

PAP1 and PAP7 are required for association of plastid-encoded RNA polymerase with DNA

Joyful Wang¹, V. Miguel Palomar^{1, 2}, Ji-Hee Min¹, and Andrzej T. Wierzbicki¹

¹Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, MI 48109, USA

²Departamento de Bioquímica, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad de México, 04510, México

Correspondence: wierzbic@umich.edu

Keywords: *Arabidopsis thaliana*, chloroplast, transcription, RNA polymerase

ABSTRACT

Plastid-encoded RNA polymerase (PEP) is a bacterial-type multisubunit RNA polymerase responsible for the majority of transcription in chloroplasts. PEP consists of four core subunits, which are orthologs of their cyanobacterial counterparts. In *Arabidopsis thaliana*, PEP is expected to associate with 14 PEP-associated proteins (PAPs), which serve as peripheral subunits of the RNA polymerase. The exact contributions of PAPs to PEP function are still poorly understood. We used ptChIP-seq to show that PAP1 (also known as pTAC3), a peripheral subunit of PEP, binds to the same genomic loci as RpoB, a core subunit of PEP. The *pap1* mutant shows a complete loss of RpoB binding to DNA throughout the genome, indicating that PAP1 is necessary for RpoB binding to DNA. A similar loss of RpoB binding to DNA is observed in a mutant defective in PAP7 (also known as pTAC14), another peripheral PEP subunit. We propose that PAPs are required for the recruitment of core PEP subunits to DNA.

KEY MESSAGE

The peripheral subunits of plastid-encoded RNA polymerase play a crucial role in recruiting the core PEP subunits to DNA in *Arabidopsis* chloroplasts.

INTRODUCTION

The chloroplast genome is transcribed by two classes of RNA polymerases: nuclear-encoded RNA polymerase (NEP) and plastid-encoded RNA polymerase (PEP). NEP is a single subunit enzyme similar to T3/T7 phage RNA polymerases, which is encoded in the nuclear genome, translated by cytosolic ribosomes, and imported into plastids. NEP is primarily active in the early stages of chloroplast development. PEP, on the other hand, is a multi-subunit protein complex similar to bacterial RNA polymerases. The chloroplast genome encodes the core subunits of PEP, which is responsible for the majority of transcription in mature chloroplasts (Pfannschmidt et al. 2015).

The core PEP enzyme is composed of two α -subunits (RpoA), the catalytic β -subunit (RpoB), a β' -subunit (RpoC1) and a β'' -subunit (RpoC2) (Pfannschmidt et al. 2015). The PEP complex also interacts with nuclear-encoded sigma factors, which are responsible for promoter recognition and sequence-specific initiation of transcription (Chi et al. 2015). Although the PEP complex has structural similarities to bacterial RNA polymerases, it can only be detected in its

core form in non-photosynthetic plastids (Pfannschmidt and Link 1994). In chloroplasts, PEP is present in a much larger protein complex that contains several additional subunits (Pfalz and Pfannschmidt 2013).

Peripheral subunits of PEP are also referred to as PEP-associated proteins (PAPs). They have been identified primarily through their physical interaction with the PEP complex (Pfannschmidt et al. 2000; Suzuki et al. 2004; Pfalz et al. 2006; Steiner et al. 2011). *Arabidopsis thaliana* is expected to have 14 PAPs, which are of eukaryotic origin and are encoded in the nuclear genome. The molecular architecture of the PEP-PAP complex has been recently determined using cryo-electron microscopy. However, functions of individual PAPs remain only partially understood (Ruedas et al. 2022; Vergara-Cruces et al. 2024; Wu et al. 2024; do Prado et al. 2024).

PAP proteins are necessary for the proper expression of chloroplast-encoded genes and the development of photosynthetic chloroplasts (Pfalz et al. 2006; Garcia et al. 2008; Arsova et al. 2010; Yagi et al. 2012; Gilkerson et al. 2012; Grübler et al. 2017; Yu et al. 2018; Liebers et al. 2020). The only exceptions are PAP4/FSD3 and PAP9/FSD2, which have weaker phenotypes due to partial redundancy (Myouga et al. 2008). It has been proposed that all PAPs are necessary for the proper functioning of PEP, which is consistent with their involvement in the PEP-PAP complex (Ruedas et al. 2022; Vergara-Cruces et al. 2024; Wu et al. 2024; do Prado et al. 2024) and explains why most PAP mutants exhibit similar non-photosynthetic phenotypes (Pfannschmidt et al. 2015).

The proposed participation of all PAPs in the PEP-PAP complex predicts that PAPs should co-localize with core PEP subunits throughout the chloroplast genome. This has been partially demonstrated in wheat for PAP1 on a few individual loci (Yagi et al. 2012). It was found that both RpoA and PAP1 bind to the promoters of PEP-transcribed genes *psbA*, *rbcL*, *psaA*, *rrn*, *psbD* and *trnE* (Yagi et al. 2012). On *psaA* and *rrn*, binding signal levels of RpoA and PAP1 were very similar while on the remaining tested loci, PAP1 binding was substantially stronger than binding of RpoA (Yagi et al. 2012). It is currently unknown if PAP1 overlaps with the binding of core PEP subunits throughout the rest of the chloroplast genome.

The structure of the PEP-PAP complex (Ruedas et al. 2022; Vergara-Cruces et al. 2024; Wu et al. 2024; do Prado et al. 2024) predicts that in the absence of the complete set of PAPs, PEP should be unable to transcribe. This explains why PAPs are required for proper expression

of plastid genes. However, the core PEP complex may retain some transcriptional activity (Pfannschmidt and Link 1994), and it is unknown if PAPs are required for all PEP activity, for only some aspects of its function, or for some posttranscriptional processes, which are known to play a significant role in chloroplast gene regulation (Barkan 2011).

