

# Mechanisms and Developmental Roles of Promoter-proximal Pausing of RNA Polymerase II

Christine Robinson, Matthew Lowe, Amanda Schwartz and Nobuaki Kikyo\*

Stem Cell Institute, Department of Genetics, Cell Biology and Development, University of Minnesota, USA

## Abstract

RNA polymerase II (Pol II) temporarily stops transcription after synthesizing 30-50 bases, and resumes elongation only after stimulations by various signaling molecules and developmental cues. This phenomenon, called promoter-proximal pausing, is observed in 10-50% of the entire genes from *Drosophila* embryos to human cells. Release of paused Pol II is primarily mediated by the activated form of positive transcription elongation factor b (P-TEFb) initially sequestered in the inhibitory 7SK small nuclear ribonucleoprotein (7SK snRNP) complex. Many proteins and RNAs have been discovered and studied in detail to explain the process of the pausing and release of Pol II in relation to P-TEFb. At the functional level, promoter-proximal pausing regulates genes involved in stimulus-response and development in *Drosophila*. In mammalian stem cell biology, pausing is important for proliferation and signaling in embryonic stem cells and the formation of induced pluripotent stem cells. Other than this, however, little is known about the biological significance of pausing in mammalian cell differentiation. Further study on pausing mechanisms as well as its functions will contribute to the development of stem cell biology and its clinical applications.

**Keywords:** 7SK snRNP; Differentiation; Gene expression; Promoter; P-TEFb; RNA polymerase II; Stem cells

**Abbreviations:** 7SK snRNP: 7SK Small Nuclear Ribonucleoprotein; BET: Bromodomain and Extraterminal Domain; BMP: Bone Morphogenetic Protein; Brd4: Bromodomain containing 4; Br-UTP: 5-bromouridine 5'-triphosphate; CTD: Carboxy-Terminal Heptapeptide Repeat Domain; DRB: 5,6-Dichloro-1-β-D-Ribofuranosylbenzimidazole; DSIF: 5,6-Dichloro-1-β-D-Ribofuranosylbenzimidazole Sensitivity-inducing Factor; EAF: ELL-Associated Factor; ELL: Eleven-nineteen Lysine-rich in Leukemia; ESC: Embryonic Stem Cell; GRO-seq: Global Run-On Sequencing; H3K9me3: Trimethylated Lysine 4 of Histone H3; HIC: Human I-mfa Domain Containing Protein; HIF1A: Hypoxia-Inducible Factor 1A; HIV: Human Immunodeficiency Virus; HMBA: Hexamethylene Bisacetamide; JMJD6: Jumonji C-Domain-Containing Protein 6; NF-κB: Nuclear Factor-κB; NELF: Negative Elongation Factor; PIC: Pre-Initiation Complex; PPP: Promoter-Proximal Pausing; Pol II: RNA Polymerase II; PRO-seq: Precision Run-On Sequencing; PPI1α: Protein Phosphatase 1α; PP2B: Protein Phosphatase 2B; PPM1G: Protein Phosphatase, Mg<sup>2+</sup>/Mn<sup>2+</sup> dependent, 1G; P-TEFb: Positive Transcription Elongation Factor b; SEC: Super-Elongation Complex; S2P: Phosphorylated Serine 2; SRSF: Serine/Arginine-Rich Splicing Factor; Thr186: Threonine at the 186<sup>th</sup> position; UTR: Untranslated Region

## Introduction

After RNA polymerase II (Pol II) and general transcription factors are assembled to form the pre-initiation complex (PIC) at a promoter, it synthesizes a short mRNA of 30-50 bases and then pauses (promoter-proximal pausing or PPP) [1-6] (Figure 1A). PPP is induced by two mechanisms. First, the PIC is directly inhibited by two protein complexes: 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) sensitivity-inducing factor (DSIF, containing Supt4a and Supt5) and negative elongation factor (NELF, composed of subunits A-E). Second, the promoter of transcription elongation, positive transcription elongation factor b (P-TEFb, comprised of Cdk9 and either CycT1, CycT2, or CycK), is sequestered in the inhibitory complex 7SK small nuclear ribonucleoprotein (7SK snRNP). In 7SK snRNP, the noncoding RNA 7SK serves as a scaffold to assemble four protein subunits (Hexim1, Hexim2, Larp7, and Mepce) around P-TEFb [7]. Various proteins and RNAs release P-TEFb from 7SK

snRNP, activating the Cdk9 kinase. As a result, Cdk9 releases NELF and converts DSIF into an elongation-promoting factor, both via phosphorylation (Figure 1B). Cdk9 also phosphorylates serine 2 (S2P) on the carboxy-terminal heptapeptide repeat domain (CTD) in the largest subunit of Pol II, facilitating the recruitment of RNA-processing machinery and chromatin remodeling factors. In contrast, paused Pol II is characterized by the phosphorylation of serine at the 5<sup>th</sup> position (S5P) on CTD. PPP is highly prevalent; between 10% and 50% of entire genes simultaneously undergo PPP in many cell types, with only around a half of them proceeding to transcription elongation [1,8-10]. As expected from this prevalence, Pol II is released from PPP through many mechanisms in different contexts. This review article will first describe experimental approaches commonly used to study PPP, and then proceed to the mechanisms underlying the release of P-TEFb from 7SK snRNP and the recruitment of P-TEFb to its target genes, two rate-limiting steps towards the release of Pol II from PPP. Finally, we will discuss the regulation of PPP during *Drosophila* development and pluripotent stem cell differentiation, as they are the most widely studied models.

