonuclease Cas9. Indeed, the Cas9 endonuclease can be guided to the desired locus (Figure 11) by the action of a single guide RNA molecule (sgRNA) composed of a crRNA (CRISPR RNA), which, by base-pairing with the genome, determines the target specificity and a tracrRNA (trans-activating CRISPR RNA), which activates Cas9. In addition to this base-pairing, Cas9 identifies its target DNA by binding to a trinucleotide sequence (usually an NGG) called the Protospacer-Adjacent-Motif (PAM). Once Cas9 has recognized its target (by the identification of the PAM and the hybridization of the crRNA with the genome) it generates a double-strand DNA break. This break can be repaired by homologous recombination with a donor repair template containing homologous sequences flanking the break (Dickinson and Goldstein, 2016). With this technique, punctual mutations, insertions or even replacement of a part of a gene can be generated (Dickinson and Goldstein, 2016).

It exists two main screening strategies after a CRISPR/Cas9 experiment to distinguish mutant worms. The first strategy is the use of an antibiotic resistance gene. Indeed the repair template can contain, in addition to the desired mutation and the homology sequences to the genome, a gene conferring resistance to an antibiotic. Thereby, if the CRISPR/Cas9 experiment works, the antibiotic resistance gene will be integrated into the organism's genome and only mutant worms will survive to the exposition to the antibiotic (Dickinson and Goldstein, 2016). In *C. elegans*, several antibiotics can be used, such as blasticidin, hygromycin, neomycin (Dickinson and Goldstein, 2016) or puromycin (Semple e*t al*, 2010). The second strategy is called the co-CRISPR strategy. This technique consists of editing two loci during the same microinjection experience. It means that in addition to Cas9, the crRNA, the tracrRNA and the repair template to mutate the desired locus, another crRNA with another repair template are injected to mutate a gene generating a visible phenotype. Consequently, there is a good chance that worms expressing the visible phenotype will be positive for the desired mutation (Dickinson and Goldstein, 2016). Several genes can be used such as *unc-58, unc-109* or *unc-43* which induces worms paralysis when edited or *rol-6, sqt-1* or *dpy-10* which generates a roller phenotype meaning that worms roll on themselves (Arribere *et al*, 2014). *dpy-10* and *sqt-1* have one main advantage compared to others, they confer a different phenotype if the mutation is made on only one allele (roller phenotype) or on the two copies of the gene (dumpy phenotype meaning that worms are shorter than *wild-type*). It is a benefit because as *C. elegans* is a hermaphrodite organism, worms carrying the desired mutation but lacking the screening mutation can be recovered in the progeny of a roller heterozygote worm but not in the progeny of a dumpy homozygote one. Therefore, generally dumpy worms are left out and scientists work with roller worms (Arribere *et al*, 2014).

3. Gene-specific function of CDK12

Regarding the CDKs mediated CTD phosphorylation during the transcription cycle, it was though for a long time that CDKs were required for viability. However, even if CDK9 is indeed essential, for example in *C. elegans* or *D. melanogaster* (Shim *et al*, 2002; Eissenberg *et al*, 2007), CDK12 seems to act in a more gene-specific manner and thus to be important for the expression of only a subset of genes.

3.1. Lsk1 in *S. pombe*

Previous work in the lab (Coudreuse *et al*, 2010) showed that in *Schizosaccharomyces pombe*, the absence of CTD-Ser2P resulting from the deletion of the Lsk1 kinase (the *S. pombe* homolog of CDK12) or from the replacement of all CTD-Ser2 by alanine (CTD-S2A), only barely affects vegetative growth. They also showed by ChIP experiment that as expected, in *wild-type* yeast, the CTD-Ser2P increase towards the 3' region of most coding genes. However, a subset of 150 genes does not follow this pattern and the CTD-Ser2P is already present upstream of the coding region and across the full length of the transcript. Among those genes, the master regulator of differentiation, *ste11*. Ste11 is required under nitrogen starvation to stop the vegetative growth (mitotic cell cycle) and induced sexual differentiation (mating and meiosis) (Van Werven and Amon, 2011). Consequently, Coudreuse and co-workers quantified the expression level of *ste11* by qRT-PCR during nitrogen starvation. They discovered that the level of *ste11* mRNA increases in a *wild-type* yeast while in mutants deleted for the CTD-Ser2P, there is no induction of *ste11*. These mutant yeasts present a strong phenotype under starvation,

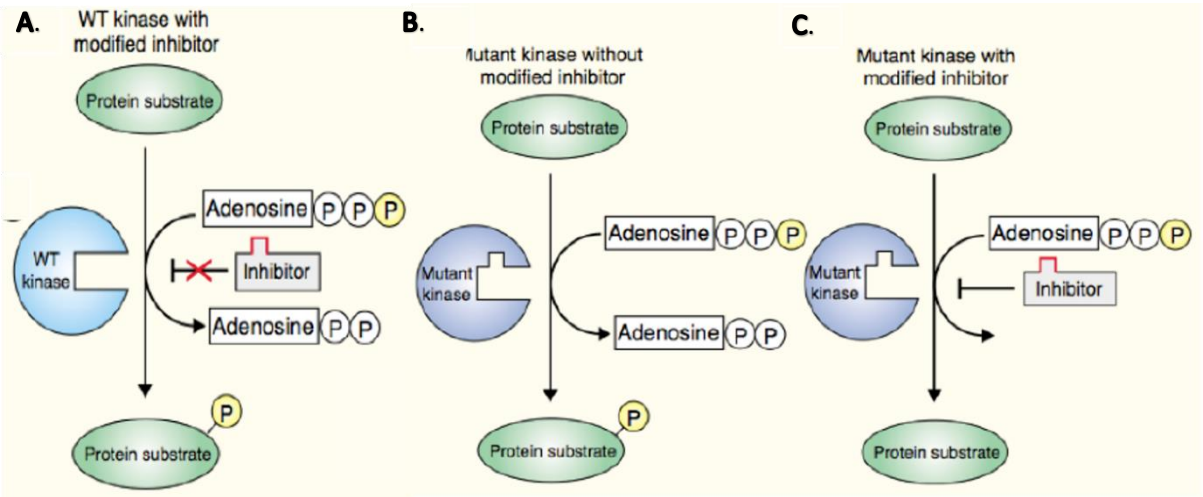


Figure 12: The analog-sensitive kinase. The Shokat analog-sensitive kinase allows to specifically inhibit a desired kinase without impacting on the others. It consists in enlarging the catalytic pocket of the desired protein in order to use an inhibitor (enlarged ATP analog) targeting only this kinase. **(A).** *Wild-type* (WT) kinase are not impacted by the presence of the inhibitor. **(B).** In the absence of inhibitor, the as- kinase (mutant) has normal activity. **(C).** In the presence of the inhibitor, the as- kinase can not play its role (Alaimo *et al*, 2001).

they are unable to mate. In this study, they consequently propose that the phosphorylation of the CTD-Ser2 is not essential for general transcription but is important for a subset of genes implicated in a certain cellular response, here the sexual differentiation (Coudreuse *et al*, 2010).

3.2. Cdk12 in *D. melanogaster*

In *D. melanogaster* (Li *et al*, 2016), the action of Cdk12 is dispensable for viability or for general transcription as its deletion does not compromise the viability and does not affect the expression of housekeeping genes. However, Cdk12 is required for the expression of a subset of genes involved in stress response, like the transcription factor, Nrf2. In a *wild-type* organism, treated with paraquat (a compound which stimulates the production of superoxide radical), the expression of Nrf2 target genes increases. However, in mutants deleted for Cdk12 and treated with paraquat, the expression of these genes does not increase. In addition, flies with Cdk12 knockdown are much more sensitive to stressors than *wild-type* flies. For example, they died more quickly than control flies when exposed to an oxidative stressor. This suggests that Cdk12 is not necessary for the transcription of every Pol II-transcribed gene but is rather important for a small subset of genes implicated in the stress response pathway (Li *et al*, 2016).

3.3. CDK-12 in *C. elegans*

Inhibiting CDK-12 is quite difficult because although RNAi is possible, it has one main disadvantage, the effect can be inherited but it is not permanent and often not fully penetrant (Conte Jr *et al*, 2015). The solution is to chemically inhibit the kinase. However, most of the existing inhibitors target all kinases. Shokat and his co-workers thus developed a method to override the use of such inhibitors. This is the creation of the Shokat analog-sensitive kinase (as- kinase) (Figure 12) (Alaimo *et al*, 2001). This approach consists of the replacement, in the catalytic pocket of the desired kinase, of a bulky amino acid by a smaller one, enlarging the catalytic pocket. This mutation has no impact on the kinase without the action of an inhibitor. Indeed, the as- kinase can link the ATP and phosphorylate its substrate. However, in the presence of an inhibitor, which consists of an enlarged ATP analog able to bind the modified pocket but not the *wild-type* one, the as- kinase cannot phosphorylate its substrate (Alaimo *et al*, 2001). In the lab, a worm strain expressing an analog-sensitive CDK-12 (*cdk-12as* mutant) was created by replacing a phenylalanine by a glycine.

The inhibition of CDK-12as in *C. elegans* induced an L1 arrest accompanied by a decrease in the CTD-Ser2P level. It means that the presence of CDK-12 and the CTD-Ser2P is not necessary for the embryonic development but is required after the hatching to drive the postembryonic development. To try to understand what is causing the L1 arrest, an RNAseq experiment was performed and it was found that the genes impacted by the absence of CDK-12 are located in position 2 and over within operons, in other words, the genes that are *trans*-spliced with the SL2 RNA, which was indeed shown to be markedly reduced. These genes are typically important for worm development.

These results suggest that CDK-12 may be required for the SL2 *trans*-splicing. Considering that the SL2 snRNP interacts with the 3' end processing factor, CstF (Cleavage stimulatory Factor) whose CstF50 subunit binds the CTD-Ser2P (Lunde *et al*, 2010), a reasonable model

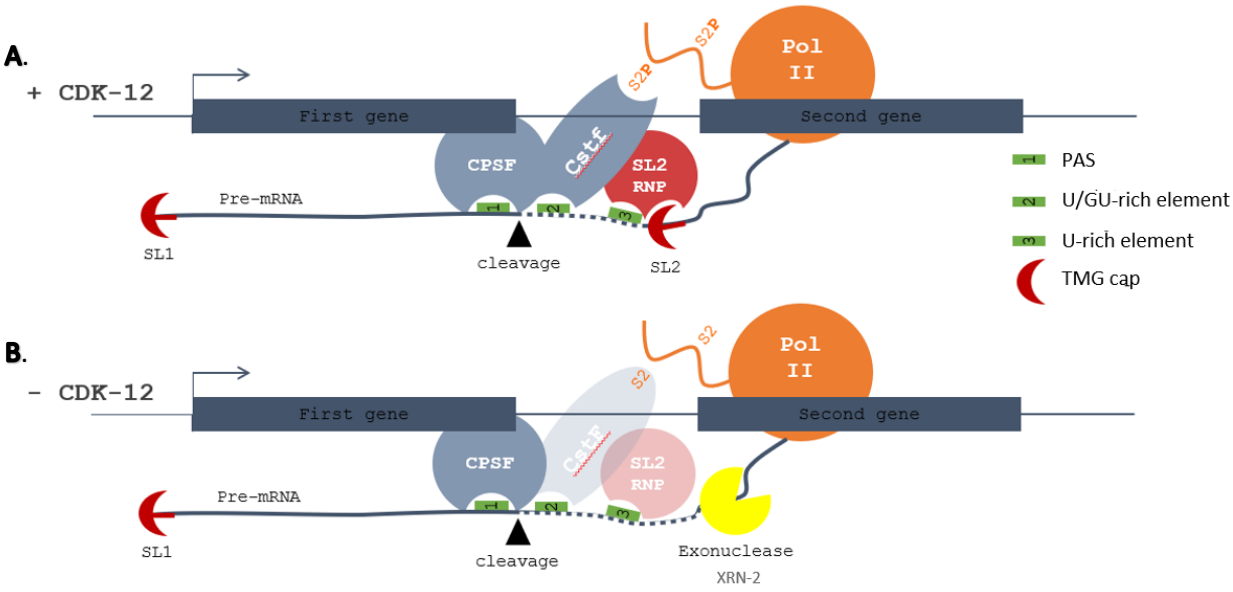


Figure 13: Suggested model to explain the L1 arrest caused by the inhibition of CDK-12as in *C. elegans*. (A). In the presence of CDK-12, the CTD-Ser2 are phosphorylated (S2P) leading to the recruitment of the CstF factor. The CstF factor next interacts with the SL2 snRNP (SL2 RNP) to allow the SL2 mediated *trans*-splicing. The mRNA from genes in position 2 and over in operons are thus protected by the spliced leader 2 (SL2) accompanied by a 5' trimethylguanosine (TMG) cap. **(B).** However, when CDK-12as is inhibited, the CTD-Ser2 are not phosphorylated (S2), the CstF factor and the SL2 snRNP are not recruited and thus the mRNA from genes in position 2 and over in operons are not protected and can be degraded by the exonuclease XRN-2.

(Figure 13) is that upon CDK-12as inhibition, the level of CTD-Ser2P drops, which dissociates CstF and compromise the SL2 *trans*-splicing.

This possibility is supported by the fact that RNAi previously performed by the lab targeting the exonuclease XRN-2 partially restored the normal worm development after CDK-12as inhibition. This suggests (Figure 13) that XRN-2 may be responsible for the degradation of the RNAs from genes in position 2 and over within operons when CDK-12as is inhibited.

Our model recapitulating the current data (Figure 13) also implicates that following the action of XRN-2, a torpedo effect may occur, leading to the detachment of the polymerase and to the termination. Consequently, the amount of protein corresponding to the genes in position 2 and over in operons is reduced at both the transcriptional and post-transcriptional levels, which results in a robust L1 arrest. It is now essential to challenge the model by initiating a new set of experiments.

Objectives of the work

The inactivation of CDK-12 in *C. elegans* results in an L1 arrest while the previous embryogenesis occurs normally. Transcriptomic analysis has revealed that the transcription of a subset of genes localized in position 2 and over within operons and required for development is specifically downregulated. It is therefore likely that the phenotype observed results from a transcriptional defect. Nevertheless, when studying a kinase-substrate relation, it is important to show that the *trans*-mutant (CDK-12as) displays the same phenotype than the *cis*-mutant. Therefore, the first objective of this work is to generate a *C. elegans* strain where each CTD-Ser2, the target of CDK-12, is replaced with an Alanine (a non-phosphorylable amino acid) (CTD-S2A strain). Using this strain, we aim to assess whether the L1 arrest observed upon CDK-12 inhibition is indeed CTD-dependent. This confirmation is particularly important as it has been shown that CDK-12 plays an unrelated role in the translation initiation process (Choi *et al*, 2019; Coordes *et al*, 2015).

Considering that the inactivation of CDK-12 mimics the well-documented L1 arrest resulting from hatching in the absence of food, we hypothesized that the recruitment of CDK-12 to operons is regulated by environmental cues including food availability. In *C. elegans*, it is known that the insulin-like IGF1 pathway connects nutrient availability to development through the Insulin-Like Receptor, DAF-2 and its downstream pathway mediated by the PI3K/Akt or by the Ras/MAPK (Templeman and Murphy, 2018). Indeed, while less well understood, it seems that the Ras/MAPK pathway is also connected downstream of DAF-2 (Nanji *et al*, 2005). We thus postulate that CDK-12 could be regulated by one of these two pathways (PI3K/Akt or Ras/MAPK), which is reminiscent of previous work of our laboratory in fission yeast (Materne et al., 2015) that has shown that the gene-specific recruitment of Lsk1 (CDK-12 yeast homolog) requires a MAP-kinase dependent phosphorylation of an N-terminal extension of the kinase conserved in nematodes (Materne *et al*, 2015). Therefore, the second aim of this work is to construct and purify a tagged version of CDK-12 from a synchronous worm population hatched in the presence or absence of food and analyze the phosphorylation status of the kinase by mass spectrometry.

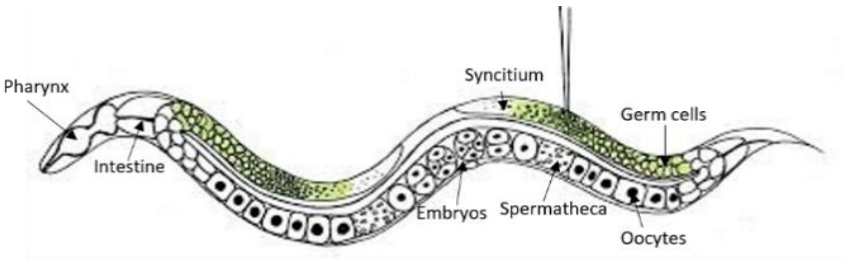


Figure 14: Microinjection in *C. elegans.* The injection is performed in the gonad syncytium (adapted from Daniel *et al*, 2013).

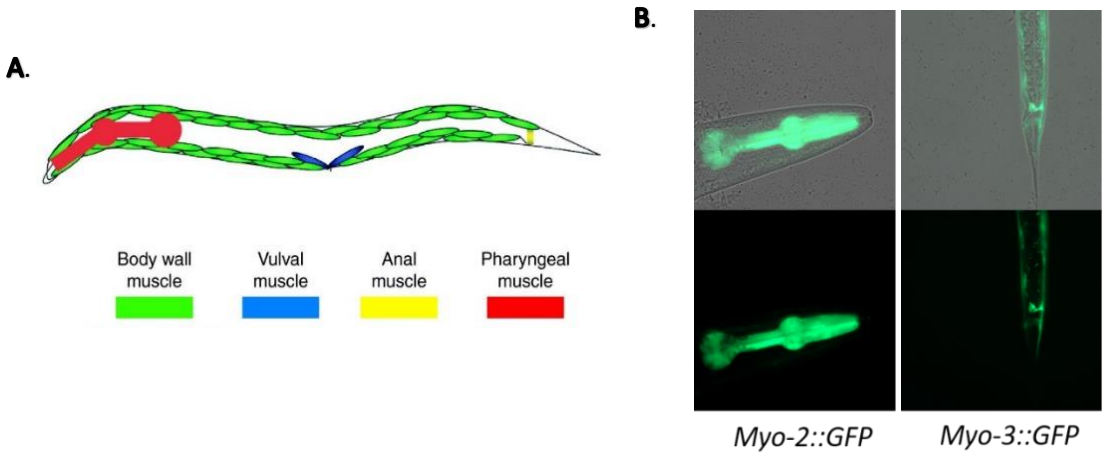


Figure 15: Microinjection training. (A). *C elegans* muscles. *myo-2* is expressed in pharyngeal muscles while *myo-3* is expressed in body-wall muscles, vulval muscles and anal muscles (adapted from Fox *et al*, 2007). **(B).** F1 worms expressing either *myo-2::GFP* or *myo-3::GFP* after microinjection of parental worms.

A.

Genomic CTD:	CTD-S2S:	CTD-S2A:
L S P R T P S	L S P R T P S	L S P R T P S
Y G G M S P G V	Y G G M S P G V	Y G G M S P G V
Y S P S S P Q	Y S P S S P Q	Y A P S S P Q
F S M T S P H	F S M T S P H	F S M T S P H
Y S P T S P S	Y S P T S P S	Y A P T S P S
Y S P T S P A A G	Y S P T S P A A G	Y A P T S P A A G
Q S P V S P S	Q S P V S P S	Q A P V S P S
Y S P T S P S	Y S P T S P S	Y A P T S P S
Y S P T S P S	Y S P T S P S	Y A P T S P S
Y S P T S P S	Y S P T S P S	Y A P T S P S
Y S P T S P S	Y S P T S P S	Y A P T S P S
Y S P T S P S	Y S P T S P S	Y A P T S P S
Y S P T S P S	Y S P T S P S	Y A P T S P S
Y S P S S P S	Y S P S S P S	Y A P S S P S
Y S P S S P S	Y S P S S P S	Y A P S S P S
Y S P S S P R	Y S P S S P R	Y A P S S P R
Y S P T S P T	Y S P T S P T	Y A P T S P T
Y S P T S P T	Y S P T S P T	Y A P T S P T
Y S P T S P T	Y S P T S P T	Y A P T S P T
Y S P T S P T	Y S P T S P T	Y A P T S P T
Y S P T S P S	Y S P T S P S	Y A P T S P S
Y E S G G G	Y E S G G G	Y E S G G G
Y S P S S P K	Y S P S S P K	Y A P S S P K
Y S P S S P T	Y S P S S P T	Y A P S S P T
Y S P T S P S	Y S P T S P S	Y A P T S P S
Y S P T S P Q	Y S P T S P Q	Y A P T S P Q
Y S P T S P Q	Y S P T S P Q	Y A P T S P Q
Y S P S S P T	Y S P S S P T	Y A P S S P T
Y T P S S P T	Y T P S S P T	Y A P S S P T
Y N P T S P R	Y N P T S P R	Y N P T S P R
G F S S P Q	G F S S P Q	G F S S P Q
Y S P T S P T	Y S P T S P T	Y A P T S P T
Y S P T S P S	Y S P T S P S	Y A P T S P S
Y T P S S P Q	Y T P S S P Q	Y A P S S P Q
Y S P T S P T	Y S P T S P T	Y A P T S P T
Y T P S P S E	Y T P S P S E	Y A P S P S E
Q P G T S A Q	Q P G T S A Q	Q P G T S A Q
Y S P T S P T	Y S P T S P T	Y A P T S P T
Y S P S S P T	Y S P S S P T	Y A P S S P T
Y S P A S P S	Y S P A S P S	Y A P A S P S
Y S P S S P T	Y S P S S P T	Y A P S S P T
Y D P N S	Y D P N S	Y D P N S

B.

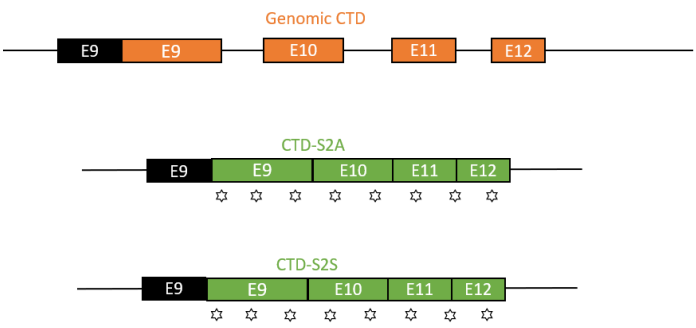


Figure 16: Sequence of the genomic CTD, CTD-S2A and CTD-S2S. (A). Protein sequence of the *wild-type* genomic CTD (adapted from Uniprot website), the CTD-S2S and the CTD-S2A. The protein sequence of the CTD-S2S is the same as the genomic CTD. However, the protein sequence of the CTD-S2A differs from the genomic CTD because of the presence of alanine instead of serine in position 2 (blue). The only exception is the CTD-Ser2 of the fourth repeat (red) which is conserved because it is not followed by a proline and thus can not be phosphorylated by CDK-12. **(B).** Cartoon of the DNA sequence coding for the genomic CTD, the CTD-S2S and the CTD-S2A. Contrary to the genomic CTD, the sequence coding for the CTD-S2A or the CTD-S2S lack introns and possess silent mutations (✡).

Results

1. Microinjection training

Transgenesis is performed in *C. elegans* by microinjecting in the gonad arm of a worm all the necessary components of CRISPR-mediated gene targeting (Figure 14). Because all germline nuclei share the same cytoplasm, the procedure targets a lot of futures oocytes and thus increases the heritability percentage by targeting the germline syncytium (Mello *et al*, 1991). Consequently, the parental injected worm is not mutated but its descendants (F1) are, most often on a single allele inherited from the oocyte. Because *C. elegans* is hermaphrodite, about 25% of the F2 should be homozygous for the mutation.