To determine the impact of PAPs on transcription, we examined the DNA binding patterns of PAP1 and RpoB. Our analysis revealed that PAP1 binds to the same genomic regions as RpoB with comparable intensities. We subsequently investigated whether PAP1 is necessary for RpoB binding to DNA. The *pap1* mutant exhibited no detectable RpoB binding to DNA, indicating that PAP1 is essential for the recruitment of PEP to its target genes. We also tested whether another peripheral PEP subunit, PAP7, is necessary for PEP binding to DNA. Similar to the impact of PAP1, the *pap7* mutant lost all detectable RpoB binding to DNA. This suggests that the requirement for PEP binding to DNA may be a more general property of PAPs.

MATERIALS AND METHODS

Plant materials and growth conditions

Wild-type *Arabidopsis thaliana* ecotype Col-0 was used in all analyses. We used the following mutant genotypes: *ptac3/pap1* (Salk_108852) (Alonso et al. 2003) and *ptac14/pap7* (SAIL_566_F06) (McElver et al. 2001). Experiments were performed on 5-day-old plants. Seeds were first stratified in darkness at 4°C for 48 hours and grown on 0.5X MS plates (0.215% MS salts, 0.05% MES-KOH pH 5.7, 1% sucrose, 0.65% agar) for 5 days at 22°C under constant white LED light (50 $\mu\text{mol m}^{-2} \text{sec}^{-1}$).

Chloroplast crosslinking

As previously described (Palomar et al. 2022), whole 5-day-old seedlings were vacuum-infiltrated with 4% formaldehyde for 10 minutes and incubated in darkness for 4 hours at 4°C. Formaldehyde was quenched by adding 2 M glycine to 125 mM final concentration and vacuum infiltrating for 5 minutes.

ptChIP-seq

Crosslinked whole 5-day-old seedlings were homogenized in ice-cold chloroplast lysis buffer (50 mM Tris-HCl [pH 8.0], 10 mM EDTA, 1% SDS) then filtered through 2 layers of Miracloth by centrifuging at 1500 x g for 10 sec to remove debris. Samples were sonicated to achieve DNA fragments ranging from 200 nt to 300 nt using a QSonica Q700 sonicator. The fragmented samples were incubated overnight with 5 µg of polyclonal anti-RpoB antibody (PhytoAB, San Jose, CA, USA; catalog number PHY1239) or anti-PAP1/pTAC3 antibody (PhytoAB, San Jose, CA, USA; catalog number PHY0391A) and with 60 µL Protein A Dynabeads (Thermo Fisher Scientific, Waltham, MA, USA; catalog number 10002D). After incubation, the beads were washed, and DNA was eluted and reverse crosslinked as described (Rowley et al. 2013). High-throughput sequencing libraries were prepared as reported (Bowman et al. 2013) and sequenced using an Illumina NovaSeq 6000 S4 flow-cell with 150 x 150 paired-end configuration at the University of Michigan Advanced Genomics Core.

Data analysis

The obtained raw sequencing reads were trimmed using trim_galore v.0.6.7 with cutadapt v.3.5 (Martin 2011) and mapped to the TAIR10 Arabidopsis plastid genome (www.arabidopsis.org) using Bowtie2 v.2.4.4 (Langmead and Salzberg 2012). Numbers of reads and accession numbers of all high throughput sequencing datasets are shown in Table 1. Read counts on defined genomic regions were determined using samtools v.1.13 (Danecek et al. 2021) and bedtools v.2.30.0 (Quinlan and Hall 2010). ptChIP-seq signals on annotated genes were calculated by dividing reads per million (RPM)-normalized read counts from anti-RpoB or anti-PAP1 ptChIP-seq by RPM-normalized read counts from input samples. ptChIP-seq enrichments on annotated genes were calculated by dividing signal levels on individual genes by the median signal level on genes in the *rpoB* operon, which is not transcribed by PEP and represents background signal levels. ptChIP-seq enrichments on genomic bins were calculated by dividing signal levels on individual bins by the signal level on the entire *rpoB* operon.

Isolation and fractionation of chloroplasts

To isolate chloroplasts, 5-day-old seedlings of Col-0 were homogenized with 1x isolation buffer (330 mM Sorbitol, 30 mM HEPES, pH 7.5, β-mercaptoethanol, and protease inhibitor) and filtered through 2 layers of Miracloth by centrifuging at 1500 x g for 5 min. The pellet was

washed with the 1x isolation buffer once, resuspended with resuspension buffer (50 mM Tris-HCl pH 8.0, 5 mM CaCl₂, β -mercaptoethanol, and protease inhibitor), and incubated on ice for 10 min to burst chloroplasts (chloroplast extracts). The sample was centrifuged at 21,000 g for 5 min at 4°C. The soluble fraction was recovered as a stroma fraction, and the pellet (membrane fraction) was washed three times in resuspension buffer and finally resuspended in resuspension buffer.

Immunoblot analysis

To detect RpoB, LHCB1, and RbcL in chloroplast extract, stroma, and membrane fractions, each fractionated sample was mixed with 2x SDS loading buffer (125 mM Tris-HCl, pH 6.8, 2% SDS, 0.05% Bromophenol blue, 20% glycerol, 200 mM β -mercaptoethanol). Anti-RpoB antibody (PhytoAB catalog numbers PHY 1239 and PHY1700), LHCB1 (Agrisera AS01 004), RbcL (PHY0096A), anti-rabbit IgG antibody conjugated with horseradish peroxidase (Cell Signaling catalog number 7074S), and anti-mouse IgG antibody conjugated with horseradish peroxidase (Cell Signaling catalog number 7076S) were used.

To detect RpoB, RpoC1, pTAC3/PAP1 and Actin in Col-0 wild-type, *pap1*, and *pap7* mutants, total proteins were extracted by 2x SDS loading buffer (125 mM Tris-HCl, pH 6.8, 2% SDS, 0.05% Bromophenol blue, 20% glycerol, 200 mM β -mercaptoethanol). Anti-RpoB antibody (PhytoAB catalog numbers PHY 1239 and PHY1700), anti-RpoC1 antibody (PhytoAB catalog number PHY0381A), anti-pTAC3/PAP1 antibody (PhytoAB catalog number PHY0391A), anti-Actin antibody (Agrisera catalog number AS13 2640), and anti-rabbit IgG antibody conjugated with horseradish peroxidase (Cell Signaling catalog number 7074) were used. Protein bands were visualized using chemiluminescence reagents (SuperSignal West Femto Maximum Sensitivity Substrate, Thermo Scientific) and a ChemiDoc Imaging System (Bio-Rad).