## Genome-wide approaches to examine PPP

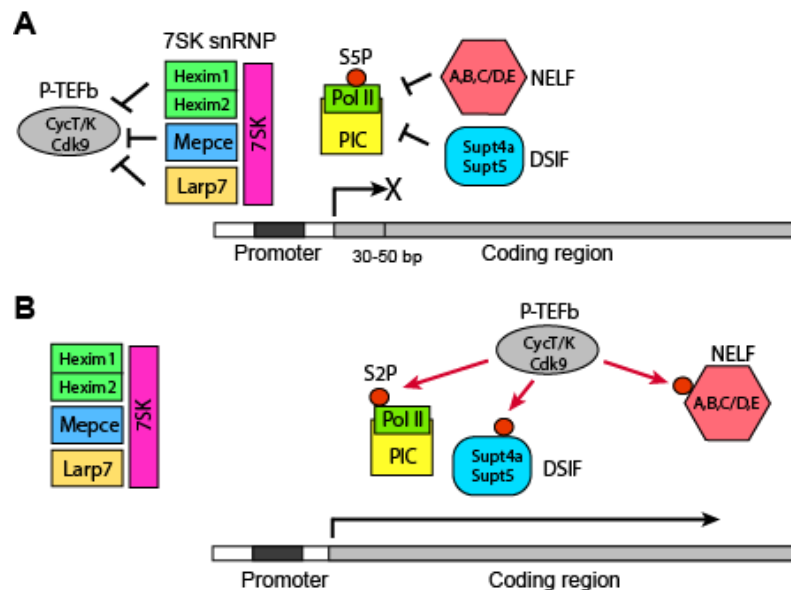
By definition, the study of PPP requires the detection of the distribution pattern of Pol II within a given gene, which is commonly performed with chromatin immunoprecipitation (ChIP) followed by a microarray analysis (ChIP-chip) or next generation sequencing (ChIP-seq) (Table 1). These approaches have the strength of detecting

**\*Corresponding author:** Nobuaki Kikyo, Stem Cell Institute, Department of Genetics, Cell Biology and Development, University of Minnesota, MTRF, 2001 6th St SE, Minneapolis, MN 55455, USA, Tel: 612-624-0498; Fax: 612-624-2436; E-mail: [kikyo001@umn.edu](mailto:kikyo001@umn.edu)

**Received** December 01, 2015; **Accepted** March 07, 2016; **Published** March 14, 2016

**Citation:** Robinson C, Lowe M, Schwartz A, Kikyo N (2016) Mechanisms and Developmental Roles of Promoter-proximal Pausing of RNA Polymerase II. J Stem Cell Res Ther 6: 330. doi:10.4172/2157-7633.1000330

**Copyright:** © 2016 Robinson C, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



**Figure 1:** Key players responsible for promoter-proximal pausing of RNA polymerase II. Paused status (A) and released status (B) of RNA polymerase II are depicted.

Factor	Principles	References
ChIP-chip and ChIP-seq	Antibodies against Pol II, Pol II S2P, and Pol II S5P as well as various modified histones are used.	[8,14,51,52]
GRO-seq	Nascent RNAs that have incorporated 5-bromouridine 5'-triphosphate are isolated with anti-deoxy bromouridine antibody and sequenced.	[9,10,14]
PRO-seq	Nascent RNAs that have incorporated biotinylated NTP are isolated with streptavidin beads for sequencing	[11]
Short-capped RNA analysis	Short mRNAs (<100 bases) with a 5' cap are isolated with anti-cap antibody and sequenced.	[12]

Pol II: RNA Polymerase II; Pol II S2P: RNA Polymerase II with phosphorylated serine at the 2<sup>nd</sup> position in the carboxy-terminal heptapeptide repeat domain; GRO-seq: Global Run-On sequencing; PRO-seq: Precision Run-On sequencing

**Table 1:** Genome-wide approaches commonly used to detect PPP.

not only Pol II and its phosphorylated forms but also several covalent histone modifications characteristic of paused genes, all within the same set of experiments. ChIP-chip studies have demonstrated that the promoters of paused genes are typically enriched with histone marks characteristic of initiated genes, such as trimethylated lysine 4 of histone H3 (H3K4me3) and acetylated H3K9 and H3K14, in addition to Pol II S5P [8]. Additionally, paused genes usually lack markers of transcriptional elongation at the coding regions, including H3K36me3, H3K79me2, and Pol II S2P.

Definitive evidence for paused transcription requires detection of transcript length. Global Run-On Sequencing (GRO-seq) is a genome-wide nuclear run-on assay, in which nascent transcripts are labeled with 5-bromouridine 5'-triphosphate (BrUTP) [9]. The labeled RNAs are then purified using anti-deoxy-BrU beads and sent for next generation sequencing. This approach is particularly useful in determining how specific stimuli release paused Pol II, allowing for elongation at a specific time point. Precision Run-On Sequencing (PRO-seq) sequencing is a modification of GRO-seq, in which reactions incorporate biotinylated NTPs. Incorporation of the first biotinylated NTP inhibits further elongation, allowing for the detection of nascent transcripts at a single-base resolution, as opposed to 30 to 50-base resolution with GRO-seq [11].

Capping of mRNA with 7-methylguanosine takes place at the 5' end during the initial stage of transcription [12]. The cap protects mRNA from exonuclease degradation, allowing for transcription to

complete. Therefore, short mRNAs (<100 bases) with a 5' cap can be isolated and sequenced to identify paused mRNAs using a technique known as short-capped RNA analysis [13].