The injection procedure is not straightforward and requires practice. As I had no microinjection experience, I trained by injecting plasmids expressing *myo-2::GFP* and *myo-3::GFP*. *myo-2* codes for myosin 2 and is expressed in pharyngeal muscles while *myo-3* codes for myosin 3 and is expressed in body-wall muscles, vulval muscles and anal muscles (WormBase website). If the injection correctly targeted the germline rather than the intestine, which is not unusual, the progeny of a given injected worm will appear green under a fluorescence microscope (Figure 15). On average, 30% of the injected worms had F1 fluorescent descendants. Based on these results, I moved on to the generation of new mutants.

2. Does the mutation of the CTD-Ser2 recapitulate the inhibition of CDK-12?

2.1. Design of the CTD-S2A or CTD-S2S versions to integrate

In *C. elegans*, the CTD serine in position 2 is conserved in 33 out of the 42 repeats composing the CTD (Uniprot website) (Figure 16A). To test if the CTD-Ser2 is the genuine target of CDK-12 explaining the phenotype of the inactivation of the kinase, we ordered a DNA fragment whose sequence encodes for a version of the CTD where every CTD-Ser2 is replaced by an alanine (CTD-S2A) except the CTD-Ser2 found within the fourth uncanonical repeat, which is followed by a methionine rather than a proline (Figure 16A). Indeed CDKs are proline-directed serine/threonine kinases (Malumbres, 2014) and that specific serine is not a predicted substrate of CDK-12.

The sequence coding for this CTD-S2A was synthesized by Eurofins Genomics but, because of the repetitive nature of the CTD (even at the DNA level), the DNA sequence had to be adapted and codon-optimized from the genuine sequence (*wild-type* worms) to be synthesizable. In addition, in order to minimize its size, the 3 endogenous introns were removed (Figure 16B).

Considering that these constraints could have unanticipated effects, another sequence was synthesized by Eurofins Genomics, the CTD-S2S, following the same strategy but retaining serines in position 2 (Figure 16A and B). This sequence, therefore, serves as a control.

2.2. Construction of a strain expressing the CTD-S2A or CTD-S2S

To create the *C. elegans* strains where the CTD-S2A or CTD-S2S replaces the endogenous locus, we used the CRISPR/Cas9 method. There are two main ways to deliver the protein Cas9

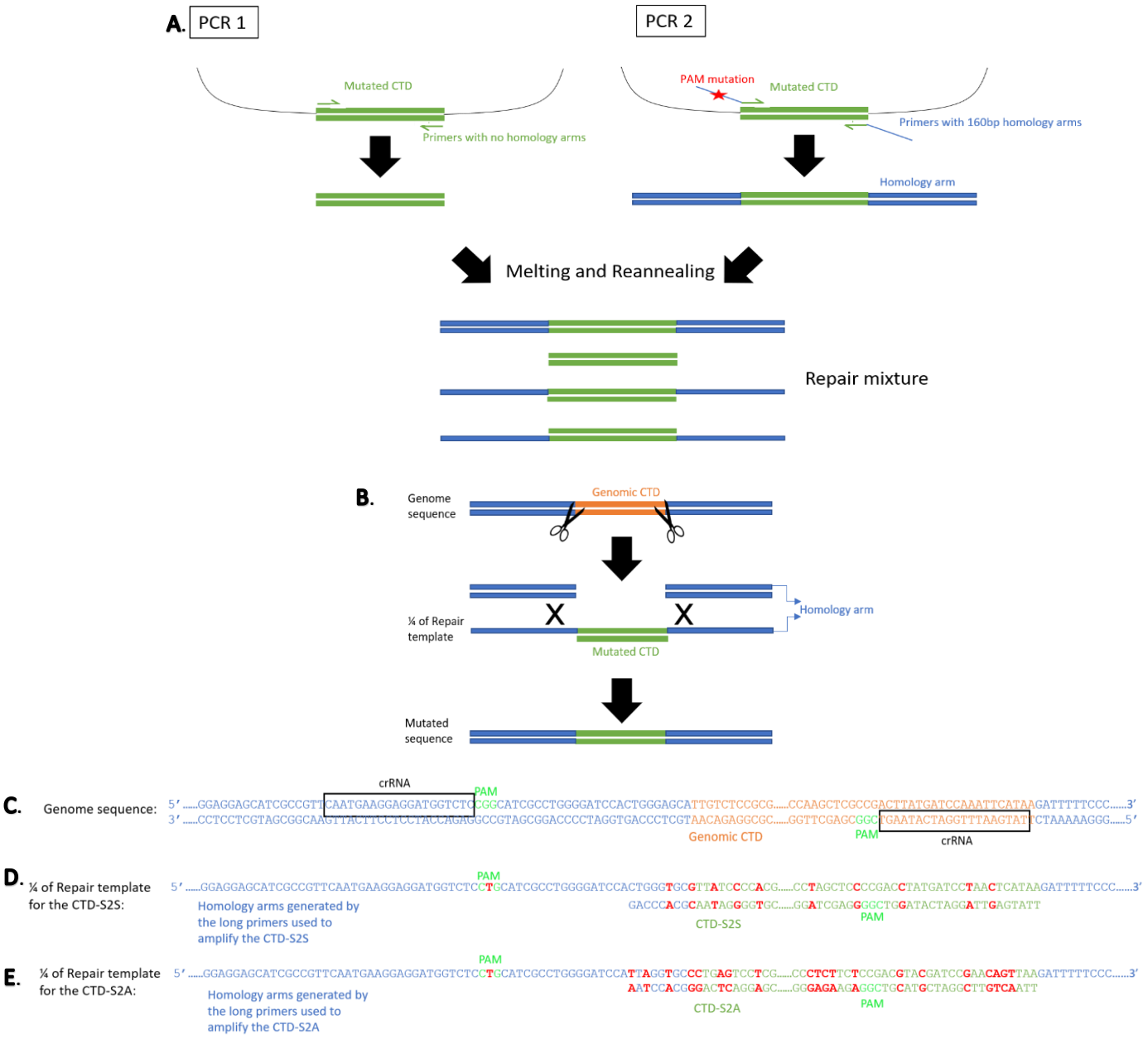


Figure 17: Construction of a strain expressing the CTD-S2A or CTD-S2S. (A). PCR reactions to obtain the repair template with the mutated CTD (either the CTD-S2A or the CTD-S2S). The first PCR is made with a set of primers with no homology arms to the genome (short primers) and the second is made with a set of primers with 160bp homology arms (long primers). The two PCR products are melted and reannealed to generate the repair mixture composed of four molecules of which two have single-strand overhangs. **(B).** Two crRNAs were designed to bring Cas9 on each side of the sequence coding for the CTD. Two double-strand DNA breaks are then generated, and a homologous recombination takes place with the injected donor template to integrate the sequence coding for the mutated CTD (CTD-S2A or CTD-S2S) instead of the one coding for the genomic CTD. **(C).** The first crRNA is designed to bind just before the beginning of the CTD and the second crRNA is designed to bind in the CTD, just before the stop codon of *ama-1*. **(D) (E).** The repair template created by the two different PCR reactions, contains the mutated CTD (green) and homology sequences to the genome (blue) generated by the long primers used in the second PCR reaction. To prevent Cas9 from cutting the modified genome in the beginning of the CTD, a mutation to mutate the PAM (red) was introduced in the forward long primer. To prevent Cas9 from cutting the modified genome in the end of the CTD, the silent mutations in the sequence coding for the mutated CTD and introduced by Eurofins Genomics (red) are enough.

and the sgRNA into the gonad, either as plasmid-borne or as a preassembled Cas9 ribonucleoprotein (RNP) composed of the purified Cas9 protein and the tracrRNA with crRNA synthesized *in vitro*. In addition, two types of repair templates can be used, a linear fragment or a circular repair plasmid (Dickinson and Goldstein, 2016).

In a first attempt to generate the mutant strains, we decided to use a linear repair template and two preassembled Cas9 RNPs, one with a crRNA targeting the upstream region of the DNA sequence encoding the CTD and one with a crRNA targeting the downstream region (Figure 17B and C).

The two crRNAs were designed using the IDT software, "Custom Alt-R CRISPR-Cas9 guide RNA". This software proposes a list of crRNA that can be used to target a desired region and gives for each crRNA the on-target score and the off-target score. The on-target score is an indication of the predicted activity of the crRNA and the off-target score is an indication of its specificity (Dickinson and Goldstein, 2016). Therefore, these scores have to be high (close to 100). For the crRNA targeting the beginning of the CTD, the software gave a lot of possibilities so we chose the one that is the closest from the first codon of the sequence encoding the CTD because the Cas9 mediated cut has to be realized as close as possible to the place where the modification is desired (Dickinson and Goldstein, 2016). This crRNA has an on-target score of 67, an off-target score of 99 and binds before the sequence encoding the CTD (Figure 17C). For the second crRNA, targeting the end of the CTD, the same approach was followed and we chose to work with a crRNA with an on-target score of 73, an off-target score of 83 and that binds within the sequence encoding the CTD just before the stop codon of *ama-1* (Figure 17C).

The linear repair template was produced by two independent, yet largely overlapping PCR reactions from the plasmid produced by Eurofins Genomics (Dokshin *et al*, 2018) (Figure 17A). The first PCR used primers with no homology arms to the *C. elegans* genome (short primers) and the second PCR used primers with 160bp homology to the genome locus (long primers). The two PCR products were melted and reannealed to obtain a mixture of four molecules of which two have single-strand overhangs. This mixture served as the repair template. We chose to work with single-stranded overhangs repair template because the work of Dokshin and co-workers supports that it works better than a template with fully double-stranded homology arms (Dokshin *et al*, 2018).

To make sure that Cas9 will not re-cut an already repaired locus, we introduced a silent mutation to change the PAM in the forward long primer used to amplify the repair template (Figure 17D and E). Consequently, after the cut and repair, the PAM will be mutated preventing Cas9 from recognizing its target in the 5' region of the CTD. Indeed, even with a crRNA perfectly matching with the genome, Cas9 cannot bind to its target region without the PAM motif (Diskinson and Golsdtein, 2016). On the reverse primer, no mutations were necessary because the PAM is located within a region that is already strongly altered in the synthetic, mutated sequence, which is enough to prevent Cas9 from re-cutting the modified genome (Figure 17D and E).

The mutant strain expressing the CTD-S2S was already obtained before I arrived in the lab, so we characterized it, but the CTD-S2A mutant had yet to be obtained. Therefore, we continued to inject *C. elegans*, using the co-CRISPR screening strategy mutating the *dpy-10* locus as reporter. Indeed, by replacing the CGT trinucleotide (coding for an arginine) by a TGC (coding for a cysteine) in the *dpy-10* locus (mutation *dpy-10* (*cn64*)), the worms become roller (if

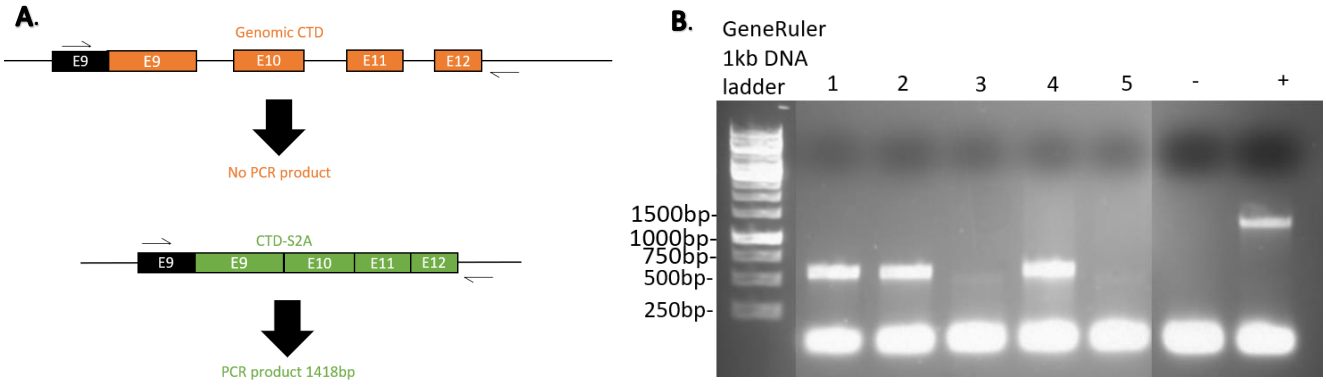


Figure 18: CTD-S2A screening. (A). CTD-S2A screening strategy. If the worm is *wild-type* (orange), no PCR product is amplified because the sequence is too long. If the worm has the CTD-S2A (green), then a 1418bp PCR product is obtained. **(B).** Agarose gel stained with ethidium bromide shows no CTD-S2A worms but worms expressing a deleted CTD (just over 500bp) (lanes 1, 2 and 4) and worms expressing the genomic CTD because no PCR product is obtained (lanes 3 and 5). The lane "-" is the negative control (F1 worm from a non-injected parental) and the lane "+" is the positive control (worm expressing the CTD-S2S because the sequence coding for the CTD-S2S or the CTD-S2A are the same length). The first single worm PCR protocol (see material and methods) was used.

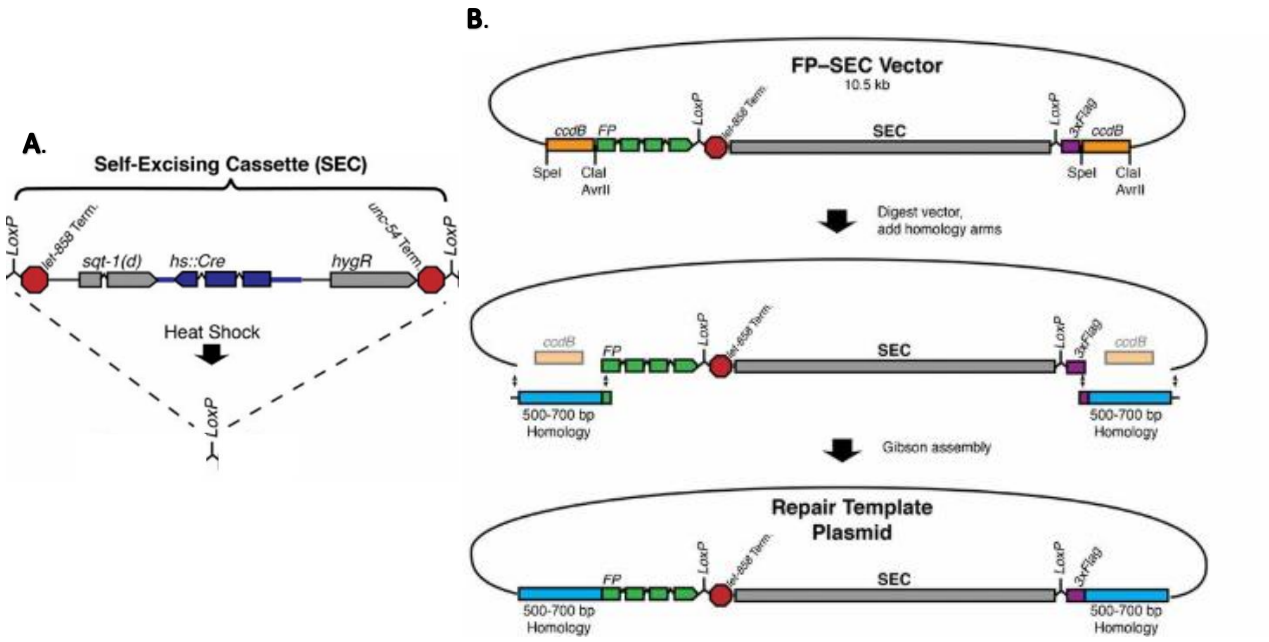


Figure 19: The self-excising cassette (SEC) (A). Composition of the SEC. This cassette is composed of a visible marker, *sqt-1(d)*, a hygromycin resistance gene (*hygR*) and an heat-shock inducible Cre recombinase (*hs::Cre)*. The SEC is flanked by two *LoxP* sites which allow its removal by the Cre enzyme after its activation by heat shock. After the excision of the cassette, one *LoxP* site remains within an intron (adapted from Dickinson *et al, 2015)*. **(B).** Use of the SEC to generate fluorescent and 3-FLAG-tagged protein. The plasmid that we can purchased contains the SEC with a fluorescent (FP) and a 3-FLAG tag flanked with *ccdB* fragments. It is possible to replace these *ccdB* fragments by homology arms to the *C. elegans* genome to use the plasmid in a CRISPR/Cas9 experiment in order to generate tagged proteins (Dickinson *et al*, 2015).

heterozygous) or dumpy (if homozygous) (Arribere *et al*, 2014). However, the CTD-S2A strain was not obtained. Indeed, roller worms were obtained and were thus tested to assess the replacement of the genomic CTD by the CTD-S2A by single worm PCR using a set of primers flanking the CTD (Figure 18A). If the worm was not mutant (he had the genomic CTD), no PCR product was obtained because the fragment is too long to be amplified by this PCR reaction. However, if the worm was mutant (the CTD-S2A replaced the genomic CTD) a PCR product of 1418bp should have been obtained because the DNA sequence coding for the CTD-S2A has no introns and is therefore shorter. All tested roller worms were negative for the mutation of the CTD but we obtained worms with a band just over 500bp (Figure 18B). We sequenced this PCR product and found that these worms were heterozygous for the expression of a deleted CTD (Figure S1). This means that the crRNAs can bind to the target DNA and that Cas9 can cut this DNA, however, the following repair process relies on non-homology end-joining with no homology-based integration of the repair template.

We thus decided to modify the method in order to obtain this strain. We thought of using a repair template with much longer homology arms to the *C. elegans* genome (2kb instead of 320bp) in order to increase the probability of homologous recombination. We also wanted to change the way to screen potential mutant worms in order to limit false-positive (worms with the *dpy-10* (*cn64*) mutation but not with the desired mutation). We, therefore, decided to use the screening method described by Dickinson and co-workers (Dickinson *et al*, 2015). In this paper, the authors introduce what they named a self-excising cassette (SEC) (Figure 19A). This SEC is composed of a hygromycin resistance gene, which is a selectable marker in *C. elegans*, and a Cre recombinase under the control of a heat-shock inducible promoter. In addition, the *sqt-1(d)* dominant allele (with a cysteine replacing an arginine) of the *sqt-1* gene, which generates a roller phenotype is also present (Dickinson *et al*, 2015). The *sqt-1(d)* allele only confers a robust roller phenotype but no a dumpy one because, by contrast to the co-CRISPR strategy where the endogenous gene is mutated, a third (mutated) copy of *sqt-1* is added while the endogenous alleles are intact. The *sqt-1(d)* allele, therefore, acts as a dominant-negative and this leads to a roller phenotype only (Dickinson *et al*, 2015).

In their paper, Dickinson and co-workers used this cassette to generate fluorescent and 3-FLAG tagged proteins. Indeed, they inject a mix of plasmids in *C. elegans*. The first plasmid of their mix contains the Cas9 gene, the second allows expression of the crRNA coupled with the tracrRNA and the third plasmid represents the repair template (Figure 19B) containing a fluorescent protein and a 3-FLAG merged to the SEC and flanked by 500-700bp homology arms (these homology arms are introduced by replacing the *ccdB* fragments in the starting plasmid by a Gibson Assembly). They screened mutant worms by first selecting hygromycin resistant worms and next by finding worms expressing the roller phenotype. Finally, once the wanted strain is obtained, the worms can be heat-shocked to induce the Cre recombinase, which will remove the SEC thanks to two *LoxP* sites flanking the cassette (Figure 19A). Therefore, a mutant that expresses a fluorescent and 3-FLAG tagged protein is obtained no longer containing the selection markers. The only disadvantage of this cassette is the residual presence of a *LoxP* site but that scare is located within a synthetic intron and will therefore not influence the produced protein sequence (Dickinson *et al*, 2015).

In order to create the two mutant strains expressing the CTD-S2A or CTD-S2S (which is repeated here as a control), we followed the protocol of Dickinson and co-workers (Dickinson

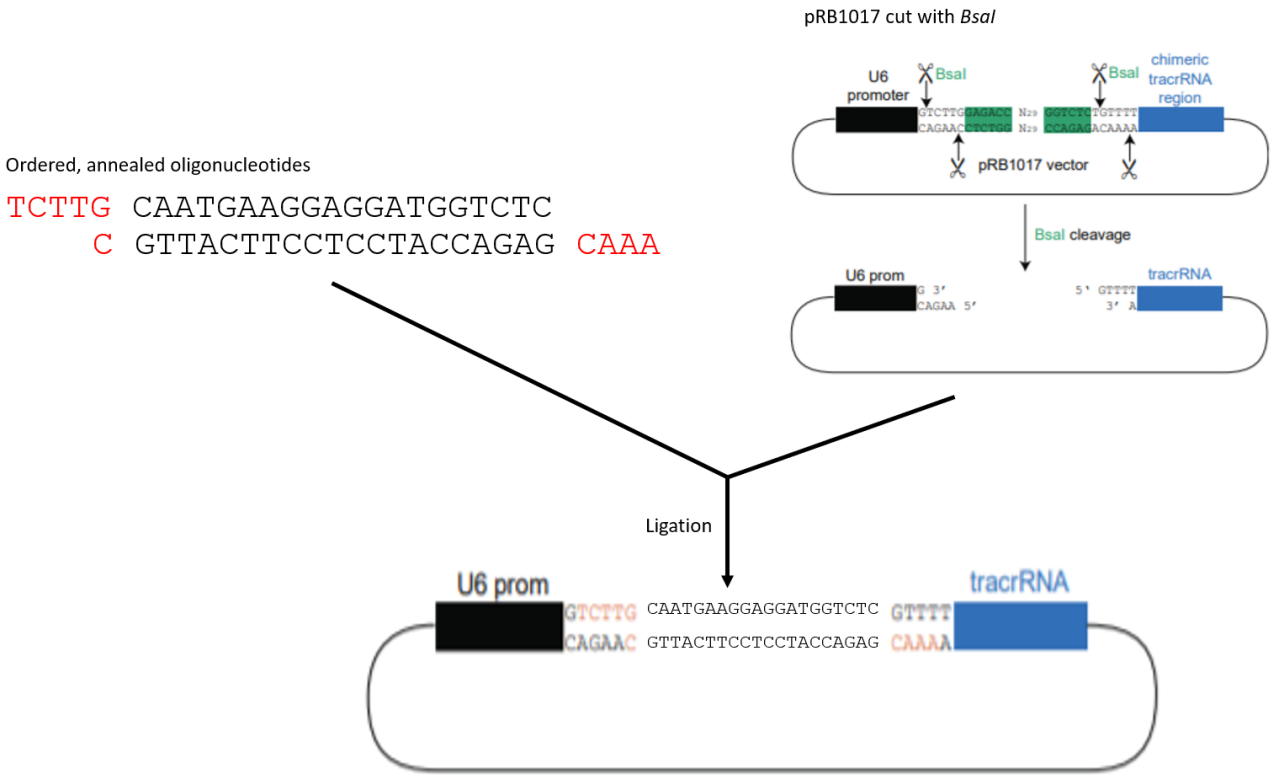


Figure 20: Construction of plasmids containing the crRNA and the tracrRNA. The plasmid pRB1017 is cut with *BsaI* to generate four nucleotides overhangs. Two oligonucleotides (the forward and the reverse sequence coding for the crRNA) are ordered and annealed. These oligonucleotides contain some more nucleotides (red) allowing their ligation in the cut plasmid. This figure show the example for the crRNA targeting the beginning of the CTD to generate a *C. elegans* mutant strain expressing the CTD-S2A or CTD-S2S (adapted from Arribere *et al*, 2014).

et al, 2015) and used a mix of plasmids instead of the preassembled Cas9 RNPs and the linear repair template.