RESULTS

PAP1 and RpoB bind the same DNA sequences

To determine the genome-wide binding pattern of PAP1, we performed ptChIP-seq with an anti-PAP1 antibody. We performed three biological replicates of the assay using Col-0 wild-type

plants and used a *pap1* knock-out mutant as a negative control. PAP1 binding to DNA had a complex and locus-specific pattern with a strong enrichment on rRNA genes in the inverted repeats (IR) and on several individual loci in both large single copy (LSC) and small single copy (SSC) regions (Fig. 1A). Anti-PAP1 ptChIP-seq signal was not detectable in the *pap1* mutant negative control, which confirms the specificity of the antibody (Fig. 1A). These and other high throughput sequencing data obtained in this study may be accessed in the Plastid Genome Visualization Tool (Plavisto) at <http://plavisto.mecdb.lsa.umich.edu>.

To test if PAP1 binds the same genomic regions as PEP core subunits, we compared the binding pattern of PAP1 to the binding pattern of RpoB, which was detected using three biological replicates of ptChIP-seq with anti-RpoB antibody in Col-0 wild-type seedlings (Fig. 1A). Specificity of the anti-RpoB antibody in ptChIP has been previously established (Palomar et al. 2022). A western blot demonstrating the specificity of the anti-RpoB antibody in total chloroplast extracts as well as stromal and membrane fractions is shown in Fig. 2. The patterns of PAP1 and RpoB binding were remarkably similar throughout the entire chloroplast genome (Fig. 1A). To quantify the correlation between PAP1 and RpoB binding, we performed regression analysis using ptChIP enrichment data calculated for coding sequences of all annotated genes, which demonstrated a very strong ($R^2 = 0.99$) and significant correlation (Fig 1B). Similar results were obtained by comparing ptChIP enrichments in 250 bp genomic bins covering the entire genome (Fig. 1C). This indicates that PAP1 binds the same regions as RpoB.

We further tested if PAP1 binding follows PEP core subunits in preferential binding to specific elements of individual protein-coding genes, particularly binding to gene promoters (Palomar et al. 2022). PAP1 binding was enriched on all analyzed gene promoters (Fig. 1D-H). On *psbA*, *psbEFLJ*, and *rbcL* genes, PAP1 binding closely followed binding of RpoB (Fig. 1D-F). However, on *psaA* and *psbB* promoters, PAP1 enrichment signal was substantially stronger than the RpoB signal, which indicates that locus-specific differences between PAP1 and RpoB binding within individual loci are possible. Overall, these results indicate that PAP1 binds the same genomic regions as core subunits of PEP. This is consistent with PAP1 working as an accessory subunit of PEP throughout the entire chloroplast genome.

PAP1 is required for RpoB association with DNA

PAP1 has been proposed to work as an accessory subunit of PEP and have an impact on the accumulation of *psaA*, *psbA* and *rbcL* mRNAs in wheat and rice (Yagi et al. 2012; Wang et al. 2016). However, the impact of PAP1 on PEP transcription and especially on the recruitment of core subunits of PEP to transcribed genes remains unknown. To test if PAP1 is required for PEP binding to DNA, we performed ptChIP-seq with the anti-RpoB antibody in Col-0 wild-type and *pap1* knock-out mutant plants. While RpoB binding was detected at the expected loci (Palomar et al. 2022) in Col-0 wild-type, no substantial binding of RpoB to DNA was detected in the *pap1* mutant (Fig. 3A). This indicates that PAP1 is required for association of RpoB with DNA throughout the entire chloroplast genome.

RpoB binding to DNA was lost in the *pap1* mutant throughout the entire lengths of the analyzed genes, including *psbA*, *psbEFLJ*, *psaA* and *rbcL* (Fig. 3B-E). This includes the loss of RpoB binding to gene promoters where RpoB is normally enriched (Palomar et al. 2022). This is consistent with PAP1 being required for binding of PEP to both gene promoters and transcribed sequences.

Regression analysis further supports the genome-wide loss of RpoB binding to DNA in the *pap1* mutant on annotated genes (Fig. 3F) and on 250 bp bins distributed throughout the entire genome (Fig. 3G). Small residual RpoB ptChIP enrichment signal may be observed in *pap1* on rRNA and tRNA genes (Fig. 3F). This low signal is unlikely to be specific and may be the outcome of sequencing bias caused by differences in CG content. Overall, we conclude that PAP1 is required for binding of core PEP subunits to DNA throughout the entire genome.

PAP1 contributes to efficient recruitment of core subunits to DNA

The loss of RpoB binding to DNA in the *pap1* mutant may be caused by defective recruitment of PEP core subunits to DNA in the absence of PAP1. Alternatively, it may be caused by core subunit expression and/or stability being dependent on the presence of PAP1. To distinguish between these possibilities, we performed western blots with anti-RpoB and anti-RpoC1 antibodies in the *pap1* mutant. The accumulation of RpoB and RpoC1 was reduced by approximately 50% and 40%, respectively (Fig. 4). This indicates that reduced stability of the core PEP complex may contribute to the observed loss of RpoB binding to DNA in the *pap1* mutant. However, this reduction alone cannot explain the complete genome-wide loss of RpoB binding to DNA in *pap1* (Fig. 3A). Therefore, we conclude that the loss of RpoB binding to

DNA in the *pap1* mutant is at least partially caused by defective recruitment of RpoB to DNA in the absence of PAP1. This indicates that PAP1 contributes to the efficient recruitment of core PEP subunits to DNA.