Among these genome-wide approaches, ChIP-based approaches are more readily available to general cell biology laboratories than mRNA-based approaches, which require extensive optimization. A recent report showed that GRO-seq data are generally consistent with ChIP-seq data obtained using anti-Pol II antibody [14]; current PPP research typically employs the combination between GRO-seq and ChIP-seq of Pol II.

### Release of P-TEFb from 7SK snRNP

About half of P-TEFb in the nucleus is sequestered in 7SK snRNP [15]. The release of P-TEFb from 7SK snRNP represents one of the first steps toward the release Pol II from PPP. A diverse range of factors, including binding proteins, phosphatases, proteases, and RNAs, are involved in this step (Table 2) and are not mutually exclusive. Several representative examples are highlighted below.

Exposure of cells to ultraviolet light or hexamethylene bisacetamide (HMBA) disrupts 7SK snRNP and releases P-TEFb [16]. Both agents activate calcium ion-calmodulin-protein phosphatase 2B (PP2B), which relaxes the structure of 7SK snRNP although the details remain to be studied. The relaxation allows the second phosphatase, protein phosphatase 1α (PP1α), to gain access to the exposed Cdk9,

Factors	Mechanisms	Reference
PP2B and PP1 $\alpha$	Sequentially dephosphorylate 7SK snRNP and Cdk9	[16]
PPM1G	Recruited by NF- $\kappa$ B and Tat, and dephosphorylates Cdk9	[17]
Tat	Changes confirmation of 7SK RNA	[18]
HIC mRNA	Displaces 7SK RNA from 7SK snRNP by competition	[20]
SRSF1 and 2	Binds the exonic-splicing enhancer sequence in nascent mRNA	[21]
Calpain 2	Cleaves Mepce	[22]
JMJD6	Decaps 7SK RNA and destabilizes 7SK snRNP	[25]
AFF1	Promotes Tat's extraction of P-TEFb from 7SK snRNP	[55]
PIP7S	Maintains 7SK snRNP integrity	[56]

PP2B: Protein Phosphatase 2B; PP1 $\alpha$ : Protein Phosphatase 1  $\alpha$ ; 7SK snRNP: 7SK small nuclear ribonucleoprotein; PPM1G: Protein Phosphatase, Mg<sup>2+</sup>/Mn<sup>2+</sup> dependent, 1G; NF- $\kappa$ B: Nuclear Factor- $\kappa$ B; Tat: Tyrosine aminotransferase; HIC: Human I-mfa domain containing protein; SRSF: Serine/Arginine-Rich Splicing Factor; JMJD6: Jumonji C-Domain-Containing Protein 6; AFF1: AF4/FMR2 Family Member 1; P-TEFb: Positive Transcription Elongation Factor b; PIP7S: P-TEFb-Interaction Protein for 7SK Stability

**Table 2:** Factors associated with the release of P-TEFb from 7SK snRNP.

dephosphorylating threonine at the 186<sup>th</sup> position (Thr186) in the T-loop of Cdk9. This, in turn, leads to the release of P-TEFb from 7SK snRNP. The dephosphorylation of Thr186 is induced in a physiological context as well. For example, nuclear factor- $\kappa$ B (NF- $\kappa$ B) and the human immunodeficiency virus (HIV)-1 protein Tat recruit PPM1G (protein phosphatase, Mg<sup>2+</sup>/Mn<sup>2+</sup> dependent, 1G) to promoters, which also dephosphorylates Thr186 [17]. However, the phosphatase-mediated process may not be the only mechanism Tat uses to release P-TEFb from 7SK snRNP. This is due to its ability to bind and release P-TEFb from 7SK snRNP without the use of other factors *in vitro* [18]. Rather, the *in vitro* process is accompanied by a conformational change of 7SK RNA and the release of Hexim1 from 7SK RNP.

The human I-mfa domain containing protein (HIC) interacts with CycT1 to inhibit P-TEFb-dependent transcription [19]. In contrast, HIC mRNA has the opposite effect on transcription. The 3' untranslated region (UTR) of HIC mRNA is necessary and sufficient to activate gene expression at the HIV-1 promoter by displacing 7SK RNA from P-TEFb, allowing it to dissociate from 7SK snRNP [20]. This is explained by a competition between the similar hairpin structures at the 3' end regions of HIC mRNA and 7SK RNA. The opposite functions of HIC protein and HIC mRNA illuminate the complexity of P-TEFb release from 7SK snRNP and the many facets of pause regulation in transcription.

Release of P-TEFb is also directly linked with splicing of nascent mRNA. Serine/arginine-rich splicing factor 1 and 2 (SRSF1 and SRSF2), major components of the splicing machinery, bind 7SK snRNP at active gene promoters [21]. When an exonic-splicing enhancer sequence in nascent mRNA, such as GAAGGA, binds SRSFs, they are released from 7SK snRNP along with P-TEFb, allowing for pausing release. Due to the wide distribution of SRSFs in the genome, this mechanism is likely to be used in many contexts of pause release.

All of the previously described release mechanisms are reversible; however, an irreversible mechanism also exists. During megakaryocyte differentiation, a calcium influx activates the protease calpain 2, causing the cleavage of the Mepce subunit of 7SK snRNP, promoting the release of P-TEFb and upregulating cytoskeleton remodeling genes [22]. GATA1 is known as an upstream transcription factor regulating calpain 2 expression. A truncated mutant GATA1 in megakaryocytic leukemia fails to upregulate calpain 2, revealing the connection between Pol II release, cell differentiation, and leukemogenesis.