The protocol requires the construction of two plasmids, designed to express one of the two crRNAs targeting the locus. We used the published pRB1017 plasmid (Arribere *et al*, 2014) (Figure 20). This plasmid already contains the DNA sequence corresponding to the tracrRNA preceded by 20 nucleotides flanked by sites for the restriction enzyme *Bsa*I. Consequently, to introduce the sequence corresponding to the crRNA, we cut the plasmid with *Bsa*I which generates 4 nucleotides overhangs. We next ordered two DNA oligonucleotides (the forward and reverse sequence corresponding to the crRNA) that we annealed. Each of these oligonucleotides contains in their 5' and 3' region, some nucleotides, complementary to the nucleotides overhangs generating by the cut of the plasmid by *Bsa*I, allowing their ligation in the plasmid. By using this technique, the crRNA is precisely fused to the tracrRNA to form a complete sgRNA after transcription.

The construction of the repair template was more challenging. Indeed, our goal is to replace the genomic sequence encoding the CTD by the CTD-S2S or CTD-S2A versions rather than tagging with a fluorescent protein or a 3-FLAG as described by Dickinson and co-workers (Dickinson *et al*, 2015). Consequently, a new more complex vector had to be constructed on the base of the Dickinson template (Figure 21).

We first cloned a PCR-amplified 5' CTD homology region (5' homo) of about 1kb into CTD-S2S and CTD-S2A containing pGEX vectors already available in the lab, using restriction and ligation (Figure 21A). The homology arm was amplified using GoTaq PCR with primers containing restriction sites for *Pvu*II and *Eco*RI at their 5' ends. The two pGEX were cut by *Mls*I (*Mls*I and *Pvu*II generate blunt extremities allowing the ligation of their product) and *Eco*RI allowing the incorporation of the 5' homology fragment in the plasmid.

The cloning of the two additional fragments needed, namely the SEC and the 3' CTD homology arm (3' homo) was performed by a Gibson Assembly strategy (Figure S2) that allows the one-step ordered joining of multiple fragments (Figure 21B). We thus made the PCR reactions to amplify the SEC and the 3' homology region with primers sharing overlapping sequences to the plasmid or between them to allow the Gibson Assembly, performed as described by the manufacturer (New England Biolabs). This way, the plasmid containing the 5' homology region, the CTD-S2A, the SEC and the 3' homology region was obtained. However, we failed to obtain the sister plasmid harboring the CTD-S2S, which was done through regular cloning replacing the CTD-S2A sequence by the CTD-S2S.

After the verification of these three plasmids, we injected them in *C. elegans* with a plasmid coding for Cas9. Because the injection of plasmids can lead to the formation of extrachromosomal arrays that can be transmitted to the progeny and expressed, we co-injected mCherry fluorescent plasmids (*myo-2::mCherry* and *myo-3::mCherry* to mark the muscles and *prab-3::mCherry* to mark the nervous system) (Dickinson *et al*, 2015; FrØkjaer-Jensen *et al*, 2008). These co-markers helped us to differentiate worms created by a proper genomic replacement event (roller, hygromycin resistant and not red when exposed to the appropriate wavelength) from worms carrying a combo of the plasmids as extrachromosomal arrays (roller, hygromycin resistant but also red) (Dickinson *et al*, 2015).

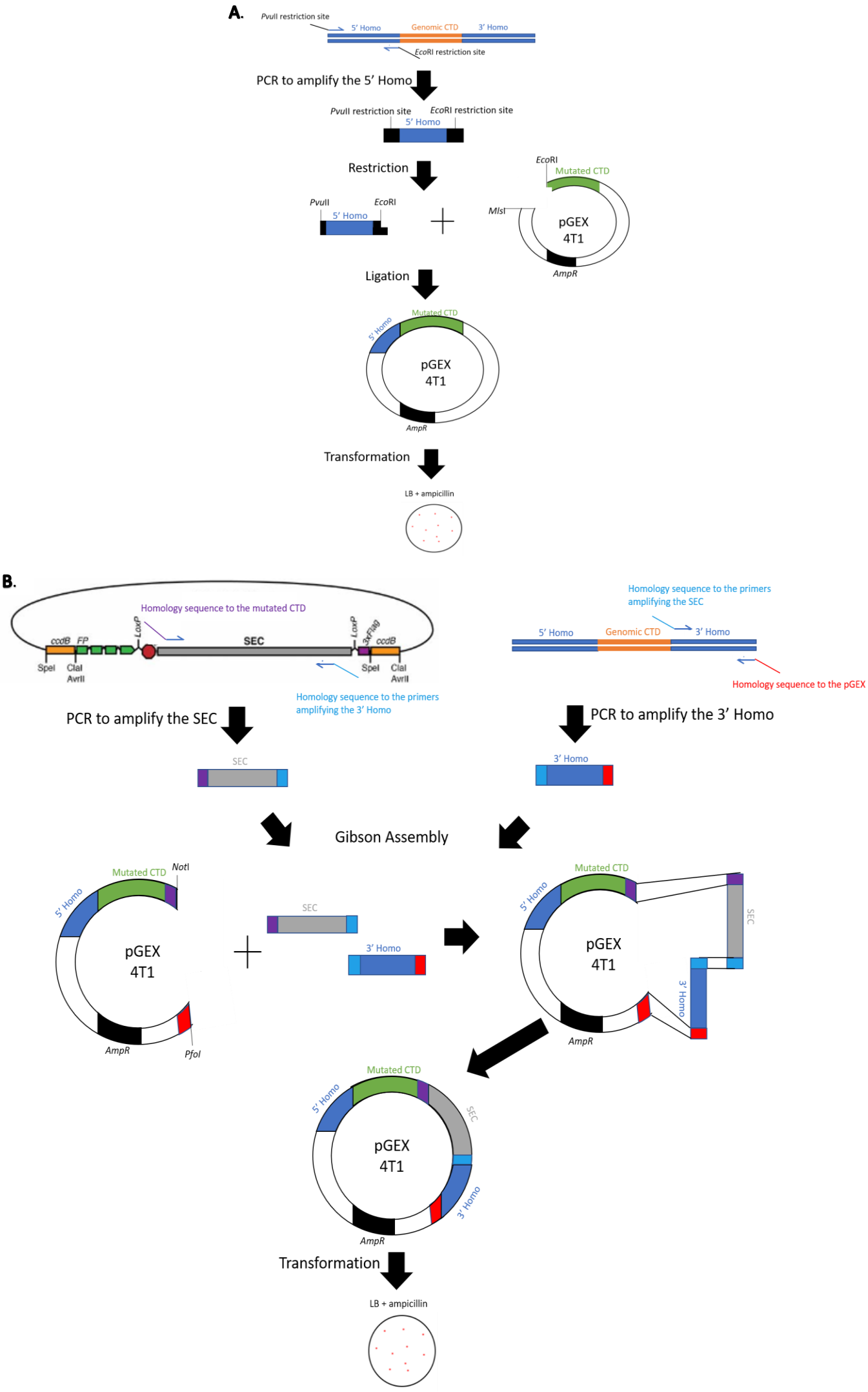


Figure 21: Construction of the repair template to create a strain expressing either the CTD-S2S or CTD-S2A (mutated CTD). (A). Integration of the 5' homo in the pGEX using restriction and ligation. The 5' CTD homology arm was amplified by PCR using a set of primers containing restriction site in their 5' part. This fragment was then cut by *Pvu*II and *Eco*RI to allow the ligation in the pGEX cut by *Mls*I and *Eco*RI . **(B).** Integration of the SEC (Dickinson *et al*, 2015) and the 3' homo in the pGEX already containing the 5' homo and the mutated CTD. The two inserts were amplified by PCR with primers sharing homology with the end of the CTD, cut with *Not*I (purple to insert the SEC), sharing homology between them (light blue to ligate the SEC with the 3' homo) and sharing homology with the other side of the plasmid, cut with *Pfo*I (red to insert the 3' homo). After the PCR reaction, a Gibson Assembly was performed following the protocol of New England Biolabs.

After the injection sessions, either for the CTD-S2A or CTD-S2S, we never obtained roller, hygromycin resistant worms. A possible explanation is that the PAM present in the 5' homology region in the repair template was not mutated due to the cloning strategy. Cas9 could thus cut the modified locus after repair as already discussed above. We thus performed an additional Q5 site-directed mutagenesis on the repair plasmids to mutate the PAM in the 5' homology region (Figure 22). We next injected the plasmid repair template containing the PAM mutation with the two plasmids coding for the sgRNAs and the one coding for Cas9 to first obtain the strain expressing the CTD-S2S. Indeed, we decided to try to generate this strain before the CTD-S2A because we knew that it is fully viable because we obtained it using our first strategy (two preassembled Cas9 ribonucleoproteins with a linear repair template). We injected 160 parental worms, however, we only obtained 21 roller, hygromycin resistant yet red worms suggesting that the plasmids formed extrachromosomal arrays and that the desired mutation was not incorporated in the genome. Indeed their progeny was not roller or hygromycin resistant due to the progressive loss of the extrachromosomal arrays.

We, therefore, reasoned that pursuing the same strategy was counter-productive and consulted again the most recent literature. Indeed, it was reported than the direct injection of a preassembled Cas9 ribonucleoprotein (as in our first design) is more efficient to edit the genome than the use of plasmids. (Paix *et al*, 2015; Dickinson and Goldstein, 2016, Au *et al*, 2019). It was also reported that for insertions larger than 1 or 2kb (we want to insert a fragment of 6500bp) or for insertions situated 30 nucleotides or more further than the Cas-9 mediated cut, plasmid repair templates with long homology arms are the most appropriate because they can accommodate large inserts (Paix *et al*, 2014; Dickinson and Goldstein, 2016; Paix *et al*, 2017). Consequently, to meet both arguments, we are currently trying to inject the two preassembled Cas9 RNPs along with the repair templates and the mCherry co-markers, therefore, combining the two strategies already tested.

2.3. Characterization of the CTD-S2S strain

The strain expressing the CTD-S2S was already obtained with the first strategy and available for further characterization.

The objective is to check that the presence of silent mutations and the removal of introns from the DNA sequence coding for the CTD do not impact the expression of AMA-1, the CTD-Ser2 phosphorylation, or do not result in phenotypes. To answer these questions, we first compared the brood size and the embryonic mortality between *wild-type* worms (expressing the genomic CTD) and worms expressing the CTD-S2S. Secondly, we compared the expression level of AMA-1 and its phosphorylation status between the two strains by Western blot and immunofluorescence.

We assessed the brood size and embryonic mortality because it is a metric that characterizes the fitness of a strain and because most developmental abnormalities affect the fertility of an individual. Consequently, we counted the larvae and the unhatched (dead) embryos (Figure 23A). No statistical difference was observed between the two strains.

The expression level of AMA-1 and CTD-Ser2P in each strain was analyzed by Western blot using antibodies anti-CTD and anti-CTD-Ser2P (Figure 23B and C). In this analysis, AMA-1 with a non-phosphorylated CTD migrates faster than AMA-1 with a phosphorylated

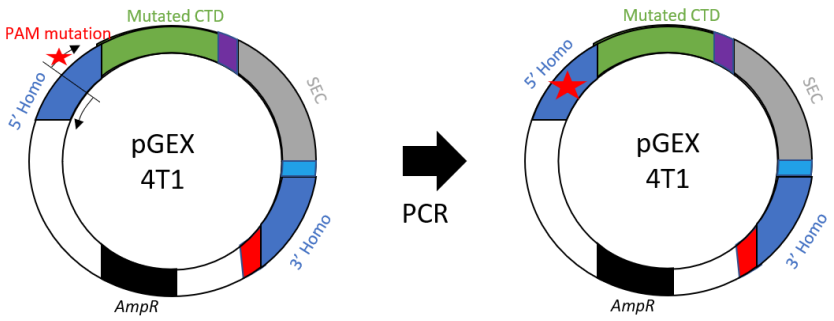


Figure 22: Q5 site-directed mutagenesis to mutate the PAM in the repair template to create a strain expressing either the CTD-S2A or CTD-S2S. By using a forward primer containing the mutation of the PAM and a reverse primer annealing back-to back to the forward primer, the entire plasmid is amplified by a PCR reaction and the desired mutation is introduced in the plasmid.

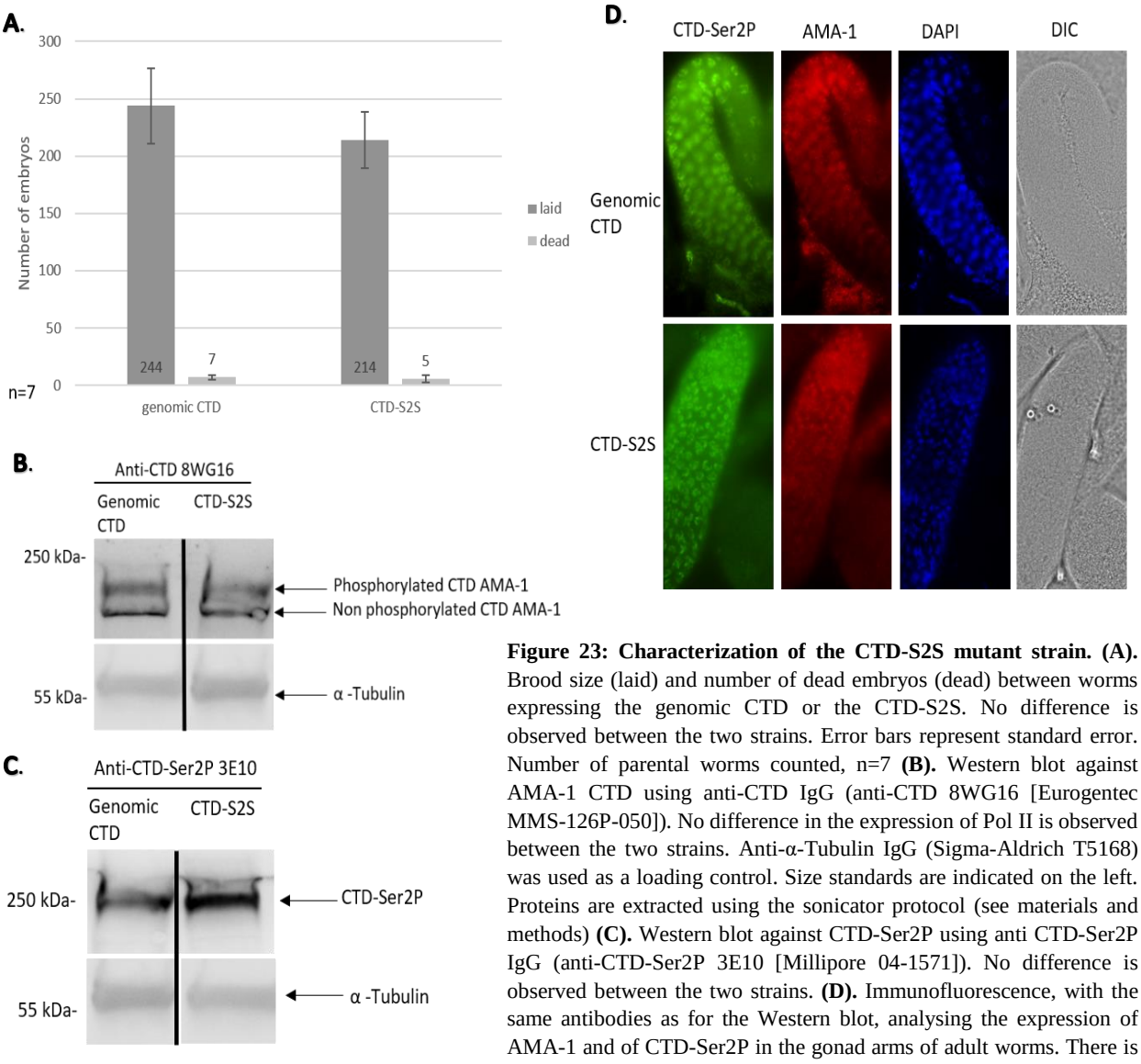


Figure 23: Characterization of the CTD-S2S mutant strain. (A). Brood size (laid) and number of dead embryos (dead) between worms expressing the genomic CTD or the CTD-S2S. No difference is observed between the two strains. Error bars represent standard error. Number of parental worms counted, n=7 **(B).** Western blot against AMA-1 CTD using anti-CTD IgG (anti-CTD 8WG16 [Eurogentec MMS-126P-050]). No difference in the expression of Pol II is observed between the two strains. Anti-α-Tubulin IgG (Sigma-Aldrich T5168) was used as a loading control. Size standards are indicated on the left. Proteins are extracted using the sonicator protocol (see materials and methods) **(C).** Western blot against CTD-Ser2P using anti CTD-Ser2P IgG (anti-CTD-Ser2P 3E10 [Millipore 04-1571]). No difference is observed between the two strains. **(D).** Immunofluorescence, with the same antibodies as for the Western blot, analysing the expression of AMA-1 and of CTD-Ser2P in the gonad arms of adult worms. There is no difference between the two strains. DAPI marks the nuclei. DIC stand for Differential Interference Contrast.

CTD and that is why the non-phosphorylated CTD RNA polymerase appears lower on the gel. The higher band corresponds to the phosphorylated CTD polymerase and can migrate at different molecular weights depending on the CTD phosphorylation status (Greenleaf, 1992). Thus, it is not straightforward to analyze the CTD phosphorylation but figures 23B and C indicate that there is no obvious difference in the expression of AMA-1 or on CTD-Ser2P between the two strains.

We also analyzed the expression level of AMA-1 and of CTD-Ser2P in each strain by immunofluorescence using the same antibodies as for the Western blot (Figure 23D). The experience was performed on adult worms, but because of their cuticle, the antibodies have difficulties to reach their target. The solution is to cut the worms in half allowing the gonad arms to come out of the cuticle. Consequently, we can only observe the worms' gonad arms in this experiment. Figure 23D shows that there is no difference either on the expression of the polymerase or on the CTD-Ser2 phosphorylation between the non-mutated worms and the mutated one.

In conclusion, the results of these experiments support that the silent mutations introduced into the DNA sequence coding for the CTD have no detectable impact.

2.4. A technical control for the CRISPR-dependent gene targeting

At the same time as performing injections to obtain a mutant strain expressing the CTD-S2A or CTD-S2S, we injected a mix of plasmids to obtain a strain expressing the protein SLD-3 tagged with an MNase and a 3-FLAG. Sld3 is a protein involved in the replication cycle. Indeed, it is the substrate of Cyclin-Dependent Kinases involved in the S phase of the cell cycle. After the phosphorylation of Sld3 by CDKs, Dpb11 can interact with it and the complex will promote the assembly of the replication fork helicase (Dhingra *et al,* 2015; Zegerman and Diffley, 2007). The MNase (Micrococcal Nuclease) is an endonuclease that can generate double-strand DNA breaks under high calcium concentration. It means that under physiological Ca^{2+} concentration, the fusion protein SLD-3-MNase will be recruited to the DNA with Dpb11 and that the MNase will remain inactive. However, when genomic DNA will be exposed to high calcium concentration, the complex will be recruited to the DNA and the endonuclease will generate double-strand DNA breaks where SLD-3 is located. The cleaved chromatin can then be analyzed for example by a ligation-mediated PCR and sequencing. This technique allows for studying protein-chromatin interactions (Kubben *et al,* 2010).

This mutant strain has no biological interest for us but has technical interest as it has already been obtained in the laboratory of a collaborator, Florian Steiner, in Geneva, using a mix of plasmids; a plasmid coding for Cas9, a plasmid containing a crRNA coupled with the tracrRNA, a plasmid repair template containing the SEC and the mCherry co-markers. Consequently, the goal of recreating this strain in our laboratory was to practice the use of the self-excising cassette and the mCherry co-markers as a screening strategy and to test the efficacy of the heat-shock.

Florian Steiner sent us all the plasmids we needed to obtain the strain so we injected them into *C. elegans*. Briefly, the plasmid containing the crRNA coupled with the tracrRNA was also produced by cutting the pRB1017 plasmid by *Bsa*I and by replacing the 20 random nucleotides before the tracrRNA by the nucleotides coding for the crRNA. This crRNA is designed to bind near the start codon of *sld-3* (Figure 24) allowing the incorporation of the sequence coding for

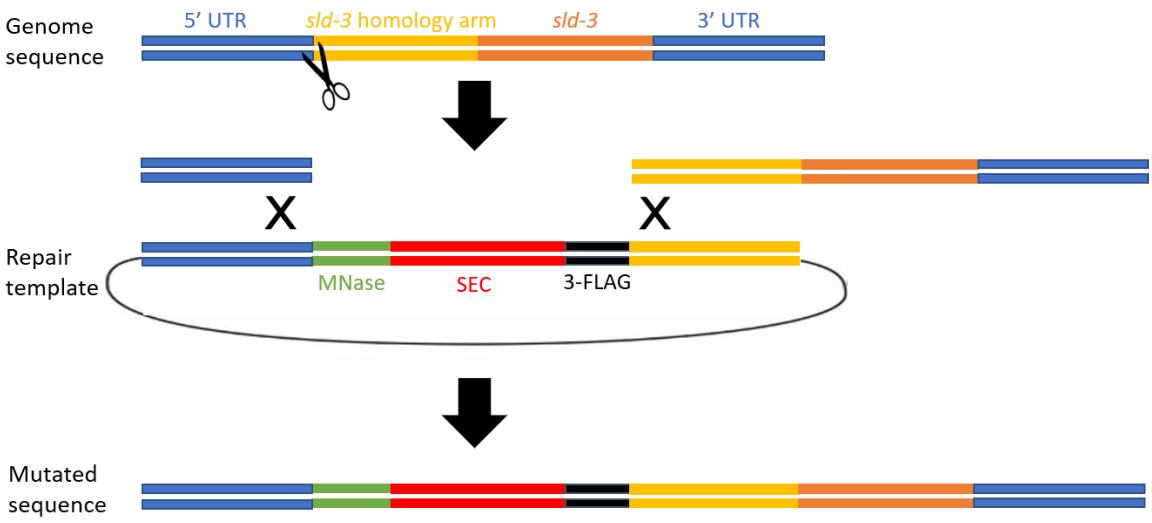


Figure 24: Construction of a strain expressing a tagged SLD-3. To obtain a strain expressing a tagged SLD-3, the crRNA was designed to bring Cas9 near the start codon of *sld-3*. Cas9 thus induces a double-strand DNA break which is repaired by homologous recombination with the plasmid repair template containing two homology arms to the genome, the SEC and the desired tags (MNase and 3-FLAG).