PAP7 is required for RpoB association with DNA

Requirement of PAP1 for PEP binding to DNA may indicate that peripheral subunits of PEP may be generally required to recruit PEP to DNA. This would be consistent with the observed reductions of PEP transcripts accumulation in most mutants defective in PAPs (Pfalz et al. 2006; Garcia et al. 2008; Arsova et al. 2010; Yagi et al. 2012; Gilkerson et al. 2012; Grübler et al. 2017; Yu et al. 2018; Liebers et al. 2020). To test this possibility, we performed ptChIP-seq with anti-RpoB antibody in a mutant defective in another accessory PEP subunit, PAP7. Compared to Col-0 wild-type, the *pap7* mutant lost RpoB binding to DNA throughout the entire chloroplast genome (Fig. 5A). Loss of RpoB binding to DNA was observed throughout the entire lengths of *psbA*, *psbEFLJ*, *psaA* and *rbcL* genes, including promoters and coding sequences (Fig. 5B-E). Very little residual signal was observed in the *pap7* mutant on annotated genes (Fig. 5F) and on 250 bp bins distributed throughout the entire genome (Fig. 5G).

RpoB and RpoC1 were still present in the *pap7* at levels very similar to that observed in *pap1* (Fig. 4). This indicates that PAP7 also contributes to efficient recruitment of core PEP subunits to DNA. Interestingly, accumulation of PAP1 was almost entirely lost in the *pap7* mutant (Fig. 4), indicating that the impact of PAP7 on the recruitment of core PEP subunits may be direct and/or indirect by stabilizing PAP1.

Overall, we conclude that the effect of PAP7 on PEP recruitment to DNA is very similar to that of PAP1. This is consistent with accessory subunits of PEP being required for the recruitment of core PEP subunits to DNA.

DISCUSSION

Our findings offer further support for the model where all PAPs are required for proper function of the entire PEP complex (Pfannschmidt et al. 2015), which is consistent with the structural roles of PAPs in the PEP-PAP complex (Ruedas et al. 2022; Vergara-Cruces et al. 2024; Wu et al. 2024; do Prado et al. 2024). The model's first prediction is genome-wide co-localization of core and peripheral subunits of PEP, which we demonstrated for PAP1 and RpoB (Fig 1).

Consistently, previous studies demonstrated genome-wide co-localization of PAP5/pTAC12 with RpoB (Palomar et al. 2022) and locus-specific co-localization of PAP1 with RpoA (Yagi et al. 2012). It is expected that other PAPs will also co-localize on DNA with core PEP subunits, although their genomic locations have not yet been tested.

The model predicts that all PAPs are required for PEP transcription (Pfannschmidt et al. 2015). Previous studies have shown that most PAPs are necessary for proper accumulation of PEP-transcribed RNAs (Pfalz et al. 2006; Garcia et al. 2008; Arsova et al. 2010; Yagi et al. 2012; Gilkerson et al. 2012; Grübler et al. 2017; Yu et al. 2018; Liebers et al. 2020). However, it remained unknown what step of gene expression is controlled by PAPs. We found that PAP1 and PAP7 contribute to PEP recruitment to DNA (Fig. 3 and 5). This indicates that both tested PAPs are essential for PEP transcription. While the effect of other PAPs on PEP transcription has not been experimentally tested, they are also expected to be required for PEP transcription.

PAP1 is a component of the scaffold module, where it works together with PAP3, PAP5, PAP7, PAP8, and PAP11 to stabilize the PEP core and connect it to three other functional modules within the PEP-PAP complex (Wu et al. 2024). Although the domain structure of PAP1 is consistent with DNA and/or RNA binding (Yagi et al. 2012; Pfannschmidt et al. 2015), it is unlikely to interact with RNA and cryo-EM studies provide inconsistent insights into its interactions with DNA (Vergara-Cruces et al. 2024; Wu et al. 2024; do Prado et al. 2024). Our ChIP data (Fig. 1) also do not provide conclusive evidence that PAP1 directly binds to DNA because ChIP includes formaldehyde crosslinking and may detect both direct and indirect protein-DNA interactions (Hoffman et al. 2015). The role of PAP1 as a component of the scaffold module (Wu et al. 2024) suggests that PAP1 is needed for the integrity of the entire complex and the *pap1* mutant may contain no functional PEP-PAP complex. This hypothesis is consistent with our results but remains to be tested experimentally.

PAP7 is another component of the scaffold module (Wu et al. 2024) and its requirement for PEP recruitment to DNA (Fig. 5) may be explained by its importance for the assembly of the PEP-PAP complex. However, PAP7 contains a SET-domain and has been proposed to function as a protein methyltransferase (Gao et al. 2011; Pfannschmidt et al. 2015). Although its enzymatic activity and potential substrates have not been identified, PAP7 may have a double role as a scaffold protein and as a protein methyltransferase. The exact mechanism responsible for the

albino phenotype and disrupted accumulation of PEP transcripts in the *pap7* mutant (Gao et al. 2011) as well as loss of PEP binding to DNA (Fig. 5) remains undetermined.

An important open question about PEP function is if the core PEP complex is transcriptionally competent in the absence of PAPs. This is supported by evidence that the core PEP complex isolated from *Sinapis alba* etioplasts and lacking PAPs (peak B), is transcriptionally active *in vitro* (Pfannschmidt and Link 1994). On the other hand, our data show that PAPs co-localize with core PEP subunits throughout the entire genome, including both gene promoters and coding sequences (Fig. 1), and that both tested PAPs are necessary for PEP binding to these regions (Fig. 3, Fig. 5). This suggests that in mature leaf chloroplasts, PAPs are required for both initiation and elongation of transcription. This is also consistent with recent cryo-EM studies (Ruedas et al. 2022; Vergara-Cruces et al. 2024; Wu et al. 2024; do Prado et al. 2024) and indicates that in mature leaf chloroplasts, all PAPs may be required for both initiation and elongation of transcription.

ACKNOWLEDGEMENTS

This work was supported by a grant from the National Science Foundation (MCB 1934703) to A.T.W and partially by UNAM-PAPIIT grant IA203424 to V.M.P. The sequencing data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE259283. Sequencing data presented in this study are available through a dedicated publicly available Plastid Genome Visualization Tool (Plavisto) at <http://plavisto.mcdb.lsa.umich.edu>.