### Recruitment of P-TEFb

P-TEFb released from 7SK RNP must to be recruited to the promoters of target genes. This recruitment process is also diverse

(Table 3), frequently using multiple mechanisms within a single recruitment, as exemplified by Brd4 and Mediator summarized below. Recruitment of P-TEFb by c-Myc and MyoD will also be discussed as examples directly relevant to cell differentiation. Note that the release of P-TEFb from 7SK snRNP and the recruitment of the released P-TEFb to promoters are not necessarily sequential steps. These processes are not chronologically indistinguishable in some cases because a single protein, like Brd4, can be involved in the both steps.

Brd4 is a member of the bromodomain and extraterminal domain (BET) family proteins [15,23]. Brd4 is ubiquitously expressed and binds free P-TEFb that is released from 7SK snRNP. Because Brd4 binds acetylated histone H3 and H4 via bromodomains, it was initially proposed to recruit P-TEFb to acetylated promoters. However, it was later found that bromodomains are not required for Brd4 to recruit P-TEFb to promoters [24]. Instead, Brd4 is recruited to promoters by Mediator or other chromatin components. Additionally, Brd4 interacts with jumonji C-domain-containing protein 6 (JMJD6) at enhancers and regulates Pol II release from PPP through long-range interactions between enhancers and promoters [25]. In this context, dual demethylase activity of JMJD6 plays a critical role in transcription elongation. First, protein demethylase activity of JMJD6 is directed toward symmetric dimethylation of H4R3 (H4R3me2s), a marker for gene suppression, and contributes to gene activation. Second, RNA demethylase activity of JMJD6 removes the methyl group at the 5' cap in 7SK RNA (decapping), which destabilizes 7SK snRNP, resulting in the release of P-TEFb. In human embryonic kidney cells HEK293T, more than 1,200 genes are co-regulated by Brd4 and JMJD6, suggesting the prevalence of this mechanism.

P-TEFb is also recruited to promoters by Mediator through multiple mechanisms. Mediator is a protein complex composed of 26 core subunits serving as a physical and functional bridge between Pol II and DNA-bound transcription factors during transcription initiation and elongation [26,27]. The MED26 subunit recruits P-TEFb as a part of the super-elongation complex (SEC) to the promoters of *c-Myc* and *hsp70* [28]. The recruitment of SEC is particularly important because it contains other subunits that promote elongation after pause release, such as the eleven-nineteen lysine-rich in leukemia (ELL) family members and ELL-associated factors (EAFs) 1 and 2. Additionally, the interaction between Mediator and SEC is induced by hypoxia-inducible factor 1A (HIF1A) in hypoxic breast cancer cells [29]. Mediator also uses another subunit MED23 to recruit P-TEFb to the promoter of the serum-responsive gene *Egr1* in mouse embryonic stem cells (ESCs) through a direct interaction with Cdk9 [30].

The oncoprotein c-Myc directly binds Cdk9 and CycT1 at different

Factor	Functions	References
Brd4	Binds free P-TEFb	[15,23]
	Binds Mediator	[24]
	Binds JMJD6 to release P-TEFb	[25]
	Binds Oct4 and regulates pluripotency genes in ESCs	[57]
Mediator (MED26)	Recruits P-TEFb as a part of SEC	[28]
Mediator (MED23)	Binds P-TEFb in ESCs	[30]
HIF1A	Induces the Mediator-SEC interaction	[29]
c-Myc	Directly binds Cdk9 and CycT1	[31,32]
	Binds a third of active genes in ESCs	[33]
MyoD	Binds P-TEFb to activate muscle differentiation genes	[37]
Ikaros	Binds Cdk9 to elongate the $\gamma$ globin gene in erythroid cells	[46]
GATA1	Binds P-TEFb during megakaryocyte differentiation	[47]
TIF1 $\gamma$	Recruits P-TEFb to erythroid genes during erythropoiesis	[48]
Klf4	Recruits P-TEFb during iPSC formation	[52]
NF- $\kappa$ B	Recruits P-TEFb to the IL-8 promoter upon TNF $\alpha$ stimulation	[58]
Cdk8	Binds Mediator and promotes elongation	[59]
STAT3	Binds Cdk9 upon induction by IL-6 in hepatoblastoma cells	[60]
Estrogen receptor $\alpha$	Overcomes pausing at intron 1 of the MYB gene after binding of estrogen	[61]
Androgen receptor	Binds P-TEFb upon stimulation by androgen in prostatic cancer cells	[62]
PPAR $\gamma$	Directly binds Cdk9 during adipogenesis	[63]
Aryl hydrocarbon receptor	Recruits P-TEFb to a cytochrome P450 gene upon binding to aryl hydrocarbon	[64]

Brd4: Bromodomain containing 4; P-TEFb: Positive Transcription Elongation Factor b; JMJD6: Jumonji C-Domain-Containing Protein 6; ESC: Embryonic Stem Cell; SEC: Super-Elongation Complex; HIF1A: Hypoxia-Inducible Factor 1A; CycT1: Cyclin T1; GATA1: GATA Binding Protein 1; TIF1 $\gamma$ : Transcriptional Intermediary Factor 1  $\gamma$ ; Klf4: Kruppel-like factor 4; NF- $\kappa$ B: Nuclear Factor- $\kappa$ B; STAT3: Signal Transducer and Activator 3; IL-6: Interleukin 6; PPAR $\gamma$ : Peroxisome Proliferator Activated Receptor  $\gamma$

**Table 3:** Factors responsible for the recruitment of P-TEFb to promoters.

domains and recruits them to target genes [31,32]. In mouse ESCs, c-Myc binds a third of actively transcribed genes [33]. Treatment of ESCs with a chemical inhibitor of c-Myc or knockdown of c-Myc with shRNA reduces Pol II in the coding regions of the target genes without affecting the level of Pol II at promoters. This provides evidence that c-Myc is responsible for regulating the release of Pol II from pausing. Because c-Myc is one of four genes used to establish induced pluripotent stem cells (iPSCs) [34], its release of Pol II from paused genes may be one of its functions during iPSC formation.