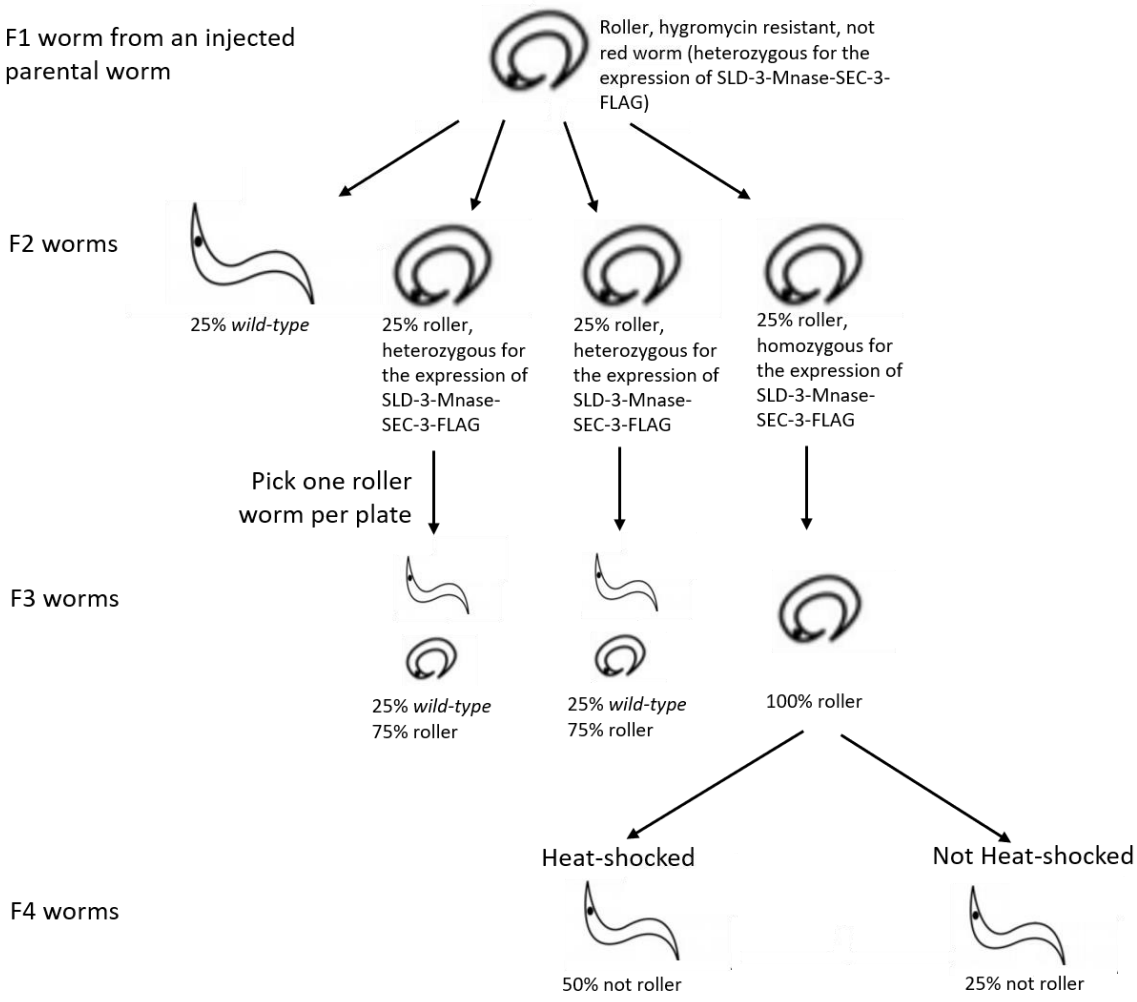


Figure 25: Selection of mutant worms expressing SLD-3 tagged with a MNase and a 3-FLAG. After injection of parental worms, we obtained one roller, hygromycin resistant and not red F1 worm meaning that the MNase with the SEC and the 3-FLAG was incorporated into its genome. This worm was heterozygous for the incorporation of the desired modification as its progeny (F2) contained non roller worms. By picking one roller F2 worm per plate, we obtained plates where 100% of the progeny (F3) were roller meaning that they were homozygous for the expression of the MNase, the SEC and the 3-FLAG. These homozygous worms were heat-shocked to induce the Cre recombinase. 50% of the heat-shocked worms were not anymore roller. However, 25% of not heat-shocked worms were also not anymore roller.

the MNase and the 3-FLAG in the genome (actually, in the Steiner lab, even if one is theoretically sufficient, they always use two plasmids coding for two crRNAs to optimize the targeting of the desired gene). The plasmid repair template is produced from the plasmid, described by Dickinson and co-workers, containing a fluorescent protein, a 3-FLAG, the SEC and *ccdB* fragments (Dickinson *et al*, 2015) by replacing the two *ccdB* fragments by homology arms to the genome and by replacing the fluorescent protein by the MNase but conserving the 3-FLAG.

After the injection sessions, we obtained a lot of roller and hygromycin resistants F1 worms due to the presence of the SEC. However, the majority expressed the red co-markers meaning that the injection had been correctly performed but that the plasmids were maintained as extrachromosomal arrays. The descendants of these F1 worms were no longer roller or hygromycin resistant because of the progressive loss of the extrachromosomal arrays. But we obtained one worm (1 to approximately 150 roller and hygromycin resistant worms) that was not red suggesting proper incorporation into the genome. Accordingly, the descendants of this worm were also roller and hygromycin resistant. We next isolated one roller per plate and screened for plates where 100% of the progeny were roller meaning that they were homozygous for the insertion (Figure 25). We performed a heat-shock on these homozygous worms to induce the Cre recombinase in order to remove the SEC. We did this by incubating 24 L1/L2 worms 3 hours and 30 minutes at 34°C and then return them at 20°C until they become adults. 62.5% of the worms submitted to the heat-shock survived and among this 62.5 %, 50% were not roller anymore. As a control, we incubated 24 L1/L2 worms only at 20°C, they all survived but surprisingly, 25% were not roller anymore. We next noticed that the roller phenotype disappeared in approximately 20% of the progeny of a roller worm supposed to be homozygous for the expression of the SEC. Consequently, we presumed that the SEC is able to excise itself spontaneously without heat-shocking. To check this hypothesis, we decided to perform three different PCR reactions (Figure 26A) using three different forward primers but the same reverse primer (number 3194).

The first PCR (3191-3194) is expected to generate a band of 1448bp only if the worm expresses the 3-FLAG (and thus most likely the MNase.) The second PCR (3192-3194) is expected to generate a 1643bp band if the worm is *wild-type*, a 2264bp band if the worm expresses the MNase and the 3-FLAG without the SEC and no PCR product if the worm expresses the MNase, the SEC and the 3-FLAG because the fragment is too long to be amplified by this reaction. The third PCR (3141-3194) is expected to generate a 2062bp band if the worm expresses the SEC. As the three forward primers are able to hybridize within the repair template, we needed to distinguish a PCR product corresponding to the integration into the genome from an extrachromosomal array, and the reverse primer was therefore designed in *sld-3* downstream of the targeted homology region present in the repair template. We first performed these reactions on a *wild-type* worm as a control (Figure 26B). The first PCR gives no PCR products, the second shows a 1643bp but the third gives aspecific bands. We next performed these PCR on worms potentially homozygous for the expression of SLD-3-MNase-3-FLAG and which are roller and hygromycin resistant (Figure 26B). The first PCR gives a 1448bp band meaning that the 3-FLAG and the MNase are expressed, the third PCR gives the same aspecific bands than in the control but also gives a 2062bp band meaning that the worms express the SEC. However, if these worms were homozygous for the expression of the MNase, the SEC and the 3-FLAG, the second PCR would not have given a product. However, we obtained a 2264bp.

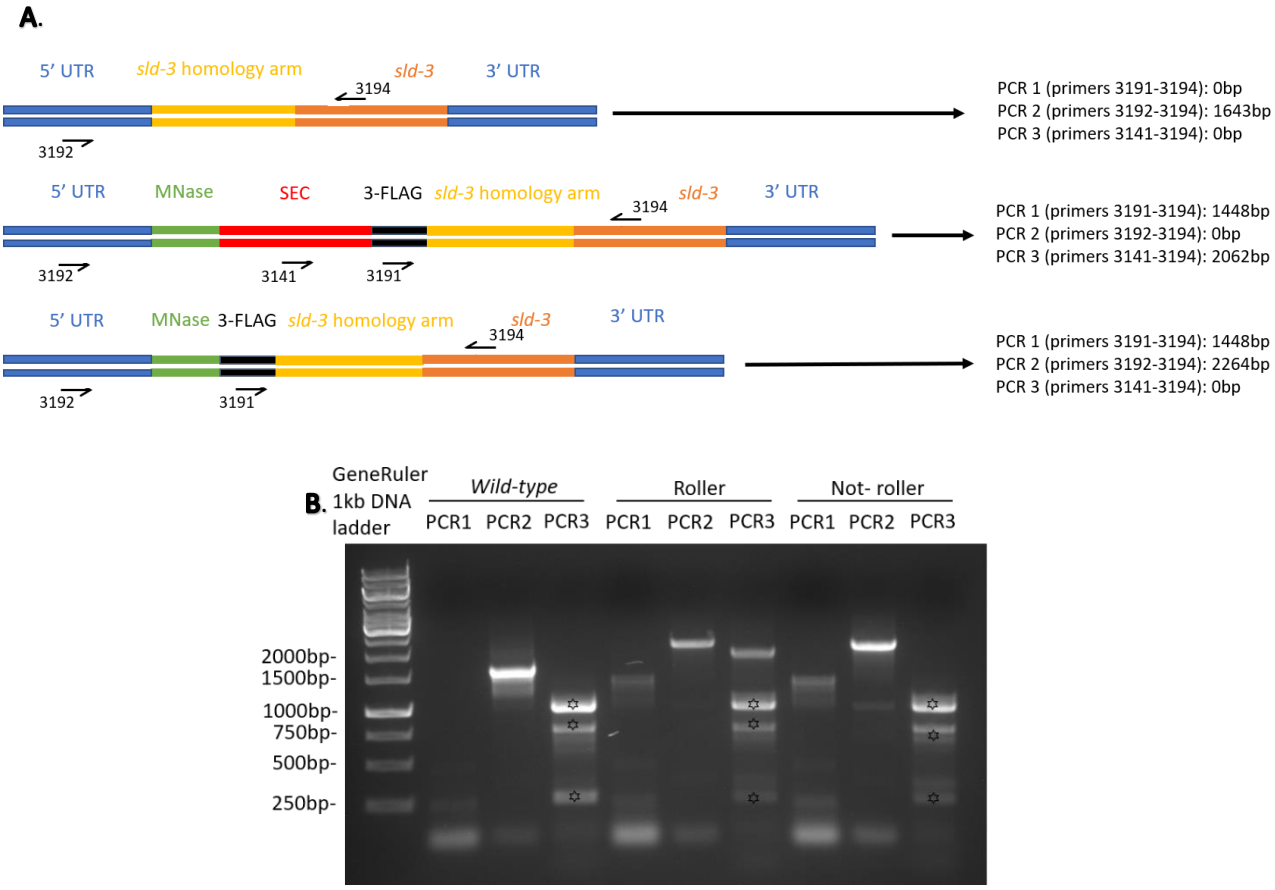


Figure 26: PCR reactions to analyse the expression of the MNase, the SEC and the 3-FLAG in worms supposed to be homozygous for their expression. (A). Cartoon of the three possible alleles that can be expressed by a worm with the length of the PCR product given by three different PCR reactions, using the same reverse primer (number 3194), hybridizing in *sld-3*. The first PCR uses the primer number 3191, hybridizing in the 3-FLAG, the second PCR uses the primer 3192, hybridizing in 5' UTR of *sld-3* and the third PCR uses the primer 3141, hybridizing in the SEC. **(B).** Agarose gel stained with ethidium bromide indeed shows that on a *wild-type* worm, the PCR1 gives nothing, the PCR2 gives a 1643bp band but the third PCR gives aspecific bands (✡). On a worm supposed to be homozygous for the expression of the SEC, the MNase and the 3-FLAG and that are roller and hygromycin resistant (Roller), the PCRs indicated that they are actually homozygous for the expression of the MNase and the 3-FLAG but they express the SEC only on one of the two alleles. On a worm that has lost the roller and hygromycin resistant phenotype (Not-roller), the PCRs indicated that they are homozygous for the expression of the MNase and the 3-FLAG but they do not more express the SEC. The second single worm PCR protocol (see material and methods) was used.

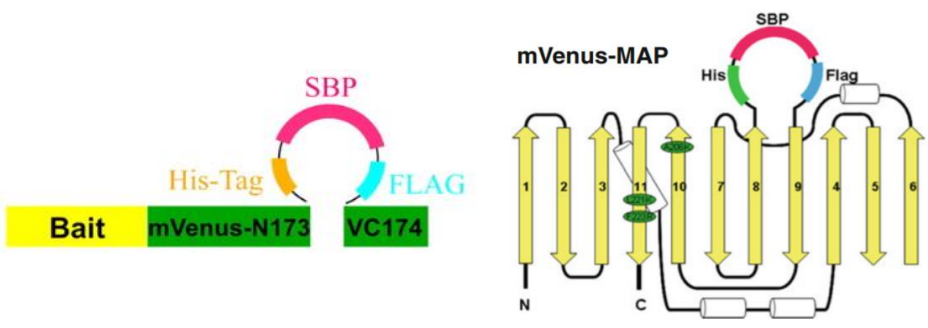


Figure 27: The MAP tag. The MAP tag is composed of the fluorescent protein mVenus with a His_8-tag, a Streptavidin Binding Peptide (SBP) and a FLAG tag inserted between β8-β9 of mVenus (Ma *et al*, 2012; Ma *et al*, 2014).

Taken together, these results suggest that the tested roller worms are actually homozygous for the expression of the MNase and of the 3-FLAG but heterozygous for the expression of the SEC, which supports that the cassette has spontaneously excised due to the leakiness of the heat-shock promoter. We also performed these PCR reactions on worms supposed to be homozygous for the expression of the tagged SLD-3 but which have lost the roller phenotype without heat-shocking (Figure 26B). These worms express the 3-FLAG and thus most likely the MNase because the first PCR gives a 1448bp band and that the second gives a 2264bp band but they do not express the SEC because the third PCR did not give the 2062bp band corresponding to the expression of the cassette. This also confirms the hypothesis that the SEC is able to excise without heat-shock at 34°C, which was confirmed by Florian Steiner (personal communication).

These results suggest that despite the self-excision of the cassette without heat-shocking, the SEC accompanied by the co-markers (mCherry plasmids) is an appropriate way to screen for potential mutant worms. Indeed the only roller, hygromycin resistant and negative for red fluorescence worm we obtained was properly mutated at the *sld-3* locus. We, therefore, should be able to obtain the CTD-S2A or CTD-S2S strains by using this screening method and by screening roller, hygromycin resistant and red fluorescence negative F1 worms after injection sessions.

3. Study of the regulation of CDK-12

3.1. Construction of a strain expressing a tagged CDK-12

To study how CDK-12 is regulated *in vivo*, we want to tag the kinase in order to identify putative physical partners and to study its phosphorylation status. To generate this strain, we first decided to use a tag that allows specific isolation of the target protein with the minimum contaminants to avoid false-positive. We first thought about using the TAP-tag (Li, 2010). However, the disadvantage of the TAP-tag is the presence of a calmodulin-binding peptide because of the existence of calmodulin-binding proteins in *C. elegans* (Shen *et al*, 2008). These could also bind to the calmodulin resin during the purification of the desired complexes and thus give false-positive results (Li, 2010). We thus looked for a derivative of the TAP-tag retaining its advantages in the absence of a calmodulin-binding peptide. Several options are available and we chose to use the MAP (Multifunctional tandem Affinity Purification) tag (Ma *et al*, 2012) because firstly, the purification of the target complex is reported to be very efficient, secondly, this tag works for the isolation of proteins in *C. elegans* and thirdly, this tag offers the possibility to study the localization of a protein thanks to the presence of a fluorescent protein (Ma *et al*, 2012). Indeed, the MAP tag is composed the mVenus with a His_8-Tag, a Streptavidin-Binding Peptide (SBP) and a FLAG tag inserted into the surface loop between the beta barrels β8-β9 (Figure 27) (Ma *et al*, 2012).

To incorporate the MAP tag, we designed a crRNA targeting the end of the *cdk-12* ORF in order to fuse the tag to the C-terminal end of the kinase (Figure 28A). We used the IDT software. We could either chose a crRNA binding in the coding sequence of *cdk-12*, just before the stop codon, or in the 3' UTR of *cdk-12*, just after the stop codon. We chose to work with the first proposition using a crRNA with an on-target score of 58 and an off-target score of 97 (Figure 28B) because the crRNAs hybridizing in the 3' UTR have bad specificity (off-target score of 0).

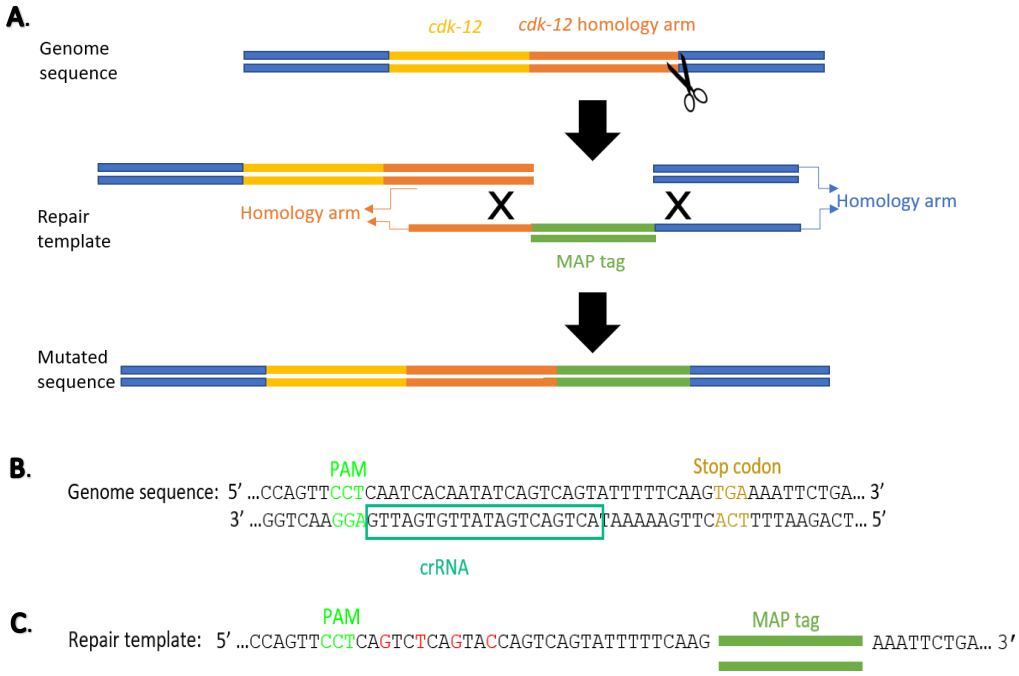


Figure 28: Construction of a strain expressing CDK-12 tagged with the MAP tag. (A). A crRNA was designed to bring Cas9 near the stop codon of *cdk-12*. A double-strand break is thus generated to open the DNA and allows the homologous recombination with the injected DNA in order to incorporate the MAP tag. **(B).** The crRNA is designed to bind just before the stop codon of *cdk-12* to put the tag in the C-terminal part of the protein. **(C).** The MAP tag is flanked by homology regions to the break but silent mutations (red) are inserted in one homology arm (the one that recognizes the crRNA) to prevent the crRNA from binding and thus Cas9 to cut the modified genome.

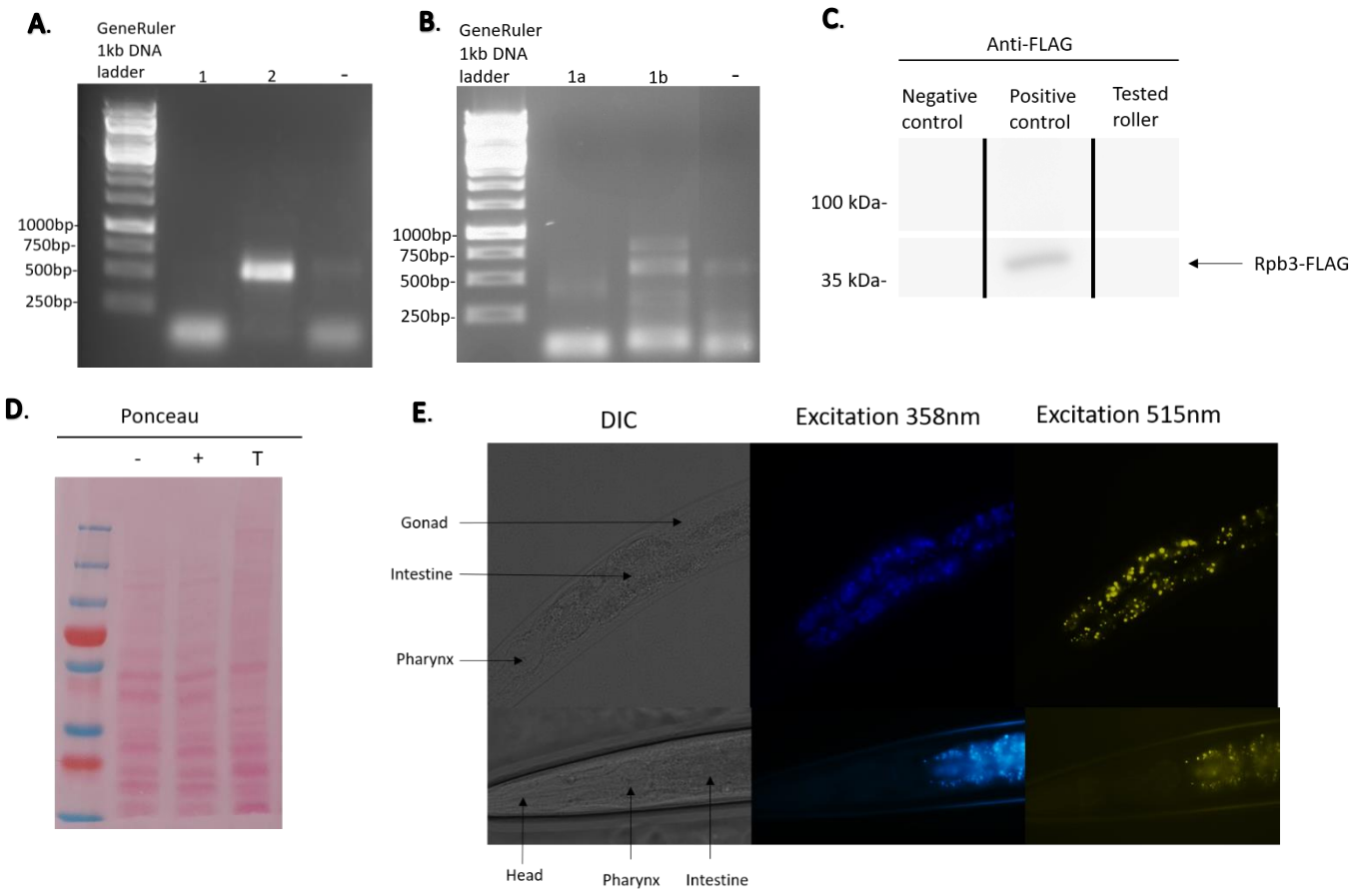


Figure 29: CDK-12-MAP tag screening. (A). An agarose gel stained with ethidium bromide shows a possible mutant worm (lane 2) that had the 541bp band, and a non-mutant worm (lane 1) with no PCR product. The lane "-" is a F1 worm from a non-injected parental worm as negative control. The first single worm PCR (see material and methods) was used **(B).** The agarose gel shows a possible mutant worm (lane 1a and 1b) that had the expected 541bp band with the first designed set of primers (lane 1a) and the expected 727bp band with the second designed set of primers (lane 1b) despite the presence of aspecific bands. Lane 2 is the negative control. **(C).** Western blot against the FLAG using anti-FLAG IgG (Sigma-Aldrich F3165). No roller worms (tested roller) showed a band for the expression of CDK-12 tagged with the MAP tag (108,6kDa). The negative control is an *S. pombe* strain which does not express the FLAG tag. The positive control is an *S. pombe* strain expressing Rpb3-FLAG (35kDa). Size standards are indicated on the left. Proteins were extracted using the NaOH extraction (see material and methods). **(D).** Ponceau staining served as a control to see if there were proteins on the membrane. The lane "-" is the negative control, the lane "+" is the positive control and the lane "T" is the tested roller. **(E).** Exposition of roller worms to 515nm (excitation wavelength of the mVenus). No roller worms expressed the mVenus because what we can see on the figure is the autofluorescence of the intestine. The way to check that the observed fluorescence was indeed due to the intestine is to exposed worms to 358nm (excitation wavelength of the DAPI). In addition, the pharynx, gonad or head do not appear fluorescent when exposed to 515nm. DIC stand for Differential Interference Contrast.