FIGURE LEGENDS

Figure 1. PAP1 and RpoB bind the same DNA sequences.

A. Genome-wide patterns of PAP1 ptChIP-seq in Col-0 wild-type and *pap1* mutant, as well as RpoB ptChIP-seq in Col-0 wild-type. ptChIP-seq enrichment was calculated in 500 genomic bins and plotted throughout the entire plastid genome. Genome annotation including genomic regions, positions of annotated genes (Palomar et al. 2022), and names of selected individual genes is provided on top of the plot. Average signal from three independent biological replicates is shown. Ribbons indicate standard deviations.

- B. Regression analysis of PAP1 and RpoB binding to DNA on annotated genes. Data points are color-coded by function and show averages from three biological replicates. Error bars indicate standard deviations. The blue line represents the linear regression model. The red line represents no differences.
- C. Regression analysis of PAP1 and RpoB binding to DNA on 250-bp genomic bins. Data points are color-coded by genomic regions and show averages from three biological replicates. The blue line represents the linear regression model. The red line represents no differences.
- D-H. Overlap of PAP1 and RpoB ptChIP-seq signals on selected annotated genes. RpoB ptChIP-seq signal was calculated in 10-bp genomic bins and plotted at *psbA* (D), *psbE* (E), *rbcL* (F), *psaA* (G), and *psbB* (H) loci. Samples are color-coded as in Fig. 1A. Average signal from independent biological replicates is shown. Ribbons indicate standard deviations. Genome annotation is shown on top.

Figure 2. Western blot confirming the specificity of the anti-RpoB antibody.

Western blot was performed using chloroplast extracts as well as stromal and membrane fractions of chloroplasts from 5-day-old seedlings of Col-0 wild-type. The anti-RpoB antibody used for ptChIP-seq (PHY1239) detects a specific band at the expected size of 121kDa and non-specific bands at smaller molecular weights (marked with asterisks). Specific signal is detected in total chloroplast and enriched in the membrane fractions, as previously reported for RpoA in tobacco (Finster et al. 2013). An alternative anti-RpoB antibody, anti-LHBC1 antibody, and anti-RbcL antibody serve as controls. Positions of size marker bands labelled in kDa are shown on the left.

Figure 3. PAP1 is required for RpoB association with DNA.

- A. Genome-wide patterns of RpoB ptChIP-seq in Col-0 wild-type and the *pap1* mutant. ptChIP-seq enrichment was calculated in 500 genomic bins and plotted throughout the entire plastid genome. Genome annotation including genomic regions, positions of annotated genes (Palomar et al. 2022), and names of selected individual genes is provided on top of the plot. Average signal from three independent biological replicates is shown. Ribbons indicate standard deviations.

B-E. RpoB ptChIP-seq signals in Col-0 wild-type and *pap1* mutant on selected annotated genes.

RpoB ptChIP-seq signal was calculated in 10-bp genomic bins and plotted at *psbA* (B), *psbE* (C), *psaA* (D), and *rbcL* (E) loci. Samples are color-coded as in Fig. 3A. Average signal from three independent biological replicates is shown. Ribbons indicate standard deviations.

Genome annotation is shown on top.

F. Regression analysis of RpoB binding to DNA in Col-0 wild-type and *pap1* mutant on annotated genes. Data points are color-coded by genomic regions and show averages from three biological replicates. Error bars indicate standard deviations. The blue line represents the linear regression model. The red line represents no differences.

G. Regression analysis of RpoB binding to DNA in Col-0 wild-type and *pap1* mutant on 250-bp genomic bins. Data points are color-coded by genomic regions and show averages from three biological replicates. The blue line represents the linear regression model. The red line represents no differences.

Figure 4. RpoB and RpoC1 are still expressed in *pap1* and *pap7* mutants.

Western blot was performed using whole cell extracts from 4-day-old seedlings of Col-0 wild-type, *pap1*, and *pap7* using anti-RpoC1, anti-RpoB, and anti-PAP1 antibodies. Anti-Actin antibody was used as a loading control. Asterisk indicates a non-specific band. Positions of size marker bands labelled in kDa are shown on the left.

Figure 5. PAP7 is required for RpoB association with DNA.

A. Genome-wide patterns of RpoB ptChIP-seq in Col-0 wild-type and *pap7* mutant. ptChIP-seq enrichment was calculated in 500 genomic bins and plotted throughout the entire plastid genome. Genome annotation including genomic regions, positions of annotated genes (Palomar et al. 2022), and names of selected individual genes is provided on top of the plot. Average signal from three independent biological replicates is shown. Ribbons indicate standard deviations.

B-E. RpoB ptChIP-seq signals in Col-0 wild-type and *pap7* mutant on selected annotated genes.

RpoB ptChIP-seq signal was calculated in 10-bp genomic bins and plotted at *psbA* (B), *psbE* (C), *psaA* (D), and *rbcL* (E) loci. Samples are color-coded as in Fig. 5A. Average signal from

independent biological replicates is shown. Ribbons indicate standard deviations. Genome annotation is shown on top.

- F. Regression analysis of RpoB binding to DNA in Col-0 wild-type and *pap7* mutant on annotated genes. Data points are color-coded by genomic regions and show averages from three biological replicates. Error bars indicate standard deviations. The blue line represents the linear regression model. The red line represents no differences.
- G. Regression analysis of RpoB binding to DNA in Col-0 wild-type and *pap7* mutant on 250-bp genomic bins. Data points are color-coded by genomic regions and show averages from three biological replicates. The blue line represents the linear regression model. The red line represents no differences.

Table 1. High throughput sequencing datasets generated in this study.