MyoD is a master regulator of muscle differentiation [35]. It interacts with P-TEFb in the proliferating and early differentiation phases of myoblasts [36]. The interaction in the early phase of differentiation brings P-TEFb to MyoD target genes specific to differentiation, such as myogenin and muscle creatine kinase [37]. The functions of the interaction in the proliferating phase remains unclear; however, the ability of Cdk9 to phosphorylate MyoD may regulate its activity independently of pausing release [36].

### Developmental roles of PPP in *Drosophila*

PPP was discovered in 1986 at the heat shock gene *hsp70* in a *Drosophila* cell line [38]. However, it took more than 20 years before genome-wide prevalence of PPP – 10-20% of all *Drosophila* genes – was discovered in 2007 by applying the ChIP-chip technique [39,40]. While a large number of paused genes are related to heat shock and other stimulus-responses, nearly a third of the paused genes play roles in development, such as cell differentiation and cell-cell communication [39]. These results led to the interpretation that PPP acts as a checkpoint to control precisely-timed gene activation during development. This notion is consistent with the initial discovery of PPP at *hsp70*, which is rapidly activated upon a heat stimulus.

Paused developmental genes include those involved in embryonic axis formation and gastrulation in *Drosophila*. The bone morphogenetic protein (BMP) pathway regulates dorsal-ventral patterning and 25 genes

involved in this pathway are paused [41]. The segmentation genes (gap, pair-rule, and segmentation polarity genes), which specify the anterior-posterior axis, are also paused. Among many paused developmental genes, timely release of Pol II is particularly important for the *Snail* gene (*sna*), which is a key regulator for epithelial-mesenchymal transition and mesoderm formation [42,43]. *sna* is expressed in around 1,000 presumptive mesoderm cells and synchronized activation of the gene is essential for coordinated invagination of the cells. When the tightly paused *sna* promoter was swapped with the moderately paused *sog* gene promoter, it resulted in a mixed embryonic population with normal and partial invagination. Replacement with the non-paused *ths* promoter completely abolished invagination in the majority of embryos, and normal invagination was rarely detectable.

Studies with *Drosophila* also provided an unexpected insight into a novel mechanism of PPP. Knockdown of NELF was supposed to primarily upregulate gene expression by releasing Pol II; however, it upregulated 71 genes and downregulated 170 genes out of 18,500 genes in total [44]. The downregulated gene promoters displayed not only decreased Pol II, as expected by the lack of NELF, but also increased nucleosome occupancy, a potential mechanism for the observed downregulation [44,45]. This finding suggested that Pol II and nucleosomes compete against each other for occupancy at the promoter regions in the downregulated genes. Because NELF knockdown also increased many genes, there must be additional regulatory mechanisms that dictate the final outcome of the gene expression. In addition, a recent report showed that pausing is not a simple on/off state; 90% of transcriptionally active genes (7,036 out of 7,734) are partially paused in *Drosophila* embryos [41]. These findings indicate the presence of multiple layers of fine tuning mechanisms that await further study.

### Roles of PPP in pluripotent stem cells

Pluripotent stem cells and hematopoietic cells have been the main cell types used for the study of the developmental role of PPP

in mammalian cells. The focus of this section will primarily be upon pluripotent stem cells due to the more in depth mechanistic and genome-wide studies compared to hematopoietic cells [22,46-48]. In fact, the first genome-wide distribution of PPP in mammalian cells was described in human ESCs [8]. The occurrence of PPP in ESCs has been analyzed using GRO-seq and ChIP-seq of Pol II in mouse ESCs as well as ChIP-chip of Pol II in human ESCs (Table 4).

In contrast to the initial findings in *Drosophila* of a direct link between PPP and the regulation of development, there appears to be no definitive link in mammalian cells. The number of RefSeq genes paused in mouse ESCs and MEFs, 39% and 34% respectively, is quite similar [10]. In both cell types, paused genes are primarily from the cell cycle, catabolism, and DNA damage response categories as shown by a gene ontology analysis of the GRO-seq data. Pol II was not detected at the pluripotency genes in MEFs, consistent with the lack of expression. Unexpectedly, the pluripotency genes of ESCs show significant enrichment of Pol II within the promoters compared to the coding regions, suggesting that at least partial PPP does occur at these loci. The meaning behind the partial PPP requires further study to elucidate its role in the maintenance of pluripotency. The prevalence of PPP in human cells is also similar between pluripotent cells and differentiated cells as observed with ChIP-chip of Pol II-S2P in hepatocytes (36% of protein-coding genes), B cells (42%), and ESCs (52%) [8]. The categories of paused genes have yet to be extensively characterized in human pluripotent stem cells.