The repair template containing the MAP tag flanked by homology arms to the *C. elegans* genome was made following Dokshin and co-workers protocol (Dokshin *et al*, 2018); by two different PCR reactions from a plasmid expressing the MAP tag (bought from Addgene). The first PCR was made using primers with no homology arms to *C. elegans* genome (short primers) and the second using primers with 120bp homology arms (long primers). The two PCR products were melted and reannealed to generate single-stranded overhangs repair template. This template should allow the homologous recombination with the genome to incorporate the MAP tag. In addition, we introduced silent mutations in the forward long primers used to amplify the MAP tag to prevent Cas9 from cutting the genome after it has been modified (Figure 28C).

We injected a preassembled Cas9 RNP (the protein Cas9 with the crRNA and the tracrRNA *in vitro* synthesized) and the linear repair template into the gonad arm of *C. elegans* to obtain a mutant worm (Figure 28A). To screen for potential mutants, we used the co-CRISPR technique with the co-injection marker, *dpy-10 (cn64)* to generate roller worms. We obtained roller worms that we screened by single worm PCR to check for the incorporation of the tag into the genome thanks to one of the two primers binding in the tag. Theoretically, the PCR screening strategy should generate a 541bp PCR product only if the worms were at least heterozygous. We obtained worms with this 541bp (Figure 29A) so we sequenced this PCR product but it surprisingly did not correspond to the MAP tag (Figure S3). It means that the set of primers we used can amplify a sequence of 500bp somewhere else into the genome. This is unexpected as we looked for off-target binding of the primers by blasting them against the entire *C. elegans* genome. To avoid sequencing all false-positive 541bp bands, we designed another primer that should give a 727bp PCR product if the worms are mutant. We obtained roller worms showing this 727bp (Figure 29B) band but with some aspecific bands also present in the negative control worm (not injected). We unsuccessfully tried to improve our PCR protocol in order to avoid these aspecific bands. We thus decided to use another method to check if the roller worms expressed the MAP tag. We performed Western blots on roller worms using an antibody anti-FLAG and analyzed the roller worms under a fluorescent microscope to check if they expressed the mVenus. The results were negative in both experiments. Indeed on Western blots, using an *S. pombe* strain expressing a FLAG-tagged protein as a positive control, no roller worms expressed the FLAG tag (Figure 29C) (to be sure that there are proteins on the membrane, a ponceau staining was performed (Figure 29D)). In addition, no roller worms expressed the mVenus when exposed to the proper wavelength (515nm) (Figure 29E). Indeed, what we can observe in figure 29E is the intestine because it contains autofluorescent “lipofuscin granules” (Clokey and Jacobson, 1986). A way to check that indeed, the observed fluorescence was due to the intestine autofluorescence, is to expose worms to 358nm, the excitation wavelength of the DAPI. As nuclei were not marked by DAPI, only the “lipofuscin granules” of the intestine appeared fluorescent at 358nm. As the two signals (the one obtained at 515nm and at 358nm) co-localise, the signal observed by exposing worms to 515nm is likely due to the intestine autofluorescence rather than the expression of the MAP tag. Moreover, the pharynx, gonad or head do not appear fluorescent when exposed to 515nm. In conclusion, the nonspecific PCR reaction, the Western blots and the observation of roller worms under the fluorescence microscope show that after the injection of *C. elegans,* we obtained a lot of roller worms but no transgenic lines expressing a MAP tag version of CDK-12.

We consequently decided to change the method and to use a mix of plasmids with a circular repair template in order to use longer homology arms to the *C. elegans* genome and the SEC as a screening strategy.

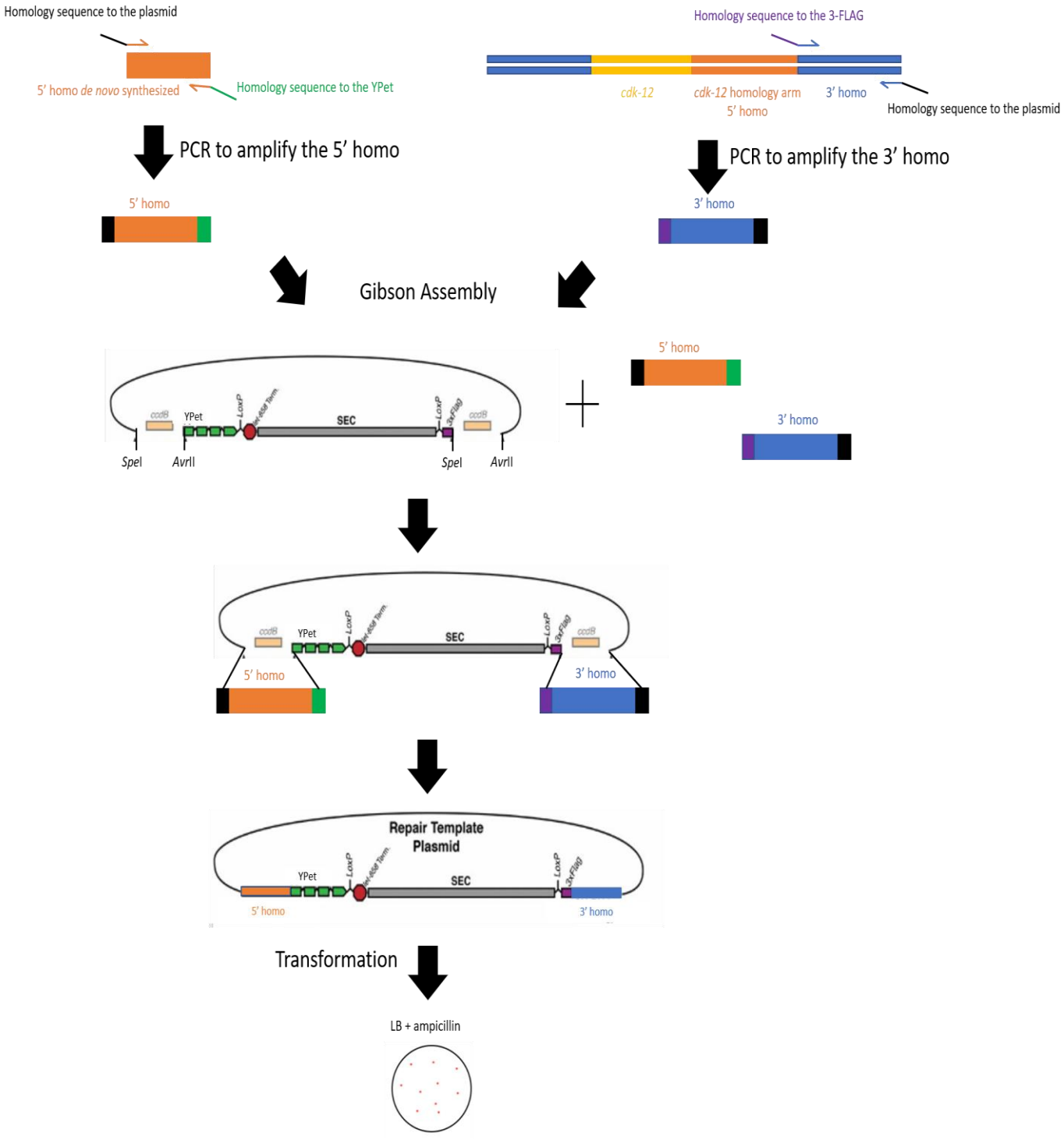


Figure 30: Construction of the repair template to obtain a strain expressing a CDK-12-YPet-3-FLAG. The 5' homology arm was synthesized *de novo* in order to incorporate silent mutations to prevent Cas9 from cutting a modified genome. It was then amplified by a PCR reaction with primers containing overlapping sequence with the plasmid pDD283 (Dickinson *et al*, 2015). The 3' homology arm was directly amplified from the genome by a PCR reaction with primers also containing overlapping sequence with the plasmid. The vector will then be cut by *Spe*I and *Avr*II to remove the *ccdB* fragments and a Gibson Assembly will be performed.

For the plasmid expressing the crRNA, we again used the pRB1017 digested with *Bsa*I to introduce the crRNA. Normally, one crRNA is sufficient because the genome has to be cut once. However, we decided to use two independent crRNAs (like Florian Steiner to obtain the strain expressing SLD-3-MNase-3-FLAG) to optimize the targeting of the gene of interest. We thus chose to work with a second crRNA hybridizing 3 nucleotides upstream of the first one (on-target score of 67 and off-target score of 96). We, therefore, constructed two plasmids each containing one crRNA coupled with the tracrRNA.

Concerning the repair template, we decided to change the tag and to use a fluorescent protein with a 3-FLAG instead of the MAP tag in order to directly use the plasmid described by Dickinson and co-workers because their technique seems to be very powerful to generate tagged protein in *C. elegans* (Dickinson *et al*, 2015). In addition, a fluorescent protein is also present so comparing to the MAP tag, we keep the advantage of being able to study the localization of the protein. Finally, the FLAG is generally used for immunoprecipitations and seems to allow specific protein purification (Gerace and Moazed, 2015). We thus requested the plasmid pDD283 containing the SEC with the 3-FLAG and the YPet as fluorescent protein (Dickinson *et al*, 2015) and we are currently making a Gibson Assembly to replace the *ccdB* fragments by homology arms to the end of *cdk*-12 (Figure 30). The downstream homology arm (3' homo) was produced by PCR from genomic DNA with primers sharing overlapping sequences to the plasmid to allow the Gibson Assembly. The upstream homology arm (5' homo), was synthesized *de novo* by IDT (Integrated DNA Technologies). Indeed, we failed to amplify this fragment by PCR reaction. In addition, it must contain the silent mutations to preclude Cas9 from cutting the modified locus (Figure 28C). We thus resorted to *de novo* synthesis of the 5' homo in order to directly incorporate these silent mutations in the fragment. The 5' homo was then amplified by PCR with primers containing overlapping sequences to the plasmid pDD283. The plasmid will next be cut with *Avr*II and *Spe*I to remove the *ccdB* markers (Dickinson *et al*, 2015) and the Gibson Assembly will be performed following the protocol of New England Biolabs. The last step will be the transformation of *E. coli* with the product resulting from the Gibson Assembly. Normally, only clones expressing the desired construction can grow because bacteria where the transformed plasmid still contained one or two *ccdB* fragments (in case where the Gibson Assembly failed) will not be able to develop. Indeed the protein CcdB is a poison for *E. coli* (Loris *et al*, 1999).

When we have all the plasmids, we plan to inject them in *C. elegans* to obtain a strain expressing a tagged CDK-12.

Discussion and perspectives

1. The co-CRISPR technique generates false-positive results

Performing CRISPR/Cas9 experiments on *C. elegans* requires screening for mutant worms. Several techniques can be used for this purpose such as the co-CRISPR strategy (Dickinson and Goldstein, 2016). This method consists of editing two loci during the same microinjection experience, the locus of interest and a marker locus that produces a visible phenotype when edited. Consequently, worms expressing the visible phenotype are normally more likely to contain the desired mutation than worms with a "*wild-type*" phenotype (Dickinson and Goldstein, 2016). In the lab, we were working with the *dpy-10* locus generating a roller phenotype when edited (Arribere *et al*, 2014). However, in our hands, this technique generates a lot of false-positive results as we obtained roller worms that were not modified at the *ama-1* or *cdk-12* loci. These false-positive results could be explained by the fact that two loci are not necessarily edited in the same way, for example, because the two crRNAs have different efficiency depending on their sequence. Indeed, a G-rich sequence upstream of the PAM containing no more than four T/U residues seems to be most active (Dickinson and Goldstein, 2016; Mouridi *et al*, 2017; Xu *et al*, 2015; Wu *et al*, 2014). False-positive (or false-negative) worms could also be obtained because the two loci are not edited in the same nuclei (Mouridi *et al*, 2017). Consequently, the presence of roller worms is more an indicator that the injection was properly performed and targeted the germline and that Cas9 is active rather than a 100% efficient screening strategy for difficult genome modifications (Mouridi *et al*, 2017). It should be noted that the size of the CTD largely exceeds the size of the commonly inserted sequences.

2. The self-excising cassette and the mCherry co-markers, a powerful technique to screen mutant worms

As the co-CRISPR strategy generates many false-positive results, we decided to change our screening method to use the self-excising cassette (SEC) described by Dickinson and co-workers (Dickinson *et al*, 2015) in a plasmid repair template. This SEC contains a hygromycin resistance gene, the dominant *sqt-1(d)* gene generating a roller phenotype and a heat-shock inducible Cre recombinase to excise the cassette once the transgenic strain is obtained. Plasmids injected in the gonad can form extrachromosomal arrays that can be transmitted to the progeny. In order to assess their formation, plasmids carrying genes encoding fluorescent proteins are co-injected. This allows us to distinguish worms created by a proper genomic replacement (they do not stably express the co-markers) from worms carrying the plasmids in extrachromosomal arrays (they express the fluorescent proteins). The injection of these co-markers is used by several laboratories and reported to be very effective (Dickinson *et al*, 2013; Frøkjaer-Jensen *et al*, 2010; Tzur *et al*, 2013).

The advantage of this cassette is the decreasing of the rate of false-positive worms because, by contrast to the co-CRISPR where a second locus is edited, the entire cassette is integrated into the genome at the targeted locus. Another advantage of this cassette is the presence of two positive selectable markers, the antibiotic resistance and the roller phenotype. The antibiotic resistance allows the detection of rare edits because even if the desired mutation occurs at a very low level, viable transgenic worms can be detected among all the dead (Dickinson and

Goldstein, 2016). This is opposite to the co-CRISPR where if the mutation of interest is rare, a lot of rollers can nevertheless be obtained. The roller phenotype normally allows the easy detection of homozygous mutant worms. Indeed, contrary to a heterozygous worm, the progeny of a homozygous is normally 100% roller (Dickinson *et al*, 2015). However, we discovered that the SEC is able to spontaneously self-excise without heat-shocking in the progeny of a homozygous worm leading to worm expressing the desired mutation but no more the roller phenotype. This result first surprised us as Dickinson and co-workers (Dickinson *et al*, 2015) do not mention this possibility but further readings from Kostrouchová and co-workers (Kostrouchová *et al*, 2017) who also used this cassette, reported a spontaneous excision of the SEC.

In this master thesis, we tried to use the SEC and the mCherry co-markers to generate two mutant strains, one expressing a mutated CTD and one expressing a tagged CDK-12. However, at the moment, we did not succeed in generating either of these two strains. Nevertheless, the use of the self-excising cassette with the mCherry co-markers seems to be a powerful method to identify mutant worms as the only hygromycin resistant, roller and fluorescence-negative worm we obtained was properly targeted at the *sld-3* locus.

3. Can we obtain our *C. elegans* mutant strains?

3.1. Strain expressing a mutated CTD

The inhibition of CDK-12 in the nematode *C. elegans* induces an L1 arrest and a decrease in the CTD-Ser2P level. Our current model to explain this data is that the CDK-12 dependent CTD-Ser2 phosphorylation is required for the efficient recruitment of the CstF/SL2 snRNP complex to genes in position 2 an over in operons. Consequently, the inhibition of CDK-12 results in un-*trans*-spliced mRNAs that become the target of the exonuclease XRN-2 or similar enzymes. The encoded proteins that are mostly involved in postembryonic development are thus not produced to the level required to overcome the L1 developmental arrest.

An important aspect of this project is to check that the phenotype induced by the inhibition of CDK-12 is indeed due to its action on the Pol II CTD and not to its role in translation (Choi *et al*, 2019; Coordes *et al*, 2015). We thus wanted to create a *C. elegans* strain expressing a CTD where the CTD-Ser2 are replaced by alanines (CTD-S2A) to analyze its phenotype. Because of the complexity of the sequence coding for the CTD, the sequence coding for the CTD-S2A had to be adapted (codon-optimized and removal of introns). Consequently, another sequence, the CTD-S2S, containing the same silent modifications than the CTD-S2A but with serines in position 2, was produced. This sequence was generated to study if these non-desired modifications have biological impacts.

To create the two mutant strains, one expressing the CTD-S2A and one expressing the CTD-S2S instead of the genomic CTD, we used the CRISPR/Cas9 method with currently no success. Therefore, we may wonder if they can be obtained or not. We think that they can be generated. Indeed, we already obtained a strain expressing the CTD-S2S using our first design (by injecting preassembled Cas9 RNPs and a linear repair template) despite the fact that only one transgenic strain was recovered. This strain is perfectly viable and shows no difference either at the level of the phenotype or at the expression of AMA-1 or CTD-Ser2P compared to *wild-type* worms. We are also confident that the strain expressing a CTD-S2A can be obtainable.

Indeed, using our first design, we obtained worms expressing a deleted CTD and even if worms homozygous for the CTD deletion are non-viable (what is in agreement with other studies showing that the deletion of the CTD is lethal (Corden, 2013)), the heterozygous worms follow the normal life cycle. This indicates that only one intact copy of the CTD is sufficient to allow the postembryonic development and that the guide RNAs properly targeted Cas9 to cut the endogenous *ama-1* locus. Consequently, there is a good probability that the CTD-S2A mutant can be obtained at least as a heterozygous. Moreover, CTD-S2A mutants had already been obtained in other organisms such as in yeast (Coudreuse *et al*, 2010) or human cells (Gu *et al*, 2012) though the construct was not integrated at the locus in the last case. The yeast CTD-S2A mutant mimics the yeast CDK-12 (named Lsk1) deletion and is viable despite showing a mating defect. We, therefore, hypothesize that the worm CTD-S2A mutant will mimic the L1 arrest observed when CDK-12 is inhibited.

3.2. Strain expressing a tagged CDK-12

As the inhibition of CDK-12 in *C. elegans* mimics the L1 developmental arrest that follows hatching in the absence of food, we postulate that the kinase could be regulated by environmental signals and thus by the insulin-like IGF-1 pathway probably by the PI3K/Akt or Ras/MAPK. To test this hypothesis, we want to create a mutant strain expressing a tagged CDK-12 in order to immunoprecipitate the kinase and analyze its phosphorylation status and partners. However, at the moment, we were unable to generate the strain. We think, yet, that a CDK-12 tagged strain must be obtainable because, in 2013, Bowman and co-workers generated a CDK-12-GFP strain by MosSCI (Bowman *et al*, 2013; Wormbase website). This strain is viable and the kinase is active suggesting that fusing a tag at the C-terminal domain of CDK-12 does not impact its expression or function. Moreover, in the lab, we already tried to create a strain expressing a GFP and FLAG-tagged version of CDK-12 using the bombardment technique. This technique consists of the bombardment of DNA-coated gold particles into the gonad of *C. elegans.* This leads to the random integration of a transgene into the genome (Schweinsberg and Grant, 2013). By using this method, worms expressing the GFP were obtained, meaning that the tagged kinase is able to express. However, the transgene was probably not integrated into the genome because the GFP signal disappeared within a few generations. This result suggests that a tagged version of CDK-12 can be expressed in *C. elegans*.

4. Perspectives

4.1. Biological perspectives

4.1.1. Strain expressing the CTD-S2A

As already mentioned, the goal of producing a *C. elegans* mutant strain expressing a CTD-S2A is to check that the phenotype induced by the inhibition of CDK-12 is well due to its action on the Pol II CTD namely the phosphorylation of the CTD-Ser2. Consequently, we are hoping that the replacement of all CTD-Ser2 by a non-phosphorylable residue such as alanines, will recapitulate the inhibition of CDK-12 and induces an L1 arrest. This will strengthen our model where the inhibition of CDK-12 leads to a developmental arrest because, in the absence of CTD-Ser2P, the SL2 snRNP is not recruited to operons and thus the mRNAs from genes involved in worm development are not protected and degraded by an exonuclease.

If our model is correct, worms homozygous for CTD-S2A should represent 25% of the progeny of a heterozygous worm and should arrest development at the L1 stage. Consequently, we will not be able to analyze the effect of the expression of a Pol II with a CTD-S2A in the other stages of development. Therefore, we thought of creating another mutant strain by MosSCI relying on the fact that the endogenous AMA-1 is inhibited by the mushroom toxin α-amanitin (Sanford *et al,* 1983). This mutant would express both endogenous alleles of *ama-1* and an additional mutated copy integrated by MosSCI, expressing a CTD-S2A with a second mutation conferring resistance to α-amanitin. This second mutation is, for example, a substitution of a guanine in adenine (*ama-1(m118)*) (Sanford *et al,* 1983; Bowman *et al,* 2011; Wormbase website). By incubating this mutant in the presence of α-amanitin, we should be able to inhibit the endogenous AMA-1 at any developmental stage. Consequently, at this chosen stage, the only active AMA-1 will be the α-amanitin resistant one expressing a CTD-S2A. This would allow us to study the impact of the non-phosphorylation of the CTD-Ser2 at all *C. elegans* developmental stages.

4.1.2. Strain expressing a tagged CDK-12

The goal of producing a tagged CDK-12 mutant strain is to study the phosphorylation pattern and the partners of the kinase. Indeed there is a likely possibility that CDK-12 is regulated by phosphorylation of its N-terminal domain because first, it is known that in *C. elegans* the nutrient availability is connected to metabolism by the insulin-like IGF-1 pathway and its downstream effector, the PI3K/Akt or Ras/MAPK. Second, because previous work in our lab showed that the *S. pombe* homolog of CDK-12 is regulated by MAPK-mediated phosphorylation of its N-terminal domain.

Therefore, we are planning to purify a tagged version of CDK-12 from worms hatched in the presence or absence of food and to analyze it by mass spectrometry. If our hypothesis that CDK-12 is regulated by the phosphorylation of its N-terminal domain by MAPK is correct, we are hoping to find that in one condition, CDK-12 interacts with MAPK and has a phosphorylated N-terminal domain and that in the other condition, CDK-12 has a non-phosphorylated N-terminal domain. This leads to the activation or not of the kinase and thus to postembryonic development or L1 arrest.

4.2. Technical perspectives

For the moment, we are trying to obtain the *C. elegans* mutant strains by using the self-excising cassette on a plasmid repair template either with a plasmid coding for Cas9 and plasmids coding for the sgRNAs or with a preassembled Cas9 RNP.

However, if this failed we could try to reuse a linear repair template because it is more efficient to allow homologous recombination with the genome than a plasmid repair template (Paix *et al,* 2014; Dickinson and Goldstein, 2016; Paix *et al,* 2017) and because it avoids the creation of extrachromosomal arrays. However as for insertions larger than 1 or 2kb (the self-excising cassette makes 5594bp), linear templates are not advisable (Paix *et al,* 2014; Dickinson and Goldstein, 2016; Paix *et al,* 2017), we should use another screening strategy. We could use for example only the gene *sqt-1(d)* or only a gene conferring an antibiotic resistance which would be integrated into the repair template. This gene would be incorporated into the genome

with the desired modification thus limiting the creation of false-positive but will allow the use of a linear repair template by decreasing the length of the inserts.