REFERENCES

- Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301:653–657. <https://doi.org/10.1126/science.1086391>
- Arsova B, Hoja U, Wimmelbacher M, Greiner E, Üstün Ş, Melzer M, Petersen K, Lein W, Börnke F (2010) Plastidial Thioredoxin z Interacts with Two Fructokinase-Like Proteins in a Thiol-Dependent Manner: Evidence for an Essential Role in Chloroplast Development in *Arabidopsis* and *Nicotiana benthamiana*. *Plant Cell* 22:1498–1515. <https://doi.org/10.1105/tpc.109.071001>
- Barkan A (2011) Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. *Plant Physiol* 155:1520–1532. <https://doi.org/10.1104/pp.110.171231>
- Bowman SK, Simon MD, Deaton AM, Tolstorukov M, Borowsky ML, Kingston RE (2013) Multiplexed Illumina sequencing libraries from picogram quantities of DNA. *BMC Genomics* 14:466. <https://doi.org/10.1186/1471-2164-14-466>
- Chi W, He B, Mao J, Jiang J, Zhang L (2015) Plastid sigma factors: Their individual functions and regulation in transcription. *Biochim Biophys Acta* 1847:770–778. <https://doi.org/10.1016/j.bbabi.2015.01.001>

- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H (2021) Twelve years of SAMtools and BCFtools. *GigaScience* 10:giab008. <https://doi.org/10.1093/gigascience/giab008>
- do Prado PFV, Ahrens FM, Liebers M, Ditz N, Braun H-P, Pfannschmidt T, Hillen HS (2024) Structure of the multi-subunit chloroplast RNA polymerase. *Mol Cell* 84:910–925.e5. <https://doi.org/10.1016/j.molcel.2024.02.003>
- Finster S, Eggert E, Zoschke R, Weihe A, Schmitz-Linneweber C (2013) Light-dependent, plastome-wide association of the plastid-encoded RNA polymerase with chloroplast DNA. *Plant J Cell Mol Biol* 76:849–860. <https://doi.org/10.1111/tpj.12339>
- Gao Z-P, Yu Q-B, Zhao T-T, Ma Q, Chen G-X, Yang Z-N (2011) A Functional Component of the Transcriptionally Active Chromosome Complex, Arabidopsis pTAC14, Interacts with pTAC12/HEMERA and Regulates Plastid Gene Expression. *Plant Physiol* 157:1733–1745. <https://doi.org/10.1104/pp.111.184762>
- Garcia M, Myouga F, Takechi K, Sato H, Nabeshima K, Nagata N, Takio S, Shinozaki K, Takano H (2008) An Arabidopsis homolog of the bacterial peptidoglycan synthesis enzyme MurE has an essential role in chloroplast development. *Plant J Cell Mol Biol* 53:924–934. <https://doi.org/10.1111/j.1365-313X.2007.03379.x>
- Gilkerson J, Perez-Ruiz JM, Chory J, Callis J (2012) The plastid-localized pfkB-type carbohydrate kinases FRUCTOKINASE-LIKE 1 and 2 are essential for growth and development of Arabidopsis thaliana. *BMC Plant Biol* 12:102. <https://doi.org/10.1186/1471-2229-12-102>
- Grübler B, Merendino L, Twardziok SO, Mininno M, Alloreant G, Chevalier F, Liebers M, Blanvillain R, Mayer KFX, Lerbs-Mache S, Ravanel S, Pfannschmidt T (2017) Light and Plastid Signals Regulate Different Sets of Genes in the Albino Mutant Pap7-1. *Plant Physiol* 175:1203–1219. <https://doi.org/10.1104/pp.17.00982>
- Hoffman EA, Frey BL, Smith LM, Auble DT (2015) Formaldehyde crosslinking: a tool for the study of chromatin complexes. *J Biol Chem* 290:26404–26411. <https://doi.org/10.1074/jbc.R115.651679>
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>
- Liebers M, Gillet F-X, Israel A, Pounot K, Chambon L, Chieb M, Chevalier F, Ruedas R, Favier A, Gans P, Boeri Erba E, Cobessi D, Pfannschmidt T, Blanvillain R (2020) Nucleo-plastidic PAP8/pTAC6 couples chloroplast formation with photomorphogenesis. *EMBO J* 39:e104941. <https://doi.org/10.15252/emboj.2020104941>
- Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17:10–12. <https://doi.org/10.14806/ej.17.1.200>

- McElver J, Tzafrir I, Aux G, Rogers R, Ashby C, Smith K, Thomas C, Schetter A, Zhou Q, Cushman MA, Tossberg J, Nickle T, Levin JZ, Law M, Meinke D, Patton D (2001) Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics* 159:1751–1763. <https://doi.org/10.1093/genetics/159.4.1751>
- Myouga F, Hosoda C, Umezawa T, Iizumi H, Kuromori T, Motohashi R, Shono Y, Nagata N, Ikeuchi M, Shinozaki K (2008) A Heterocomplex of Iron Superoxide Dismutases Defends Chloroplast Nucleoids against Oxidative Stress and Is Essential for Chloroplast Development in *Arabidopsis*. *Plant Cell* 20:3148–3162. <https://doi.org/10.1105/tpc.108.061341>
- Palomar VM, Jaksich S, Fujii S, Kuciński J, Wierzbicki AT (2022) High-resolution map of plastid-encoded RNA polymerase binding patterns demonstrates a major role of transcription in chloroplast gene expression. *Plant J* 111:1139–1151. <https://doi.org/10.1111/tpj.15882>
- Pfalz J, Liere K, Kandlbinder A, Dietz K-J, Oelmüller R (2006) pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell* 18:176–197. <https://doi.org/10.1105/tpc.105.036392>
- Pfalz J, Pfannschmidt T (2013) Essential nucleoid proteins in early chloroplast development. *Trends Plant Sci* 18:186–194. <https://doi.org/10.1016/j.tplants.2012.11.003>
- Pfannschmidt T, Blanvillain R, Merendino L, Courtois F, Chevalier F, Liebers M, Grüber B, Hommel E, Lerbs-Mache S (2015) Plastid RNA polymerases: orchestration of enzymes with different evolutionary origins controls chloroplast biogenesis during the plant life cycle. *J Exp Bot* 66:6957–6973. <https://doi.org/10.1093/jxb/erv415>
- Pfannschmidt T, Link G (1994) Separation of two classes of plastid DNA-dependent RNA polymerases that are differentially expressed in mustard (*Sinapis alba* L.) seedlings. *Plant Mol Biol* 25:69–81. <https://doi.org/10.1007/BF00024199>
- Pfannschmidt T, Ogrzewalla K, Baginsky S, Sickmann A, Meyer HE, Link G (2000) The multisubunit chloroplast RNA polymerase A from mustard (*Sinapis alba* L.). Integration of a prokaryotic core into a larger complex with organelle-specific functions. *Eur J Biochem* 267:253–261. <https://doi.org/10.1046/j.1432-1327.2000.00991.x>
- Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma Oxf Engl* 26:841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- Rowley MJ, Böhmendorfer G, Wierzbicki AT (2013) Analysis of long non-coding RNAs produced by a specialized RNA polymerase in *Arabidopsis thaliana*. *Methods San Diego Calif* 63:160–169. <https://doi.org/10.1016/j.ymeth.2013.05.006>
- Ruedas R, Muthukumar SS, Kieffer-Jaquinod S, Gillet F-X, Fenel D, Effantin G, Pfannschmidt T, Couté Y, Blanvillain R, Cobessi D (2022) Three-Dimensional Envelope and Subunit