Mouse ESCs have traditionally been cultured in the presence of serum; however, a serum-free condition called the 2i system, which uses two kinase inhibitors, has seen widespread use as a more chemically-defined culture system recently [49]. After being cultured with serum, ESCs display leaky expression of many lineage-specific genes and heterogeneous expression of pluripotency genes [50]. Unexpectedly, cells cultured under the 2i condition demonstrate lower expression of lineage-specific genes as well as a more homogenous expression of pluripotency genes than the cells grown in serum-based media [49,51]. This is not due to a lack of PPP in ESCs cultured in the presence of serum. In a comparative GRO-seq analysis of PPP in the cells cultured with serum and the 2i, there was no significant difference in the gene families that were paused between the two conditions [14]. The genes that showed the highest levels of PPP in both culture conditions were those involved in signal transduction and the regulation of cell cycle, not differentiation or pluripotency. This suggests that PPP does not play a significant role in the regulation of developmental genes in ESCs.

In a more functional sense, the manipulation of PPP during reprogramming of MEFs to iPSCs can provide a new approach to improve the quality and efficiency of cell reprogramming. During reprogramming there are two waves of gene activation, the first being genes related to proliferation, metabolism, and cytoskeleton whereas the second is primarily the pluripotency genes and their regulators [52]. During reprogramming there is significant PPP in 55% of the second wave genes, most notably Oct4 [52]. The role of PPP in reprogramming was further studied by knocking down several key genes involved in the sequestration and release of P-TEFb. The elimination of PPP via downregulation of Hexim1 or upregulation of Brd4 accelerates the formation of iPSCs as well as increasing the number of iPSC colonies [52]. Alternatively, the upregulation of Hexim1 or downregulation of Brd4 leads to an arrest in reprogramming prior to the second wave. Beyond the factors directly involved in the PPP machinery, the pluripotency factor Klf4 recruits Cdk9 to pluripotency genes during reprogramming as well as in ESCs [52]. Knockdown of Klf4 leads to a significant decrease of Cdk9 loading on the pluripotency genes and their expression. Additionally, a majority of the genes involved in the late stage of reprogramming are targets of Klf4, suggesting a major role for PPP release in the formation of iPSCs. Beyond iPSC formation, by tightly controlling the expression of factors linked to the PPP maintenance and release, it may be possible to more accurately reprogram cells into a wide range of lineages, such as the conversion of fibroblasts to neural cells [53-64].

## Future Directions

Following its discovery, the study PPP has resulted in as many new questions as it has answers. As technology continues to improve, our genome-wide picture of PPP becomes more complex and diverse. The major protein and RNA players have been uncovered; however, our understanding of the mechanisms behind the release and recruitment of those molecules is still incomplete. While the developmental roles for pausing have become more apparent in *Drosophila* and pluripotent stem cells, very little is known about the roles of PPP in the differentiation of many other mammalian lineages. HIF1A-mediated pause release, mentioned above, is evidence that PPP is also involved in cancer and potentially other pathological conditions. In cardiomyocytes, PPP is present at metabolism-related genes, controlling energy homeostasis [54]. These studies help outline the versatility of PPP as a regulatory mechanism whose potential has yet to be understood in its entirety.

Cell Types	Methods of detection	Paused percentage	Categories of paused genes	Pluripotency relation	Reference
Mouse ESCs cultured with serum	GRO-seq	39%	Pluripotency, DNA damage response, translation, cell cycle, and catabolism	Pol II is not significantly paused at developmental regulators. Pluripotency genes are partially paused.	[10]
MEFs	GRO-seq	34%	DNA damage response, translation, cell cycle, and catabolism	Not Discussed	[10]
Mouse ESCs cultured with 2i	GRO-seq and ChIP-seq of Pol II	Not discussed	Cell cycle, signal transduction, and cell proliferation	Pol II is not significantly paused at developmental regulators.	[14]
Mouse ESCs cultured with 2i	ChIP-seq of Pol II	Not discussed	Not Discussed	A general increase in PPP after growth in cells cultured with 2i than with serum.	[51]
Mouse iPSCs	ChIP-seq of Pol II	Not discussed	Oct4 in the later stage of reprogramming	Klf4 recruits Cdk9 to the pluripotency promoters.	[52]
Human ESCs cultured with serum	ChIP-chip of Pol II-S5P	52%	Not discussed	Not discussed	[8]
Human hepatocytes	ChIP-chip of Pol II	36%	Not discussed	Not discussed	[8]
Human B cells	ChIP-chip of Pol II	42%	Not discussed	Not discussed	[8]

**Table 4:** A Comparison of PPP analyses in various cell types.

## Acknowledgements

C. R. was supported by the NIH (T32 GM113846). N.K. was supported by the NIH (R01 GM098294, R21 CA187232, and R21 AR066158); Engdahl Family Foundation; Adjacent Possible Grant from the Stem Cell Institute, University of Minnesota; and Grain-in-Aid of Research, Artistry and Scholarship, University of Minnesota (22802).