Conclusion

In this master thesis, we investigated the function of the kinase CDK-12 in the development of the nematode *C. elegans*. In the lab, it was previously shown that the inhibition of CDK-12 induced a decrease in the CTD-Ser2P level and that this decrease is correlated with an L1 arrest and a strong decrease of the mRNAs corresponding to genes in position 2 and over within operons. In this work, we tried to continue the study of CDK-12 by confirming that the phenotype induced by its inhibition is due to its function of phosphorylating the CTD-Ser2 and by finding how it is regulated. To do this, we tried to generate *C. elegans* mutant strains expressing either a Pol II with a CTD-S2A or a tagged CDK-12. To generate these strains, we tried to use variations of the CRISPR/Cas9 method but currently failed to obtain any of our strains of interest even though a technical control strain was obtained, indicating that the method itself is efficient.

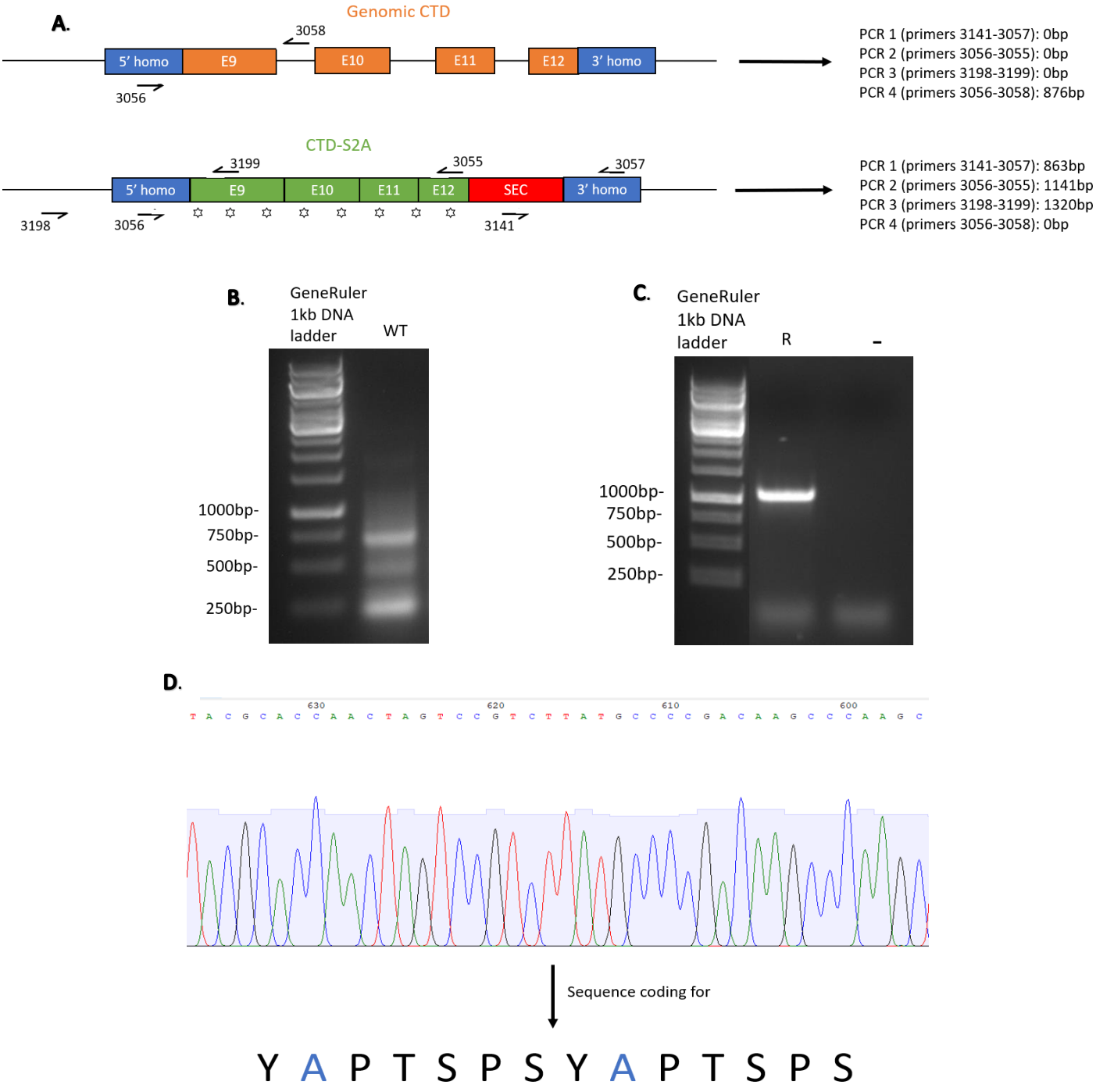


Figure 31: PCR reactions to analyse the expression of the CTD-S2A. (A). Cartoon of the two possible alleles that can be expressed by a worm with the length of the PCR product given by four different PCR reactions. **(B).** Agarose gel stained with ethidium bromide shows that **t**he first PCR (3141-3057) gives too much aspecific bands on a *wild-type* (WT) worm to be tested on potential mutant worms for the expression of the CTD-S2A . The second single worm PCR protocol (see material and methods) was used. **(C).** The agarose gel shows that the second PCR (3056-3055) gives a 1141bp band when performed on a roller, hygromycin resistant and not red worm (R) indicating the presence of the sequence encoding the CTD-S2A. The lane "-" is a *wild-type* worm as negative control. **(D).** Confirmation that the PCR product generated by the second PCR corresponds to the sequence encoding the CTD-S2A. The figure shows the sequence encoding two repeats of the heptapeptide corresponding to the CTD-S2A.

Addendum

From the injection of 86 worms in order to obtain a strain expressing the CTD-S2A by using two Cas9 RNPs, the plasmid repair template containing the SEC and the mCherry co-markers, I obtained 5 roller, hygromycin resistant but mCherry-positive worms, meaning that the injected plasmids most likely formed extrachromosomal arrays. However, while the writing of this master thesis was nearly completed, I obtained one roller, hygromycin resistant worm that was mCherry-negative, suggesting that it possibly expresses the CTD-S2A. Its progeny was indeed also roller and resistant to the presence of the antibiotic. To check that they express the desired modification and the self-excising cassette, I performed PCR reactions on these roller worms (Figure 31A). The first PCR (Figure 31A) uses the forward primer 3141 hybridizing in the cassette and a reverse primer (3057) hybridizing after the sequence encoding the CTD (in the region corresponding to the 3' CTD homology arm used in the repair plasmid). This PCR should give a 863bp band only if the sequence encoding the SEC is present. However, when I performed this PCR on *wild-type* worms, it gives numerous aspecific bands (Figure 31B) so I decided not to perform this PCR on potential mutant worms. The second PCR (Figure 31A) uses a forward primer (3056) hybridizing before the sequence encoding the CTD (in the region corresponding to the 5' CTD homology arm used in the repair plasmid) and a reverse primer hybridizing in the CTD-S2A (3055). This reverse primer is unable to bind in the sequence encoding the genomic CTD due to the silent mutations introduced in the mutated sequence by Eurofins Genomics. This second PCR should give a 1141bp band only if the sequence encoding the CTD-S2A is present. I thus performed this PCR on *wild-type* worms where indeed, no PCR product was obtained (Figure 31C) and on potential mutant worms where a 1141bp band was obtained (Figure 31C). I next sequenced this PCR product checking that it corresponds to the CTD-S2A (Figure 31D).

However, as primers 3056 and 3055 bind within the long homology regions of the repair template (as they were designed early on when we used a much shorter homology region), the obtained 1141bp amplicon does not warrant that the integration occurred at the targeted locus and could still result from the presence of an extrachromosomal array even though this is unlikely because worms are mCherry-negative. Consequently, to check that the roller, hygromycin resistant and mCherry-negative worms express the CTD-S2A from the endogenous locus, I am planning to perform a third PCR reaction (Figure 31A) using a forward primer (3198) hybridizing upstream of the sequence used as the 5' homology arm and a reverse primer (3199) hybridizing within the CTD-S2A. This PCR should give a 1320bp band only if the sequence encoding the CTD-S2A is incorporated at the *ama-1* locus.

I am also planning to perform a fourth PCR reaction (Figure 31A) using the forward primer 3056 and a reverse primer (3058) hybridizing in the first endogenous intron of the genomic CTD. This PCR will be performed to check if the obtained roller, hygromycin resistant and mCherry-negative worms are heterozygous or homozygous for the expression of the CTD-S2A. I expect heterozygous worms as it is very unlikely that the CTD-S2A mutant can be propagated.

Material and methods

Growth conditions and Strains

C. elegans strains are derived from the Bristol N2 strain. In the lab, worms are cultured on petri plates or in liquid medium. On petri plates, they are maintained at 20°C on NGM-agar plates (Nematode Growth Media – agar: 51,3mM Nacl; 1mM $CaCl_2$; 1mM $MgSO_4$; 25mM KH_2PO_4; 5µg/mL cholesterol and 1,25g peptone with 8,5g agar for 500mL)) seeded with *E. coli* OP50. In liquid medium, they are cultured on a shaker at 20°C in S-Basal (0,1M NaCl; 50mM KH_2PO_4; 3mM $MgSO_4$; 4mM $CaCl_2$; 1mM potassium citrate; 5µg/mL cholesterol) with *E. coli* HB101.

Brood size analysis and Embryonic mortality

On day 0, one L4 worm of each genotype is transferred to individual plates. On day 1, worms are adult and allow to lay eggs. On day 2, they are moved to another plate and on day 3, they are again removed. On day 4, the first plate is counted. On day 5, the second plate is counted and on day 6, the third one. On each counting day, larvae are counted as well as unhatched embryos. The sum of larvae plus unhatched embryos on all 3 plates of a single worm is its total brood size. The embryos unhatched after 48 hours are dead. The ratio of dead eggs to the total brood size is the embryonic mortality.

Protein extraction using the sonicator

Worms (± 18000 L4 in liquid culture) are pelleted, washed two times in M9 (42,2mM Na_2HPO_4; 22mM KH_2PO_4; 85,7mM NaCl; 1mM $MgSO_4$) and frozen. Thawed worms are resuspended in 4X Laemmli Sample Buffer (200mM Tris pH8,5; 8% SDS; 40% glycerol), boiled for 5 minutes at 95°C, vortexed and then sonicated on a Diagenode Bioruptor Sonicator for 10 cycles (30 seconds ON and 30 seconds OFF) to break the worms' cuticle. Next, the worms are spun and the supernatant (containing the soluble proteins) is collected. Protein concentration is determined by a Pierce assay and equal amounts of protein are loaded on Bio-Rad mini-PROTEAN TGX gels 4-15%. This protocol is mainly used for quantitative analysis.

Protein extraction using a NaOH extraction

Worms from one nearly starved plate (with lots of adult worms) are collected in M9 + Triton 0,01%. They are washed two times in M9 and one in demineralized water before being snap-freezed. They are then exposed to 300µL of water with 300µL of NaOH 0,6M for 10 minutes at room temperature and next spun 3 minutes at 14000rpm. The supernatant is discarded and the pellet (containing the protein) is resuspended in loading buffer (60mM Tris-HCl pH6,8; 4% β-mercaptoethanol; 4% SDS; 0,01% Bromophenol blue; 5% glycerol), incubated 10 minutes at 70°C and spun. The supernatant is finally collected and directly loaded on a Bio-Rad mini-PROTEAN TGX gels 4-15%. This protocol is used for qualitative analysis but not quantitative because protein concentration can not be determined. Indeed the basic pH of the NaOH interferes with the Pierce assay.

Western blot

Proteins are loaded on a Bio-Rad mini-PROTEAN TGX gel 4-15% and run for 45 minutes at 150V. Proteins are next transferred to a nitrocellulose membrane using the Bio-Rad Trans-Blot Turbo Transfer System (settle on High Molecular Weight 1,3A – 25V – 10 minutes or Mixed Molecular Weight 1,3A – 25 – 7 minutes). The membrane is blocked with Skin Milk Powder (5%) (Sigma-Aldrich) and exposed to primary antibodies overnight at 4°C (anti-α-Tubulin 1/1000 [Sigma-Aldrich T5168]; anti-CTD 8WG16 1/1000 [Eurogentec MMS-126P-050]; anti-CTD-Ser2P 3E10 1/1000 [Millipore 04-1571]; anti-FLAG 1/1000 [Sigma-Aldrich F3165]). After that, the membrane is washed 3 times in PBS-Tween (0,05%) and then exposed to secondary antibody (anti-mouse IgG Perox 1/5000 [GE NA931]; anti-rat IgG Perox 1/5000 [Dako P0450]). Finally, the membrane is washed two times with PBS-Tween (0,05%), one last time with PBS and revealed by PerkinElmer Western Lightning Plus-ECL on an ImageQuant LAS 4000 machine.

Ponceau staining

After the revelation, the membrane is washed with demineralized water and incubated in Ponceau solution (0,1% Ponceau S (Sigma-Aldrich P3504), 5% acetic acid) for 3 minutes. The membrane is then washed with water and finally photographed.

Immunofluorescence

Worms are killed in NaN_3 (50mM) on poly-lysine coated slides and cut in half with a 22G needle (BD Microlance 3). They are next frozen between slide and coverslip, for at least one hour, at -80°C on a metal plate, before being freeze-cracked by flipping the coverslip from the slide. The slides are then fixed using MetOH 100% (20 minutes at room temperature) and a 3,7% formaldehyde solution (30 minutes at room temperature) (10mL of Formaldehyde solution 37% [Sigma-Aldrich 252549] in 30 mL of PBS). They are then washed twice with PBS, once with PBS-tween (0,05%) and incubated with primary antibodies in wet chambers overnight at 4°C (anti-CTD 8WG16 1/2000 [Eurogentec MMS-126P-050]; anti-CTD-Ser2P 3E10 1/2000 [Millipore 04-1571]). The next day, the slides are washed three times with PBS-Tween (0,05%) and incubated with secondary antibodies one hour at room temperature in wet chambers (anti-rat IgG Alexa488 1/2000 [Invitrogen A21208]; anti-mouse IgG Alexa594 1/2000 [Invitrogen A11005]). Finally, the slides are washed twice with PBS-Tween (0,05%), once with PBS and mounted using Fluoroshield (containing DAPI) (Sigma-Aldrich F6057).

Slides are observed on a Zeiss Axio Imager Z1 microscope with a Hamamatsu Digital Camera C11440 using the Zeiss software ZEN. The time of the exposition was 3 seconds for Alexa488, 1 second for Alexa594 and 5 milliseconds for DAPI.

Single worm PCR, first protocol

Prior to PCR, worms are digested with proteinase K (0,2mg/mL final) (Roche Diagnostics 03 115 828 001) in 1X Colorless GoTaq Reaction Buffer (Promega M792A), for 1 hour at 65°C to obtain a worm lysate. The proteinase K is inactivated 15 minutes at 95°C before the worm lysate containing DNA is used for PCR.

The PCR reaction is performed using GoTaq polymerase (GoTaq G2 DNA Polymerase, Promega M748B) with 3µL of worm lysate from proteinase K treated worms, 5µL of 5X Green GoTaq Reaction Buffer (Promega M791A), 2µL of primers at 10µM, 0,25µL of dNTP at

20mM, 0,125µL of GoTaq polymerase and 12,7µL of water. The annealing temperature corresponds to the T_m (Primer Melting Temperature) of the primer with the lowest T_m less 2 and the extension time is 1 minute per kb.

Single worm PCR, second protocol

Prior to PCR, worms are digested with proteinase K (1mg/mL final) (Roche Diagnostics 03 115 828 001) in 1X Colorless GoTaq Reaction Buffer (Promega M792A), first by snap-freezing for 15 minutes and then by incubating 1 hour at 65°C to obtain a worm lysate. The proteinase K is inactivated 15 minutes at 95°C before the worm lysate containing DNA is used for PCR.

The PCR reaction is performed using GoTaq polymerase (GoTaq G2 DNA Polymerase, Promega M748B) with 1,5µL of worm lysate from proteinase K treated worms, 5µL of 5X Green GoTaq Reaction Buffer (Promega M791A), 2µL of primers at 10µM, 0,25µL of dNTP at 20mM, 0,125µL of GoTaq polymerase and 14,1µL of water. The annealing temperature corresponds to the T_m (Primer Melting Temperature) of the primer with the highest T_m plus 1,3 and the extension time is 1 minute per kb.

PCR

The PCRs to generate the linear repair templates for the CTD-S2A or CTD-S2S were performed using Expand polymerase (Expand High Fidelity PCR System, Roche Diagnostics 11732650001) on the plasmids created by Eurofins Genomics.

The PCRs for the amplification of the “MAP tag” were performed using GoTaq polymerase (GoTaq G2 DNA Polymerase, Promega M748B) on the plasmid pcDNA4 (order on Addgene).

The PCR for the amplification of the “5’ CTD homology arm” was performed using GoTaq polymerase (GoTaq G2 DNA Polymerase, Promega M748B) on worm lysate. The PCR for the amplification of the Self-Excising Cassette (SEC) and the “3’ CTD homology arm” were performed using Q5 polymerase (Q5 High-Fidelity DNA polymerase, New England Biolabs M0491S) on the plasmid pDD283 and on commercial *C. elegans* genomic DNA (Zyagen GC-280) respectively.

The PCRs to amplify the “5’ homology arm” and “3’ homology arm” to generate the plasmid repair template to obtain a tagged CDK-12 were made using Q5 polymerase (Q5 High-Fidelity DNA polymerase, New England Biolabs M0491S) on the gBlocks synthesized by IDT and commercial *C. elegans* genomic DNA (Zyagen GC-280) respectively.

The protocol provided by the manufacturer was followed for each enzyme.

Restriction reaction

The restrictions reactions were performed using the ThermoFisher Scientific FastDigest enzymes, *Bsa*I (FD0293), *Pvu*II (FD0634), *Mls*I (FD1214), *Eco*RI (FD0274), *Not*I (FD0593), *Pfo*I (FD1754), *Spe*I (FD1254) and *Avr*II (FD1564) following the protocol provided by the manufacturer for each enzyme.

Ligation of the “5’ CTD homology arm” in the pGEX

The “5’ CTD homology arm” amplified by PCR was ligated into the cut-pGEX using the T4 DNA ligase (New England BioLabs M0202) following the protocol of the manufacturer. 20µL of ligation products is next added to 50µL of competent cells (*E. coli* DH10B). The competent cells are left 10 minutes on ice, heat-shocked one minute at 42°C and directly returned to ice for 5 minutes. 950µL of LB Broth Base (Lennox L Broth Base, Invitrogen 12780-029) is added and bacteria are incubated 1 hour at 37°C, 300rpm. Then, bacteria are pelleted, 900µL of supernatant is removed and the cells are resuspended in the remaining volume of LB. This volume is finally spread onto an LB-amp-agar plate (Lennox L Agar, Invitrogen 22700-041; Ampicillin sodium salt, Sigma-Aldrich A9518-25G).

Gibson Assembly

The Gibson Assembly reactions were performed using the Gibson Assembly Cloning Kit of New England Biolabs (E5510S) following their protocols (Gibson Assembly Protocol (E5510) and Gibson Assembly Chemical Transformation Protocol (E2611)). The primers were designed with NEBuilder Assembly Tool.

Q5 Site-Directed Mutagenesis

The Q5 Site-Directed mutagenesis was made using the Q5 Site-Directed Mutagenesis Kit of New England Biolabs (E0554S) following their protocol (Q5 Site-Directed Mutagenesis Kit Quick Protocol (E0554)). The primers were designed with NEBaseChanger.

crRNA cloning in the pRB1017 plasmid

2µg of each oligonucleotide is resuspended in 50µL of annealing buffer (10mM Tris pH7,5; 50mM NaCl; 1mM EDTA). They are annealed by heating 5 minutes at 95°C. After allowing to come back to room temperature, annealed oligos are diluted into nuclease-free water to obtain a final concentration of 8ng/µL. These annealed diluted oligonucleotides are ligated into the *Bsa*I-cut pRB1017 vector using the T4 DNA ligase (Promega, M180A) by incubating 50ng of *Bsa*I-cut plasmid, 8ng of annealed diluted oligonucleotides, 1µL of Ligase 10X Buffer, 1µL of T4 DNA ligase and nuclease-free water for 3 hours at room temperature. The last step is the transformation of competent cells (*E. coli* DH10B) with 10µL of ligation product. The competent cells are left 10 minutes on ice, heat-shocked one minute at 42°C and directly returned to ice for 5 minutes. 950µL of LB Broth Base (Lennox L Broth Base, Invitrogen 12780-029) is added and bacteria are incubated 1 hour at 37°C, 300rpm. Then, bacteria are pelleted, 900µL of supernatant is removed and the cells are resuspended in the remaining volume of LB. This volume is finally spread onto an LB-kan-agar plate (Lennox L Agar, Invitrogen 22700-041; Kanamycin Monosulphate, Formedium KAN0005).

Miniprep

Minipreps are made with the Sigma-Aldrich GenElute Plasmid Miniprep Kit (PLN350-1KT) following the protocol provided by the manufacturer.

Midiprep

Midipreps are made with the Macherey-Nagel NucleoBond Xtra Midi Kit (740410) following the manufacturer protocol.

Sanger Sequencing

All Sanger sequencings are sent to Eurofins Genomics (Mix2Seq Kit).

CRISPR mix preparation

Preparation of the injection mix with a preassembled Cas9 ribonucleoprotein and the linear repair template is made as previously described (Dokshin et *al*, 2018). Briefly, the protein Cas9 (5µg), the tracrRNA (2µg), the crRNA against *Dpy-10* (0,56µg) and the crRNAs against the gene of interest (0,56µg of each) are incubated 10 minutes at 37°C to allow the formation of the Cas9 ribonucleoprotein. At the same time, 2µg of the linear repair template without homology arm (generated by PCR with the short primers), with 2µg of the linear repair template with homology arms (generated by PCR with the long primers) are heated to 95°C and cooled to 4°C to allow the melting and reannealing of the two PCR products. Finally, the repair template mixture is added to the Cas9 ribonucleoprotein with 2,2µg of the repair template for *Dpy-10 (cn64)* and water to obtain a final volume of 20µL. To avoid needle clogging, the CRISPR mix is spun at 14000 rpm for 2 minutes and about 17µL is transferred to a fresh tube.

Preparation of the injection mix using a mix of plasmids is made by adding in the same tube, the plasmid coding for Cas9 (50ng/µL final), the two plasmids coding for the crRNAs and tracrRNA (25ng/µL final), the plasmid repair template (100ng/µL final), the plasmid coding for *prab-3::m-Cherry* (10 ng/µL final), the plasmid coding for *myo-3::m-Cherry* (5ng/µL final), the plasmid coding for *myo-2::m-Chery* (2,5ng/µL final), and water to obtain a final volume of 20µL. To avoid needle clogging, the mix is spun at 14000 rpm for 2 minutes and about 17µL is transferred to a fresh tube (protocol adapted from Dickinson *et al*, 2015).