- Interactions of the Plastid-Encoded RNA Polymerase from *Sinapis alba*. *Int J Mol Sci* 23:9922. <https://doi.org/10.3390/ijms23179922>
- Steiner S, Schröter Y, Pfalz J, Pfannschmidt T (2011) Identification of Essential Subunits in the Plastid-Encoded RNA Polymerase Complex Reveals Building Blocks for Proper Plastid Development. *Plant Physiol* 157:1043–1055. <https://doi.org/10.1104/pp.111.184515>
- Suzuki JY, Jimmy Ytterberg A, Beardslee TA, Allison LA, van Wijk KJ, Maliga P (2004) Affinity purification of the tobacco plastid RNA polymerase and in vitro reconstitution of the holoenzyme. *Plant J* 40:164–172. <https://doi.org/10.1111/j.1365-313X.2004.02195.x>
- Vergara-Cruces Á, Pramanick I, Pearce D, Vogirala VK, Byrne MJ, Low JKK, Webster MW (2024) Structure of the plant plastid-encoded RNA polymerase. *Cell* 187:1145–1159.e21. <https://doi.org/10.1016/j.cell.2024.01.036>
- Wang L, Wang C, Wang Y, Niu M, Ren Y, Zhou K, Zhang H, Lin Q, Wu F, Cheng Z, Wang J, Zhang X, Guo X, Jiang L, Lei C, Wang J, Zhu S, Zhao Z, Wan J (2016) WSL3, a component of the plastid-encoded plastid RNA polymerase, is essential for early chloroplast development in rice. *Plant Mol Biol* 92:581–595. <https://doi.org/10.1007/s11103-016-0533-0>
- Wu X-X, Mu W-H, Li F, Sun S-Y, Cui C-J, Kim C, Zhou F, Zhang Y (2024) Cryo-EM structures of the plant plastid-encoded RNA polymerase. *Cell* 187:1127–1144.e21. <https://doi.org/10.1016/j.cell.2024.01.026>
- Yagi Y, Ishizaki Y, Nakahira Y, Tozawa Y, Shiina T (2012) Eukaryotic-type plastid nucleoid protein pTAC3 is essential for transcription by the bacterial-type plastid RNA polymerase. *Proc Natl Acad Sci U S A* 109:7541–7546. <https://doi.org/10.1073/pnas.1119403109>
- Yu Q-B, Zhao T-T, Ye L-S, Cheng L, Wu Y-Q, Huang C, Yang Z-N (2018) pTAC10, an S1-domain-containing component of the transcriptionally active chromosome complex, is essential for plastid gene expression in *Arabidopsis thaliana* and is phosphorylated by chloroplast-targeted casein kinase II. *Photosynth Res* 137:69–83. <https://doi.org/10.1007/s11120-018-0479-y>

STATEMENTS AND DECLARATIONS

Funding

This work was supported by a grant from the National Science Foundation (MCB 1934703) to A.T.W and partially by UNAM-PAPIIT grant IA203424 to V.M.P.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

JW, VMP and JHM performed experiments; JW, VMP and ATW analyzed the data; ATW wrote the manuscript.

Data Availability

The sequencing data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE259283. Sequencing data presented in this study are available through a dedicated publicly available Plastid Genome Visualization Tool (Plavisto) at <http://plavisto.mcdb.lsa.umich.edu>.