## References

- Levine M (2011) Paused RNA polymerase II as a developmental checkpoint. *Cell* 145: 502-511. [PubMed]
- Adelman K, Lis JT (2012) Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* 13: 720-731. [PubMed]
- Guo J, Price DH (2013) RNA polymerase II transcription elongation control. *Chem Rev* 113: 8583-8603. [PubMed]
- Zhou Q, Li T, Price DH (2012) RNA polymerase II elongation control. *Annu Rev Biochem* 81: 119-143. [PubMed]
- Jonkers I, Lis JT (2015) Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* 16: 167-177. [PubMed]
- Liu X, Kraus WL, Bai X (2015) Ready, pause, go: regulation of RNA polymerase II pausing and release by cellular signaling pathways. *Trends Biochem Sci* 40: 516-525. [PubMed]
- Lenasi T, Barboric M (2010) P-TEFb stimulates transcription elongation and pre-mRNA splicing through multilateral mechanisms. *RNA Biol* 7: 145-150. [PubMed]
- Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130: 77-88. [PubMed]
- Core LJ, Waterfall JJ, Lis JT (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322: 1845-1848. [PubMed]
- Min IM, Waterfall JJ, Core LJ, Munroe RJ, Schimenti J, et al. (2011) Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev* 25: 742-754. [PubMed]
- Kwak H, Fuda NJ, Core LJ, Lis JT (2013) Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* 339: 950-953. [PubMed]
- Moteki S, Price D (2002) Functional coupling of capping and transcription of mRNA. *Mol Cell* 10: 599-609. [PubMed]
- Nechaev S, Fargo DC, dos Santos G, Liu L, Gao Y, et al. (2010) Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* 327: 335-338. [PubMed]
- Williams LH, Fromm G, Gokey NG, Henriques T, Muse GW, et al. (2015) Pausing of RNA polymerase II regulates mammalian developmental potential through control of signaling networks. *Mol Cell* 58: 311-322. [PubMed]
- Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN, et al. (2005) The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* 19: 523-534. [PubMed]
- Chen R, Liu M, Li H, Xue Y, Ramey WN, et al. (2008) PP2B and PP1alpha cooperatively disrupt 7SK snRNP to release P-TEFb for transcription in response to Ca<sup>2+</sup> signaling. *Genes Dev* 22: 1356-1368. [PubMed]
- McNamara RP, McCann JL, Gudipaty SA, D'Orso I (2013) Transcription factors mediate the enzymatic disassembly of promoter-bound 7SK snRNP to locally recruit P-TEFb for transcription elongation. *Cell Rep* 5: 1256-1268. [PubMed]
- Krueger BJ, Varzavand K, Cooper JJ, Price DH (2010) The mechanism of release of P-TEFb and HEXIM1 from the 7SK snRNP by viral and cellular activators includes a conformational change in 7SK. *PLoS One* 5: e12335. [PubMed]
- Young TM, Wang Q, Pe'ery T, Mathews MB (2003) The human I-mfa domain-containing protein, HIC, interacts with cyclin T1 and modulates P-TEFb-dependent transcription. *Mol Cell Biol* 23: 6373-6384. [PubMed]
- Young TM, Tsai M, Tian B, Mathews MB, Pe'ery T (2007) Cellular mRNA activates transcription elongation by displacing 7SK RNA. *PLoS One* 2: e1010. [PubMed]
- Ji X, Zhou Y, Pandit S, Huang J, Li H, et al. (2013) SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell* 153: 855-868. [PubMed]
- Elagib KE, Rubinstein JD, Delehanty LL, Ngho VS, Greer PA, et al. (2013) Calpain 2 activation of P-TEFb drives megakaryocyte morphogenesis and is disrupted by leukemogenic GATA1 mutation. *Dev Cell* 27: 607-620. [PubMed]
- Yang Z, Yik JH, Chen R, He N, Jang MK, et al. (2005) Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* 19: 535-545. [PubMed]
- Kanno T, Kanno Y, LeRoy G, Campos E, Sun HW, et al. (2014) BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones. *Nat Struct Mol Biol* 21: 1047-1057. [PubMed]
- Liu W, Ma Q, Wong K, Li W, Ohgi K, et al. (2013) Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. *Cell* 155: 1581-1595. [PubMed]
- Yin JW, Wang G (2014) The Mediator complex: a master coordinator of transcription and cell lineage development. *Development* 141: 977-987. [PubMed]
- Allen BL, Taatjes DJ (2015) The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol* 16: 155-166. [PubMed]
- Takahashi H, Parmely TJ, Sato S, Tomomori-Sato C, Banks CA, et al. (2011) Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* 146: 92-104. [PubMed]
- Galbraith MD, Allen MA, Bensard CL, Wang X, Schwinn MK, et al. (2013) HIF1A employs CDK8-mediator to stimulate RNAPII elongation in response to hypoxia. *Cell* 153: 1327-1339. [PubMed]
- Wang W, Yao X, Huang Y, Hu X, Liu R, et al. (2013) Mediator MED23 regulates basal transcription in vivo via an interaction with P-TEFb. *Transcription* 4: 39-51. [PubMed]
- Eberhardy SR, Farnham PJ (2002) Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter. *J Biol Chem* 277: 40156-40162. [PubMed]
- Kanazawa S, Soucek L, Evan G, Okamoto T, Peterlin BM (2003) c-Myc recruits P-TEFb for transcription, cellular proliferation and apoptosis. *Oncogene* 22: 5707-5711. [PubMed]
- Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, et al. (2010) c-Myc regulates transcriptional pause release. *Cell* 141: 432-445. [PubMed]
- Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126: 663-676. [PubMed]
- Berkes CA, Tapscott SJ (2005) MyoD and the transcriptional control of myogenesis. *Semin Cell Dev Biol* 16: 585-595. [PubMed]
- Simone C, Stiegler P, Bagella L, Pucci B, Bellan C, et al. (2002) Activation of MyoD-dependent transcription by cdk9/cyclin T2. *Oncogene* 21: 4137-4148. [PubMed]
- Giacinti C, Bagella L, Puri PL, Giordano A, Simone C (2006) MyoD recruits the cdk9/cyclin T2 complex on myogenic-genes regulatory regions. *J Cell Physiol* 206: 807-813. [PubMed]
- Gilmour DS, Lis JT (1986) RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in *Drosophila melanogaster* cells. *Mol Cell Biol* 6: 3984-3989. [PubMed]
- Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, et al. (2007) RNA polymerase is poised for activation across the genome. *Nat Genet* 39: 1507-1511. [PubMed]
- Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, et al. (2007) RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat Genet* 39: 1512-1516. [PubMed]
- Saunders A, Core LJ, Sutcliffe C, Lis JT, Ashe HL (2013) Extensive polymerase pausing during *Drosophila* axis patterning enables high-level and pliable transcription. *Genes Dev* 27: 1146-1158. [PubMed]
- Lagha M, Bothma JP, Esposito E, Ng S, Stefanik L, et al. (2013) Paused Pol II coordinates tissue morphogenesis in the *Drosophila* embryo. *Cell* 153: 976-987. [PubMed]