Preparation of the injection mix using a preassembled Cas9 RNP with the plasmid repair template is made by incubating 10 minutes at 37°C, the protein Cas9 (5µg), the tracrRNA (2µg), and the crRNAs against the gene of interest (0,56µg of each). The plasmid repair template (100ng/µL final), the plasmid coding for *myo-3::m-Cherry* (5ng/µL final), the plasmid coding for *myo-2::m-Chery* (2,5ng/µL final) and water to obtain a final volume of 20µL are then added. To avoid needle clogging, the mix is spun at 14000 rpm for 2 minutes and about 17µL is transferred to a fresh tube (protocol adapted from Au *et al*, 2019).

Microinjection

The day before the injection, L4 worms are picked onto a new NGM-agar plate. The objective is to obtain a synchronized population of young adults (containing about 10 embryos) to perform the injection (Kadandale *et al*, 2009). Indeed, L4 worms are difficult to inject as the gonad is not completely open and too old adult worms are also difficult to inject because they contain lots of eggs compressing the gonad and thus making it barely visible.

15 minutes before the injection, the young adult worms are transferred to an NGM-agar plate without bacteria because the presence of bacteria onto the worm's cuticle can clog the needle (Kadandale *et al*, 2009).

To perform the injection, a young adult is transferred to a 2% agarose pad covered in halocarbon oil 700 (Sigma-Aldrich H8898) using an eyelash mounted on a tip. This worm is next attached to the pad by pushing it with the eyelash. This allows preventing the worm from moving during the injection. The pad is next placed under the microscope (Zeiss Axio Vert A1) and the needle (Eppendorf, FemtoTip2) is loaded with 2µL of the injection mix and mounted onto the

micromanipulator (Eppendorf, FemtoJet2 with Eppendorf TransferMan NK 2). The needle is then placed near the worm and injection is performed outside the worm to check the needle's opening. If enough liquid is ejected, the worm can be injected but if no liquid is ejected, the needle has to be open. To open the needle, it is scraped against the agarose pad.

To inject the worm, the needle is inserted into the gonad syncytium and injection is performed. A sign that the injection was correctly performed is the visualization of the swelling of the gonad. After the injection, the worm is transferred onto a new culture plate using the eyelash and washed with 10µL of M9.

Selection of mutant worms using the self-excising cassette

After the injection (on day 0), parental worms are allowed to lay eggs for 4 days. On day 4, 250µL of hygromycin 10mg/mL (Hygromycin B from Streptomyces hygroscopicus, Sigma-Aldrich, H9773) is added per plate. From day 4 to day 11, plates are checked to find roller and hygromycin resistant F1 worms. These potential mutant worms are isolated into new plates (one worm per plate) and checked for the expression of the red co-markers using the Modular Stereo Microscope for Fluorescent Imaging Leica MZ10 F with the Leica Filter set ET mCherry (10450098). Only the not red worms are kept because they probably express the desired mutation. The progeny (F2) of these F1 roller, hygromycin resistant and not red worms are also isolated and the homozygous plates (where the entire progeny (F3) is roller) are identified. These homozygous worms are heat-shocked by incubating L1/L2 worms for 3,5 hours at 34°C. After heat-shock, worms are returned to 20°C for 2-3 days until they become adults (protocol adapted from Dickinson *et al*, 2015).

Annexes

1. Abbreviations

°C	celsius degree
µg	microgram
µL	microliter
µM	micromolar
3’ homo	3’ homology region
3’ss	3’ splice site
5’ homo	5’ homology region
5’ss	5’ splice site
A	ampere
Akt	Protein Kinase B (PKB)
AMA-1	Amanitin-binding subunit of RNA polymerase II
as	analog-sensitive
ATP	Adenosine Tri-Phosphate
bp	base pair
c-Abl	Abl tyrosine kinase isoform C
Cas9	CRISPR associated protein 9
CDK	Cyclin-Dependent Kinase
ChIP	Chromatin Immunoprecipitation
CPSF	Cleavage and Polyadenylation Specificity Factor
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
CstF	Cleavage stimulatory Factor
CTD	carboxy-terminal domain
CTD-S2A/CTD-S5A	RNA polymerase II carboxy-terminal domain where the serines in position 2/5 are replaced by alanines
CTD-S2S	RNA polymerase II carboxy-terminal domain containing silent mutations but retaining serines in position 2
CTD-Ser2/CTD-Ser5/CTD-Ser7	serine in position 2/5/7 of the RNA polymerase II carboxy-terminal domain
CTD-Ser2P/CTD-Ser5P/CTD-Ser7P	phosphorylated serine in position 2/5/7 of the RNA polymerase II carboxy-terminal domain
CTD-Thr4	threonine in position 4 of the RNA polymerase II carboxy-terminal domain
CTD-Tyr1	tyrosine in position 1 of the RNA polymerase II carboxy-terminal domain
DAF-2	Abnormal Dauer Formation 2
DAPI	4',6-diamidino-2-phénylindole
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
Dpb11	DNA polymerase B (II)
DSIF	DRB-Sensitivity-Inducing-Factor
dsRNA	double-stranded RNA
DTC	distal tip cell
Fcp1	CTD phosphatase FCP1
G	gauge
g	gram
GFP	Green Fluorescent Protein
His_8-tag	polyhistidine tag

IGF-1	Insulin-Like Growth Factor-1
ILP	Insulin-Like Peptide
kb	kilobase
Lsk1	Latrunculin Sensitive Kinase 1
M	molar
MAP tag	Multifunctional tandem Affinity Purification tag
MAPK	Mitogen-Activated Protein Kinase
MceI	mRNA-Capping Enzyme
mg	milligram
mL	milliliter
mM	millimolar
MNase	Micrococcal Nuclease
MosSCI	*Mos1* mediated Single Copy Insertion
MosTIC	*Mos1*-excision Transgene-Instructed gene Conversion
mRNA	messenger RNA
ncRNA	non-coding RNA
NELF	Negative Transcription Elongation Factor
ng	nanogram
NGM	Nematode Growth Media
nm	nanometer
Nrf2	Nuclear factor erythroid 2-Related Factor 2
ORF	Open Reading Frame
PAM	Protospacer-Adjacent-Motif
PAS	Polyadenylation Signal
Pcf11	Pre-mRNA Cleavage complex II protein
PCR	Polymerase Chain Reaction
PI3K	Phosphoinositide 3-Kinase
PIC	Pre-Initiation Complex
Pol	RNA polymerase
pre-mRNA	pre-messenger RNA
P-TEFb	Positive Transcription Elongation Factor b
qRT-PCR	quantitative Reverse Transcription-Polymerase Chain Reaction
Rat1	Ribonucleic Acid-Trafficking protein 1
RNA	ribonucleic acid
RNAi	RNA interference
RNAseq	RNA sequencing
RNP	ribonucleoprotein
RPAP2	RNA Polymerase II Associated Protein 2
Rpb	RNA Polymerase B
rpm	revolutions per minute
SBP	Streptavidin-Binding Peptide
SEC	self-excising cassette
sgRNA	single guide RNA
SL	spliced leader
SL RNA	splice leader RNA
SLD-3	Synthetically Lethal with Dpb11-1
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
Ssu72	CTD phosphatase SSU72
Ste11	Serine/Threonine-protein kinase STE11
TAP-tag	Tandem Affinity Purification tag
TBP	TATA Binding Protein
TFIIB/TFIID/TFIIE/TFIIH/TFIIF	transcription factor II B/D/E/F/H
T_m	Primer Melting Temperature
TMG cap	Trimethylguanosine cap

tracrRNA	trans-activating CRISPR RNA
UTR	Untranslated Region
V	volt
XRN2	5’-3’ exoribonuclease 2
YPet	Yellow Fluorescent Protein for energy transfer

2. Supplemental information

Figure S1: Sequencing of the 500bp PCR product expressed by worms injected for the expression of the CTD-S2A. These worms express a deleted CTD.

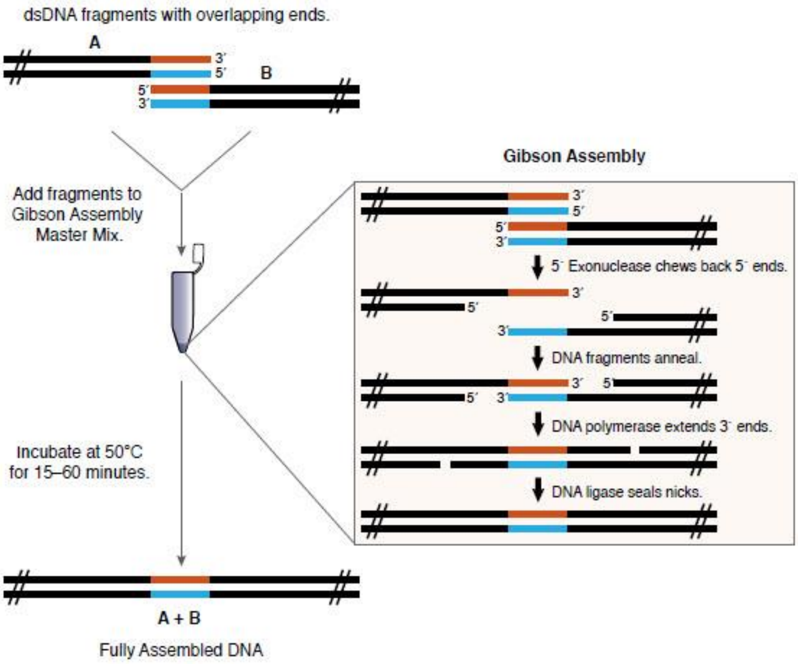
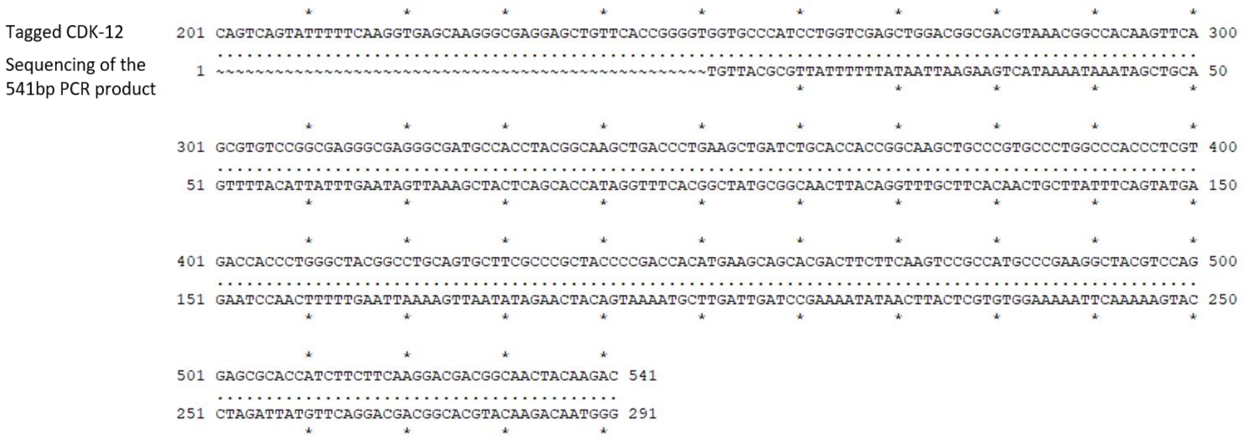


Figure S2: Gibson Assembly. The Gibson Assembly allows to join several DNA fragments (A and B) using overlapping regions. From these overlapping regions, an exonuclease will digest the 5' extremities to create single-strand 3' overhangs. These 3' free extremities can then anneal, and a polymerase will extend them using the other fragment as a template. Finally, a ligase will seal the nicks. All fragments that possess overlapping regions will thus be assemble (A+B) (adapted from New England Biolabs website).



Tagged CDK-12

Sequencing of the 541bp PCR product

```
             *         *         *         *         *         *         *         *         *         *
201 CAGTCAGTATTTTTCAAGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCA 300
    ....................................................................................................
  1 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~TGTTACGCGTTATTTTTTATAATTAAGAAGTCATAAAATAAATAGCTGCA 50
                                                               *         *         *         *         *

             *         *         *         *         *         *         *         *         *         *
301 GCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGT 400
    ....................................................................................................
 51 GTTTTACATTATTTGAATAGTTAAAGCTACTCAGCACCATAGGTTTCACGGCTATGCGGCAACTTACAGGTTTGCTTCACAACTGCTTATTTCAGTATGA 150
             *         *         *         *         *         *         *         *         *         *

             *         *         *         *         *         *         *         *         *         *
401 GACCACCCTGGGCTACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAG 500
    ....................................................................................................
151 GAATCCAACTTTTTGAATTAAAAGTTAATATAGAACTACAGTAAAATGCTTGATTGATCCGAAAATATAACTTACTCGTGTGGAAAAATTCAAAAAGTAC 250
             *         *         *         *         *         *         *         *         *         *

             *         *         *         *
501 GAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC 541
    .........................................
251 CTAGATTATGTTCAGGACGACGGCACGTACAAGACAATGGG 291
             *         *         *         *
```

Figure S3: Sequencing of a potential CDK-12 tagged mutant strain. Worms with a 541bp PCR product potentially express a tagged CDK-12. This PCR product was sequenced to see if it corresponded to CDK-12 with the incorporation of the MAP tag but it didn't. It means that theses worms are not mutants.

Table S1: Lists of primers used in this master thesis

Usage	5'-sequence- 3'	Forward/Reverse
crRNA targeting the beginning of the CTD	CAAUGAAGGAGGAUGGUCUC	/
crRNA targeting the end of the CTD	UAUGAAUUUGGAUCAUAAGU	/
Short primer to amplify the CTD-S2S from the plasmid produced by Eurofins Genomics	CTGGGTGCGTTATCCCCACG	Forward
Short primer to amplify the CTD-S2S from the plasmid produced by Eurofins Genomics	TTATGAGTTAGGATCATAGGTCGG	Reverse
Long primer to amplify the CTD-S2S from the plasmid produced by Eurofins Genomics	GTGGAACTCTGGAGTCACACCGACGTATGCTGG AGCCGCCTGGTCGCCTACCACAGGTGGAATGTC GCCTGGTGCTGGATTTTCACCGGCTGGAAATAC GGATGGAGGAGCATCGCCGTTCAATGAAGGAG GATGGTCTCCtGCATCGCCTGGGGATCCA-CTGGGTGCGTTATCCCCACG	Forward
Long primer to amplify the CTD-S2S from the plasmid produced by Eurofins Genomics	AATTAGAGGAGGAAAAAATTTTGAAAAATTAA AAGAAAAATAAATAAATAATGCTCAACATTTTT TTGGGGTAAAATAGATAGAAATCGGCGAAAAT TGCCCAAAAAAACGGGAAAAATC-TTATGAGTTAGGATCATAGGTCGG	Reverse
Short primer to amplify the CTD-S2A from the plasmid produced by Eurofins Genomics	TTAGGTGCCCTGAGTCCTCG	Forward
Short primer to amplify the CTD-S2A from the plasmid produced by Eurofins Genomics	TTAACTGTTCGGATCGTACGTC	Reverse
Long primer to amplify the CTD-S2A from the plasmid produced by Eurofins Genomics	GTGGAACTCTGGAGTCACACCGACGTATGCTGG AGCCGCCTGGTCGCCTACCACAGGTGGAATGTC GCCTGGTGCTGGATTTTCACCGGCTGGAAATAC GGATGGAGGAGCATCGCCGTTCAATGAAGGAG GATGGTCTCCTGCATCGCCTGGGGATCCA-TTAGGTGCCCTGAGTCCTCG	Forward

Long primer to amplify the CTD-S2A from the plasmid produced by Eurofins Genomics	AATTAGAGGAGGAAAAAATTTTGAAAAATTAA AAGAAAAATAAATAAATAATGCTCAACATTTTT TTGGGGTAAAATAGATAGAAATCGGCGAAAAT TGCCCAAAAAAACGGGAAAAATC-TTAACTGTTCGGATCGTACGTC	Reverse
Primer to screen for the expression of the mutated CTD	GCAGAATGTTGTAATGGGCG	Forward
Primer to screen for the expression of the mutated CTD	TGGCTCATGCTCACGCTCTG	Reverse
Primer used to sequence the PCR product corresponding to a deleted CTD	GCAGAATGTTGTAATGGGCG	Forward
Oligonucleotide coding for the crRNA targeting the beginning of the CTD and integrated in the pRB1017 plasmid	TCTTGCAATGAAGGAGGATGGTCTC	Forward
Oligonucleotide coding for the crRNA targeting the beginning of the CTD and integrated in the pRB1017 plasmid	AAACGAGACCATCCTCCTTCATTGC	Reverse
Oligonucleotide coding for the crRNA targeting the end of the CTD and integrated in the pRB1017 plasmid	TCTTGTATGAATTTGGATCATAAGT	Forward
Oligonucleotide coding for the crRNA targeting the end of the CTD and integrated in the pRB1017 plasmid	AAACACTTATGATCCAAATTCATAC	Reverse
Primer used to amplify the 5’ homo of the CTD and containing the restriction site for *Pvu*II	TTCATTAATGCAGCTGTCGGATTCACGGAGGCT	Forward
Primer used to amplify the 5’ homo of the CTD and containing the restriction site for *EcoR*I	CCGGAATTCTGGATCCCCAGGCGAT	Reverse
Primer used to amplify the SEC in order to insert it in the plasmid with the CTD-S2S	TGATCCTAACTCATAAGCGGCCGCGTAAGTTTA AAATAAC	Forward

Primer used to amplify the SEC in order to insert it in the plasmid with the CTD-S2S	AAAAATCGGTACCCTGAAAATAACTTCG	Reverse
Primer used to amplify the SEC in order to insert it in the plasmid with the CTD-S2A	GATCCGAACAGTTAAGCGGCCGCGTAAGTTTAAAATAAC	Forward
Primer used to amplify the SEC in order to insert it in the plasmid with the CTD-S2A	AAAAATCGGTACCCTGAAAATAACTTCG	Reverse
Primer used to amplify the 3' homo of the CTD	TTCAGGGTACCGATTTTTCCCGTTTTTTTGGGCAATTTTC	Forward
Primer used to amplify the 3' homo of the CTD	CAAGCTGTGACCGTCTACCGCCCACCCCTGATCT	Reverse
Primer used to perform the Q5 site-directed mutagenesis on the plasmid repair template for the mutated CTD	ATGGTCTCCAGCATCGCCTG	Forward
Primer used to perform the Q5 site-directed mutagenesis on the plasmid repair template for the mutated CTD	CCTCCTTCATTGAACGGCGA	Reverse
Primer 3191 used to screen the tagged SLD-3 strain	ACTACAAGGACGACGACGACAAG	Forward
Primer 3192 used to screen the tagged SLD-3 strain	CCTCCAACCACTCAAATTGTCCAC	Forward
Primer 3141 used to screen the tagged SLD-3 strain	GGGTCGATGCGACGCAA	Forward
Primer 3194 used to screen the tagged SLD-3 strain	GTATCATCAGAGAATGGCCTGTACC	Reverse
crRNA targeting the end of *cdk-12*	ACTGACTGATATTGTGATTG	/
Short primer to amplify the MAP tag	GTGAGCAAGGGCGAGGAGCT	Forward
Short primer to amplify the MAP tag	TTACTTGTACAGCTCGTCCA	Reverse
Long primer to amplify the MAP tag	CTGCCACCATCAGGAGGTCATGCACCTCCACCACCACCTCCACCAACACAGGCTTCTTCAACGTCGCATAACAATCATCAACCAGTTCCTCAGTCTCAGTACCAGTCAGTATTTTTCAAGGTGAGCAAGGGCGAGGAGCT	Forward

Long primer to amplify the MAP tag	TCAATTAAAAATTTAGACAGAAGAGCAGAAAA AACCAAAAAAAAAAGGAATTTGTTGTCTATAAC TAATAGAATAAAGAGAGTATACGTGTAAAAAT AATCAACAAAAAAATCAGAATTTTTACTTGTAC AGCTCGTCCA	Reverse
Primer used to screen for the expression of the MAP tag	AAGCAGCTCCTACACAACCATC	Forward
Primer used to screen for the expression of the MAP tag	GTCTTGTAGTTGCCGTCGTCC	Reverse 1
Primer used to screen for the expression of the MAP tag	TTGTGGCGGATCTTGAAGTTGG	Reverse 2
Oligonucleotide coding for the first crRNA targeting the end of *cdk-12* and integrated in the pRB1017	TCTTGACTGACTGATATTGTGATTG	Forward
Oligonucleotide coding for the first crRNA targeting the end of *cdk-12* and integrated in the pRB1017	AAACCAATCACAATATCAGTCAGTC	Reverse
Oligonucleotide coding for the second crRNA targeting the end of *cdk-12* and integrated in the pRB1017	TCTTGTGATATTGTGATTGAGGAAC	Forward
Oligonucleotide coding for the second crRNA targeting the end of *cdk-12* and integrated in the pRB1017	AAACGTTCCTCAATCACAATATCAC	Reverse
Primer used to amplify the 5’ homo of *cdk-12*	ACGTTGTAAAACGACGGCCAGTCGCCGGCAAG ATGAATTGGAAGAAATAAAGATTTC	Forward
Primer used to amplify the 5’ homo of *cdk-12*	CATCGATGCTCCTGAGGCTCCCGATGCTCCCTT GAAAAATACAGATTGGTACTG	Reverse
Primer used to amplify the 3’ homo of *cdk-12*	CGTGATTACAAGGATGACGATGACAAGAGATG AAAATTCTGATTTTTTTGTTG	Forward
Primer used to amplify the 3’ homo of *cdk-12*	GGAAACAGCTATGACCATGTTATCGATTTCTTC CGGAAAATCAAAAAATC	Reverse

References

Alaimo, P. J., Shogren-Knaak, M. A. & Shokat, K. M. (2001). Chemical genetic approaches for the elucidation of signaling pathways. *Current Opinion in Chemical Biology*, *5*(4), 360–367.

Arribere, J. A., Bell, R. T., Fu, B. X., Artiles, K. L., Hartman, P. S. & Fire, A. Z. (2014). Efficient Marker-Free Recovery of Custom Genetic Modifications with CRISPR/Cas9 in *Caenorhabditis elegans. Genetics*, *198*(3), 837–846.

Au, V., Li-Leger, E., Raymant, G., Flibotte, S., Chen, G., Martin, K., Fernando, L., Doell, C., Rosell, F. I., Wang, S. & Edgley, M. L. (2019). CRISPR/Cas9 Methodology for the Generation of Knockout Deletions in *Caenorhabditis elegans*. *G3: Genes, Genomes, Genetics*, *9*(1), 135–144.

Bartkowiak, B., Liu, P., Phatnani, H. P., Fuda, N. J., Cooper, J. J., Price, D. H., Adelman, K., Lis, J. T. & Greenleaf, A. L. (2010). CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes & Development*, *24*(20), 2303–2316.

Bartkowiak, B. & Greenleaf, A. L. (2011). Phosphorylation of RNAPII: To P-TEFb or not to P-TEFb?. *Transcription*, *2*(3), 115-119.

Baskaran, R., Dahmus, M. E. & Wang, J. Y. (1993). Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. *Proceedings of the National Academy of Sciences*, *90*(23), 11167-11171.

Baugh, L. R. (2013). To Grow or Not to Grow: Nutritional Control of Development During *Caenorhabditis elegans* L1 Arrest. *Genetics*, *194*(3), 539–555.

Björklund, S. & Gustafsson, C. M. (2005). The yeast Mediator complex and its regulation. *Trends in Biochemical Sciences*, *30*(5), 240–244.