43. Leptin M (2005) Gastrulation movements: the logic and the nuts and bolts. *Dev Cell* 8: 305-320. [[PubMed](#)]
44. Gilchrist DA, Nechaev S, Lee C, Ghosh SK, Collins JB, et al. (2008) NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev* 22: 1921-1933. [[PubMed](#)]
45. Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao Y, et al. (2010) Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* 143: 540-551. [[PubMed](#)]
46. Bottardi S, Zmiri FA, Bourgoin V, Ross J, Mavoungou L, et al. (2011) Ikaros interacts with P-TEFb and cooperates with GATA-1 to enhance transcription elongation. *Nucleic Acids Res* 39: 3505-3519. [[PubMed](#)]
47. Elagib KE, Mihaylov IS, Delehanty LL, Bullock GC, Ouma KD, et al. (2008) Cross-talk of GATA-1 and P-TEFb in megakaryocyte differentiation. *Blood* 112: 4884-4894. [[PubMed](#)]
48. Bai X, Kim J, Yang Z, Jurynek MJ, Akie TE, et al. (2010) TIF1gamma controls erythroid cell fate by regulating transcription elongation. *Cell* 142: 133-143. [[PubMed](#)]
49. Wray J, Kalkan T, Smith AG (2010) The ground state of pluripotency. *Biochem Soc Trans* 38: 1027-1032. [[PubMed](#)]
50. Kobayashi H, Kikyo N (2015) Epigenetic regulation of open chromatin in pluripotent stem cells. *Transl Res* 165: 18-27. [[PubMed](#)]
51. Marks H, Kalkan T, Menafra R, Denissov S, Jones K, et al. (2012) The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* 149: 590-604. [[PubMed](#)]
52. Liu L, Xu Y, He M, Zhang M, Cui F, et al. (2014) Transcriptional pause release is a rate-limiting step for somatic cell reprogramming. *Cell Stem Cell* 15: 574-588. [[PubMed](#)]
53. Ang CE, Wernig M (2014) Induced neuronal reprogramming. *J Comp Neurol* 522: 2877-2886. [[PubMed](#)]
54. Pan H, Qin K, Guo Z, Ma Y, April C, et al. (2014) Negative elongation factor controls energy homeostasis in cardiomyocytes. *Cell Rep* 7: 79-85. [[PubMed](#)]
55. Lu H, Li Z, Xue Y, Schulze-Gahmen U, Johnson JR, et al. (2014) AFF1 is a ubiquitous P-TEFb partner to enable Tat extraction of P-TEFb from 7SK snRNP and formation of SECs for HIV transactivation. *Proc Natl Acad Sci USA* 111: E15-24. [[PubMed](#)]
56. He N, Jahchan NS, Hong E, Li Q, Bayfield MA, et al. (2008) A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol Cell* 29: 588-599. [[PubMed](#)]
57. Wu T, Pinto HB, Kamikawa YF, Donohoe ME (2015) The BET family member BRD4 interacts with OCT4 and regulates pluripotency gene expression. *Stem Cell Reports* 4: 390-403. [[PubMed](#)]
58. Barboric M, Nissen RM, Kanazawa S, Jabrane-Ferrat N, Peterlin BM (2001) NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol Cell* 8: 327-337. [[PubMed](#)]
59. Donner AJ, Ebmeier CC, Taatjes DJ, Espinosa JM (2010) CDK8 is a positive regulator of transcriptional elongation within the serum response network. *Nat Struct Mol Biol* 17: 194-201. [[PubMed](#)]
60. Hou T, Ray S, Brasier AR (2007) The functional role of an interleukin 6-inducible CDK9-STAT3 complex in human gamma-fibrinogen gene expression. *J Biol Chem* 282: 37091-37102. [[PubMed](#)]
61. Mitra P, Pereira LA, Drabsch Y, Ramsay RG, Gonda TJ (2012) Estrogen receptor-alpha recruits P-TEFb to overcome transcriptional pausing in intron 1 of the MYB gene. *Nucleic Acids Res* 40: 5988-6000. [[PubMed](#)]
62. Lee DK, Duan HO, Chang C (2001) Androgen receptor interacts with the positive elongation factor P-TEFb and enhances the efficiency of transcriptional elongation. *J Biol Chem* 276: 9978-9984. [[PubMed](#)]
63. Iankova I, Petersen RK, Annicotte JS, Chavey C, Hansen JB, et al. (2006) Peroxisome proliferator-activated receptor gamma recruits the positive transcription elongation factor b complex to activate transcription and promote adipogenesis. *Mol Endocrinol* 20: 1494-1505. [[PubMed](#)]
64. Tian Y, Ke S, Chen M, Sheng T (2003) Interactions between the aryl hydrocarbon receptor and P-TEFb. Sequential recruitment of transcription factors and differential phosphorylation of C-terminal domain of RNA polymerase II at cyp1a1 promoter. *J Biol Chem* 278: 44041-44048. [[PubMed](#)]