Blumenthal, T., Davis, P. & Garrido-Lecca, A. (2018). Operon and non-operon gene clusters in the *C. elegans* genome. *Wormbook: The online Review of C. elegans.*

Bowman, E. A., Riddle, D. L. & Kelly, W. G. (2011). Amino Acid Substitutions in the *Caenorhabditis elegans* RNA Polymerase II Large Subunit AMA-1/RPB-1 that Result in α-Amanitin Resistance and/or Reduced Function. *G3: Genes, Genomes, Genetics*, *1*(6), 411–416.

Bowman, E. A., Bowman, C. R., Ahn, J. H. & Kelly, W. G. (2013). Phosphorylation of RNA polymerase II is independent of P-TEFb in the *C. elegans* germline. *Development*, *140*(17), 3703–3713.

Bowman, E. A. & Kelly, W. G. (2014). RNA Polymerase II transcription elongation and Pol II CTD Ser2 phosphorylation. *Nucleus*, *5*(3), 224–236.

Brenner, S. (1974). The Genetics of *Caenorhabditis Elegans*. *Genetics*, *77*(1), 71–94.

Brenner, S. (2003). Nobel lecture: nature's gift to science. *Bioscience reports*, *23*(5), 225-237.

Briand, J. B. (2015). Etude du contrôle de la transcription envahissante par la terminaison de la transcription. *Université Paris Sud – Paris XI.*

Chapman, R. D., Conrad, M. & Eick, D. (2005). Role of the Mammalian RNA Polymerase II C-Terminal Domain (CTD) Nonconsensus Repeats in CTD Stability and Cell Proliferation. *Molecular and Cellular Biology*, *25*(17), 7665–7674.

Cheung, A. C. M. & Cramer, P. (2012). A Movie of RNA Polymerase II Transcription. *Cell*, *149*(7), 1431–1437.

Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J. & Buratowski, S. (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes & Development*, *15*(24), 3319–3329.

Choi, S. H., Martinez, T. F., Kim, S., Donaldson, C., Shokhirev, M. N., Saghatelian, A. & Jones, K. A. (2019). CDK12 phosphorylates 4E-BP1 to enable mTORC1-dependent translation and mitotic genome stability. *Genes & development*, *33*(7-8), 418-435.

Clancy, S. & Brown, W. (2008). Translation: DNA to mRNA to protein. *Nature Education*, *1*(1), 101.

Clokey, G. V. & Jacobson, L. A. (1986). The autofluorescent "lipofuscin granules" in the intestinal cells of *Caenorhabditis elegans* are secondary lysosomes. *Mechanisms of Ageing and Development*, *35*(1), 79–94.

Conte Jr, D., MacNeil, L. T., Walhout, A. J. & Mello, C. C. (2015). RNA Interference in *Caenorhabditis elegans*. *Current Protocols in Molecular Biology*, *109*(1), 26-3.

Cooper, G. M. (2000). Eukaryotic RNA Polymerases and General Transcription factors. In *The Cell: A Molecular Approach. 2nd Edition*. Sunderland (MA): Sinauer Associates.

Coordes, B., Brünger, K. M., Burger, K., Soufi, B., Horenk, J., Eick, D., Olsen, J. V. & Sträßer, K. (2015). Ctk1 Function Is Necessary for Full Translation Initiation Activity in *Saccharomyces cerevisiae*. *Eukaryotic Cell*, *14*(1), 86–95.

Corden, J. L. (1990). Tails of RNA polymerase II. *Trends in biochemical sciences*, *15*(10), 383-387.

Corden, J. L. (2013). RNA Polymerase II C-Terminal Domain: Tethering Transcription to Transcript and Template. *Chemical Reviews*, *113*(11), 8423–8455.

Corsi, A. K., Wightman, B. & Chalfie, M. (2015). A Transparent Window into Biology: A Primer on *Caenorhabditis elegans*. *Genetics*, *200*(2), 387–407.

Coudreuse, D., Van Bakel, H., Dewez, M., Soutourina, J., Parnell, T., Vandenhaute, J., Cairns, B., Werner, M. & Hermand, D. (2010). A Gene-Specific Requirement of RNA Polymerase II CTD Phosphorylation for Sexual Differentiation in *S. pombe*. *Current Biology*, *20*(12), 1053–1064.

Couillault, C. & Kurz, C. L. (2010). Du nématode *Caenorhabditis elegans* et de son utilisation en laboratoire. *Stal*, *36*, 59-74.

Cramer, P., Armache, K. J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G. E., Dengl, S., Geiger, S. R., Jasiak, A. J. & Jawhari, A. (2008). Structure of Eukaryotic RNA Polymerases. *Annual Review of Biophysics*, *37*(1), 337–352.

Daniel, J., Stegehake, D. & Liebau, E. (2013). Microinjection into the gonads of *C. elegans using* Eppendorf InjectMan® 4. *Eppendorf Application Note*.

David, C. J., Boyne, A. R., Millhouse, S. R. & Manley, J. L. (2011). The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65–Prp19 complex. *Genes & Development*, *25*(9), 972–983.

Dhingra, N., Bruck, I., Smith, S., Ning, B. & Kaplan, D. L. (2015). Dpb11 Protein Helps Control Assembly of the Cdc45·Mcm2-7·GINS Replication Fork Helicase. *Journal of Biological Chemistry*, *290*(12), 7586–7601.

Dickinson, D. J., Ward, J. D., Reiner, D. J. & Goldstein, B. (2013). Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. *Nature Methods*, *10*(10), 1028–1034.

Dickinson, D. J., Pani, A. M., Heppert, J. K., Higgins, C. D. & Goldstein, B. (2015). Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette. *Genetics*, *200*(4), 1035–1049.

Dickinson, D. J. & Goldstein, B. (2016). CRISPR-Based Methods for *Caenorhabditis elegans* Genome Engineering. *Genetics*, *202*(3), 885–901.

Dokshin, G. A., Ghanta, K. S., Piscopo, K. M. & Mello, C. C. (2018). Robust Genome Editing with Short Single-Stranded and Long, Partially Single-Stranded DNA Donors in *Caenorhabditis elegans*. *Genetics*, *210*(3), 781–787.

Egloff, S., Zaborowska, J., Laitem, C., Kiss, T. & Murphy, S. (2012). Ser7 Phosphorylation of the CTD Recruits the RPAP2 Ser5 Phosphatase to snRNA Genes. *Molecular Cell*, *45*(1), 111–122.

Eissenberg, J. C., Shilatifard, A., Dorokhov, N. & Michener, D. E. (2007). Cdk9 is an essential kinase in Drosophila that is required for heat shock gene expression, histone methylation and elongation factor recruitment. *Molecular Genetics and Genomics*, *277*(2), 101–114.

Evans, D., Perez, I., MacMorris, M., Leake, D., Wilusz, C. J. & Blumenthal, T. (2001). A complex containing CstF-64 and the SL2 snRNP connects mRNA 3′ end formation and trans-splicing in *C. elegans* operons. *Genes & Development*, *15*(19), 2562–2571.

Evans, T. (2006). Transformation and microinjection. *WormBook: The Online Review of C. elegans.*

Ferris, H. & Hieb, W. F. (2015). Ellsworth C. Dougherty: A Pioneer in the Selection of *Caenorhabditis elegans* as a Model Organism. *Genetics*, *200*(4), 991–1002.

Fox, R. M., Watson, J. D., Von Stetina, S. E., McDermott, J., Brodigan, T. M., Fukushige, T., Krause, M. & Miller, D. M. (2007). The embryonic muscle transcriptome of *Caenorhabditis elegans*. *Genome Biology*, *8*(9), R188.

Frézal, L. & Félix, M. A. (2015). *C. elegans* outside the Petri dish. *ELife*, *4*, e05849.

Frøkjær-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., Grunnet, M. & Jorgensen, E. M. (2008). Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nature Genetics*, *40*(11), 1375–1383.

Frøkjær-Jensen, C., Davis, M. W., Hollopeter, G., Taylor, J., Harris, T. W., Nix, P., Lofgren, R., Prestgard-Duke, M., Bastiani, M., Moerman, D. G. & Jorgensen, E.M. (2010). Targeted gene deletions in *C. elegans* using transposon excision. *Nature Methods*, *7*(6), 451–453.

Gerace, E. & Moazed, D. (2015). Affinity Pull-Down of Proteins Using Anti-FLAG M2 Agarose Beads. In *Methods in Enzymology*. Cambridge: Academic Press.

Greenleaf, A. L. (1992). 3 RNA Polymerase II Subunits, Transcription Factors, and Kinases: Investigations in *Drosophila melanogaster* and *Saccharomyces cerevisiae*. *Cold Spring Harbor Monograph Archive*, *22*, 55-80.

Gu, B., Eick, D. & Bensaude, O. (2012). CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. *Nucleic Acids Research*, *41*(3), 1591–1603.

Hall, D. H., Herndon, L. A. & Altun, Z. (2017). Introduction to *C. elegans* Embryo Anatomy. *Wormatlas.*

Hsin, J. P., Sheth, A. & Manley, J. L. (2011). RNAP II CTD Phosphorylated on Threonine-4 Is Required for Histone mRNA 3′ End Processing. *Science*, *334*(6056), 683–686.

Huang, T., Kuersten, S., Deshpande, A. M., Spieth, J., MacMorris, M. & Blumenthal, T. (2001). Intercistronic Region Required for Polycistronic Pre-mRNA Processing in *Caenorhabditis elegans*. *Molecular and Cellular Biology*, *21*(4), 1111–1120.

Kadandale, P., Chatterjee, I. & Singson, A. (2009). Germline Transformation of *Caenorhabditis elegans* by Injection. In *Microinjection: Methods and Applications*. New York: Humana Press.

Kaplan, R. E. W. & Baugh, L. R. (2016). L1 arrest, daf-16/FoxO and nonautonomous control of post-embryonic development. *Worm*, *5*(2), e1175196.

Karp, G. (2010). L'expression des gènes: de la transcription à la traduction. In *Biologie cellulaire et moléculaire: Concepts et expériences*. Louvain-la-Neuve: De Boeck Supérieur.

Kim, W. Y. & Dahmus, M. E. (1989). The major late promoter of adenovirus-2 is accurately transcribed by RNA polymerases IIO, IIA, and IIB. *Journal of Biological Chemistry*, *264*(6), 3169-3176.

Kim, M., Krogan, N. J., Vasiljeva, L., Rando, O. J., Nedea, E., Greenblatt, J. F. & Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*, *432*(7016), 517–522.

Kornblihtt, A. R. (2004). Shortcuts to the end. *Nature Structural & Molecular Biology*, *11*(12), 1156–1157.

Kostrouchová, M., Kostrouch, D., Chughtai, A. A., Kaššák, F., Novotný, J. P., Kostrouchová, V., Benda, A., Krause, M. W., Saudek, V., Kostrouchová, M. & Kostrouch, Z. (2017). The nematode homologue of Mediator complex subunit 28, F28F8. 5, is a critical regulator of C. elegans development. *PeerJ*, *5*, e3390.

Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C. & Hampsey, M. (2004). Ssu72 Is an RNA Polymerase II CTD Phosphatase. *Molecular Cell*, *14*(3), 387–394.

Kubben, N., Voncken, J. W. & Misteli, T. (2010). Mapping of protein- and chromatin-interactions at the nuclear lamina. *Nucleus*, *1*(6), 460–471.

Kumar, K. P., Akoulitchev, S. & Reinberg, D. (1998). Promoter-proximal stalling results from the inability to recruit transcription factor IIH to the transcription complex and is a regulated event. *Proceedings of the National Academy of Sciences*, *95*(17), 9767–9772.

Kutscher, L. M. & Shaham, S. (2014). Forward and reverse mutagenesis in *C. elegans*. *WormBook : The Online Review of C. elegans.*

Lasda, E. L. & Blumenthal, T. (2011). Trans-splicing. *Wiley Interdisciplinary Reviews: RNA*, *2*(3), 417–434.

Lee, H., Choi, M. K., Lee, D., Kim, H. S., Hwang, H., Kim, H., Park, S., Paik, Y. K. & Lee, J. (2012). Nictation, a dispersal behavior of the nematode *Caenorhabditis elegans*, is regulated by IL2 neurons. *Nature Neuroscience*, *15*(1), 107–112.

Li, Y. (2010). Commonly used tag combinations for tandem affinity purification. *Biotechnology and Applied Biochemistry*, *55*(2), 73–83.

Li, X., Chatterjee, N., Spirohn, K., Boutros, M. & Bohmann, D. (2016). Cdk12 Is A Gene-Selective RNA Polymerase II Kinase That Regulates a Subset of the Transcriptome, Including Nrf2 Target Genes. *Scientific Reports*, *6*(1), 1–13.

Lints, R. & Hall, D. H. (2009). Reproductive system overview. *Wormatlas*.

Loris, R., Dao-Thi, M. H., Bahassi, E. M., Van Melderen, L., Poortmans, F., Liddington, R., Couturier, M. & Wyns, L. (1999). Crystal structure of CcdB, a topoisomerase poison from *E. coli*. *Journal of Molecular Biology*, *285*(4), 1667–1677.

Lunde, B. M., Reichow, S. L., Kim, M., Suh, H., Leeper, T. C., Yang, F., Mutschler, H., Buratowski, S., Meinhart, A. & Varani, G. (2010). Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nature Structural & Molecular Biology*, *17*(10), 1195–1201.

Luo, W., Johnson, A. W. & Bentley, D. L. (2006). The role of Rat1 in coupling mRNA 3′-end processing to transcription termination: implications for a unified allosteric–torpedo model. *Genes & development*, *20*(8), 954-965.

Ma, H., McLean, J. R., Chao, L. F. I., Mana-Capelli, S., Paramasivam, M., Hagstrom, K. A., Gould, K. L. & McCollum, D. (2012). A Highly Efficient Multifunctional Tandem Affinity Purification Approach Applicable to Diverse Organisms. *Molecular & Cellular Proteomics*, *11*(8), 501–511.

Ma, H., McLean, J. R., Gould, K. L. & McCollum, D. (2014). An Efficient Fluorescent Protein-Based Multifunctional Affinity Purification Approach in Mammalian Cells. In *Protein Affinity Tags: Methods and Protocols*. New York: Humana Press.

Malumbres, M. (2014). Cyclin-dependent kinases. *Genome Biology*, *15*(6), 122.

Materne, P., Anandhakumar, J., Migeot, V., Soriano, I., Yague-Sanz, C., Hidalgo, E., Mignion, C., Quintales, L., Antequera, F. & Hermand, D. (2015). Promoter nucleosome dynamics regulated by signalling through the CTD code. *ELife*, *4*, e09008.

Mattick, J. S. & Makunin, I. V. (2006). Non-coding RNA. *Human Molecular Genetics*, *15*(suppl_1), R17–R29.

Mayer, A., Landry, H. M. & Churchman, L. S. (2017). Pause & go: From the discovery of RNA polymerase pausing to its functional implications. *Current Opinion in Cell Biology*, *46*, 72–80.

Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. (1991). Efficient gene transfer in *C.elegans*: Extrachromosomal maintenance and integration of transforming sequences. *The EMBO Journal*, *10*(12), 3959–3970.

Mouridi, E. S., Lecroisey, C., Tardy, P., Mercier, M., Leclercq-Blondel, A., Zariohi, N. & Boulin, T. (2017). Reliable CRISPR/Cas9 Genome Engineering in *Caenorhabditis elegans* Using a Single Efficient sgRNA and an Easily Recognizable Phenotype. *G3: Genes, Genomes, Genetics*, *7*(5), 1429–1437.

Mukhopadhyay, A. & Tissenbaum, H. A. (2007). Reproduction and longevity: Secrets revealed by *C. elegans*. *Trends in Cell Biology*, *17*(2), 65–71.

Murthy, K. G. & Manley, J. L. (1995). The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes & Development*, *9*(21), 2672–2683.

Nanji, M., Hopper, N. A. & Gems, D. (2005). LET-60 RAS modulates effects of insulin/IGF-1 signaling on development and aging in *Caenorhabditis elegans*. *Aging Cell*, *4*(5), 235–245.

Nigon, V. M. & Félix, M. A. (2018). History of research on *C. elegans* and other free-living nematodes as model organisms. *Wormbook: The Online Review of C. elegans.*

Nonet, M., Sweetser, D. & Young, R. A. (1987). Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. *Cell*, *50*(6), 909–915.

Orphanides, G., Lagrange, T. & Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes & Development*, *10*(21), 2657–2683.

Paix, A., Wang, Y., Smith, H. E., Lee, C. Y. S., Calidas, D., Lu, T., Smith, J., Schmidt, H., Krause, M. W. & Seydoux, G. (2014). Scalable and Versatile Genome Editing Using Linear DNAs with Microhomology to Cas9 Sites in *Caenorhabditis elegans*. *Genetics*, *198*(4), 1347–1356.

Paix, A., Folkmann, A., Rasoloson, D. & Seydoux, G. (2015). High Efficiency, Homology-Directed Genome Editing in *Caenorhabditis elegans* Using CRISPR-Cas9 Ribonucleoprotein Complexes. *Genetics*, *201*(1), 47–54.

Paix, A., Folkmann, A. & Seydoux, G.. (2017). Precision genome editing using CRISPR-Cas9 and linear repair templates in *C. elegans*. *Methods*, *121*, 86–93.

Plasterk, R. H. & Groenen, J. T. (1992). Targeted alterations of the *Caenorhabditis elegans* genome by transgene instructed DNA double strand break repair following Tc1 excision. *The EMBO Journal*, *11*(1), 287–290.

Ranuncolo, S. M., Ghosh, S., Hanover, J. A., Hart, G. W. & Lewis, B. A. (2012). Evidence of the Involvement of O-GlcNAc-modified Human RNA Polymerase II CTD in Transcription in Vitro and in Vivo. *Journal of Biological Chemistry*, *287*(28), 23549–23561.

Riddle, D. L., Blumenthal, T., Meyer, B. J. & Priess, J. R. (1997). Section II Origins of the model. In *C. elegans II. 2nd edition.* New York: Cold Spring Harbor Laboratory Press.

Sanford, T., Golomb, M. & Riddle, D. L. (1983). RNA polymerase II from wild type and alpha-amanitin-resistant strains of *Caenorhabditis elegans*. *Journal of Biological Chemistry*, *258*(21), 12804–12809.

Schüller, R., Forné, I., Straub, T., Schreieck, A., Texier, Y., Shah, N., Decker, T. M., Cramer, P., Imhof, A. & Eick, D. (2016). Heptad-Specific Phosphorylation of RNA Polymerase II CTD. *Molecular Cell*, *61*(2), 305–314.

Schweinsberg, P. J. & Grant, B. D. (2013). *C. elegans* gene transformation by microparticle bombardment. *WormBook: The Online Review of C. elegans.*

Schwer, B. & Shuman, S. (2011). Deciphering the RNA Polymerase II CTD Code in Fission Yeast. *Molecular Cell*, *43*(2), 311–318.

Semple, J. I., Garcia-Verdugo, R. & Lehner, B. (2010). Rapid selection of transgenic *C. elegans* using antibiotic resistance. *Nature Methods*, *7*(9), 725–727.

Shapiro, J. A. (2009). Revisiting the central dogma in the 21st century. *Annals of the New York Academy of Sciences*, *1178*, 6–28.

Shen, X., Valencia, C. A., Gao, W., Cotten, S. W., Dong, B., Huang, B. C. & Liu, R. (2008). Ca2+/Calmodulin-binding proteins from the *C. elegans* proteome. *Cell Calcium*, *43*(5), 444–456.

Shim, E. Y., Walker, A. K., Shi, Y. & Blackwell, T. K. (2002). CDK-9/cyclin T (P-TEFb) is required in two postinitiation pathways for transcription in the *C. elegans* embryo. *Genes & Development*, *16*(16), 2135–2146.

Sterken, M. G., Snoek, L. B., Kammenga, J. E. & Andersen, E. C. (2015). The laboratory domestication of *Caenorhabditis elegans*. *Trends in Genetics*, *31*(5), 224–231.

Svejstrup, J. Q. (2004). The RNA polymerase II transcription cycle: Cycling through chromatin. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*, *1677*(1), 64–73.

Templeman, N. M. & Murphy, C. T. (2018). Regulation of reproduction and longevity by nutrient-sensing pathways. *The Journal of Cell Biology*, *217*(1), 93–106.

Tsutsui, T., Fukasawa, R., Tanaka, A., Hirose, Y. & Ohkuma, Y. (2011). Identification of target genes for the CDK subunits of the Mediator complex. *Genes to Cells*, *16*(12), 1208–1218.

Tzur, Y. B., Friedland, A. E., Nadarajan, S., Church, G. M., Calarco, J. A. & Colaiácovo, M. P. (2013). Heritable Custom Genomic Modifications in *Caenorhabditis elegans* via a CRISPR–Cas9 System. *Genetics*, *195*(3), 1181–1185.

Van Werven, F. J. & Amon, A. (2011). Regulation of entry into gametogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1584), 3521–3531.

Will, C. L. & Lührmann, R. (2006). 13 Spliceosome Structure and Function. *Cold Spring Harbor Monograph Archive*, *43*, 369-400.

Wong, K. H., Jin, Y. & Struhl, K. (2014). TFIIH Phosphorylation of the Pol II CTD Stimulates Mediator Dissociation from the Preinitiation Complex and Promoter Escape. *Molecular Cell*, *54*(4), 601–612.

Wood, W. B. (1988). 1 Introduction to *C. elegans* Biology. *Cold Spring Harbor Monograph Archive*, *17*, 1–16.

Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., Dadon, D. B., Cheng, A. W., Trevino, A. E., Konermann, S., Chen, S. & Jaenisch, R. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nature Biotechnology*, *32*(7), 670–676.

Xu, H., Xiao, T., Chen, C. H., Li, W., Meyer, C. A., Wu, Q., Wu, D., Cong, L., Zhang, F., Liu, J. S. & Brown, M. (2015). Sequence determinants of improved CRISPR sgRNA design. *Genome Research*, *25*(8), 1147–1157.

Yamaguchi, Y., Shibata, H. & Handa, H. (2013). Transcription elongation factors DSIF and NELF: Promoter-proximal pausing and beyond. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1829*(1), 98–104.

Zaborowska, J., Egloff, S. & Murphy, S. (2016). The pol II CTD: new twists in the tail. *Nature structural & molecular biology*, *23*(9), 771.

Zegerman, P. & Diffley, J. F. X. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature*, *445*(7125), 281–285.

Zhang, Z., Fu, J. & Gilmour, D. S. (2005). CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3′-end processing factor, Pcf11. *Genes & Development*, *19*(13), 1572–1580.

Zhang, D. W., Rodríguez-Molina, J. B., Tietjen, J. R., Nemec, C. M. & Ansari, A. Z. (2012a). Emerging views on the CTD code. *Genetics Research International*, *2012*, 347214.

Zhang, D. W., Mosley, A. L., Ramisetty, S. R., Rodríguez-Molina, J. B., Washburn, M. P. & Ansari, A. Z. (2012b). Ssu72 Phosphatase-dependent Erasure of Phospho-Ser7 Marks on the RNA Polymerase II C-terminal Domain Is Essential for Viability and Transcription Termination. *Journal of Biological Chemistry*, *287*(11), 8541–8551.

Zhou, M. & Law, J. A. (2015). RNA Pol IV and V in gene silencing: Rebel polymerases evolving away from Pol II's rules. *Current Opinion in Plant Biology*, *27*, 154–164.

Resource pages

https://international.neb.com/

https://www.mpg.de/en

https://www.uniprot.org/

https://www.wormatlas.org/

https://www.wormbase.org/#012-34-5