Figure 1

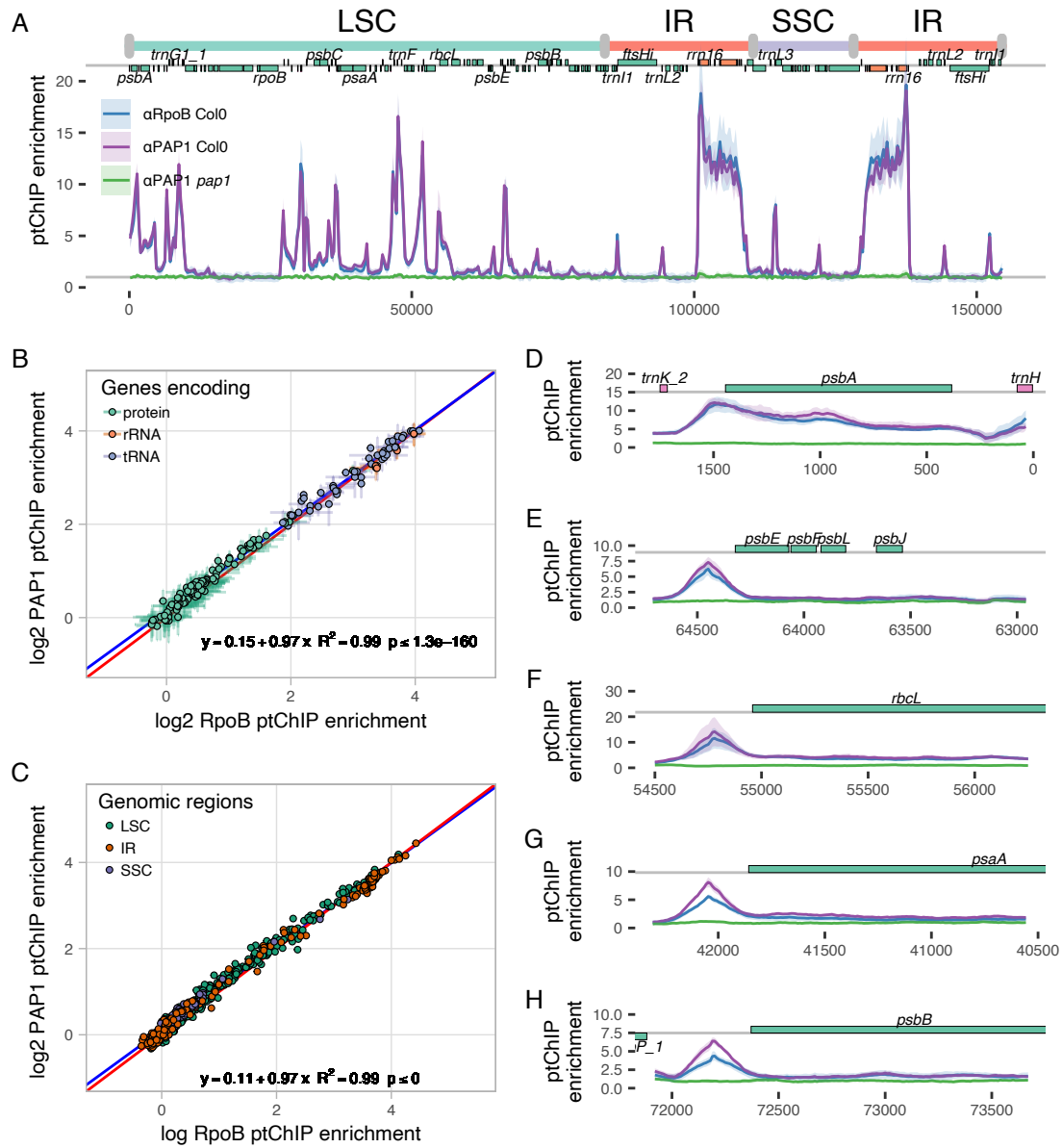
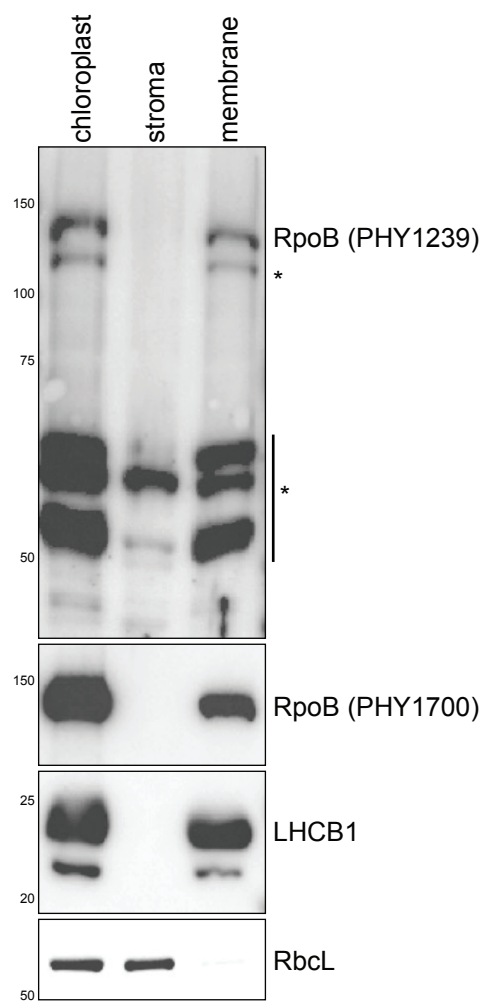


Figure 2



A

RpoB ptChIP enrichment

Col0
pap1

0 50000 100000 150000

B

ptChIP enrichment

0 5 10 15

1500 1000 500 0

C

ptChIP enrichment

0 2 4 6 8

64500 64000 63500 63000

D

ptChIP enrichment

0 2 4 6

42000 41500 41000 40500

E

ptChIP enrichment

0 4 8 12

54500 55000 55500 56000

F

log₂ pap1 RpoB ptChIP enrichment

log₂ Col-0 RpoB ptChIP enrichment

Genes encoding

- protein
- rRNA
- tRNA

G

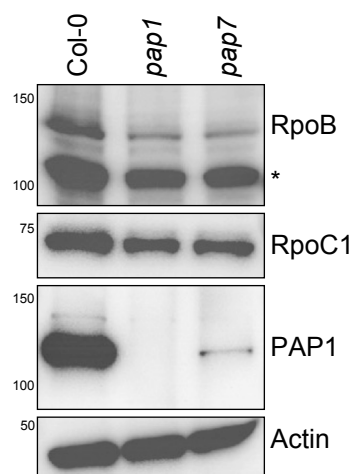
log₂ pap1 RpoB ptChIP enrichment

log Col-0 RpoB ptChIP enrichment

Genomic regions

- LSC
- IR
- SSC

Figure 4



[illegible]

Experiment	Genotype and Treatment	Repl.	Numbers of paired-end sequenced reads			Plastid Genome Coverage	Median Fragment Size	GEO accession	SRA data
			Raw	Mapped to whole genome	Mapped to plastid genome				
RpoB ChIP	Col0 (control for pap1) ChIP	1	2404550	1888259	1378496	2677.1	172	GSM8113044	SRR28106324
		2	1138407	909536	461756	896.7	196	GSM8113045	SRR28106323
		3	750595	646869	235536	457.4	156	GSM8113046	SRR28106322
	Col0 (control for pap1) input	1	936977	864265	498015	967.2	226	GSM8113062	SRR28106306
		2	921259	852732	413192	802.4	197	GSM8113063	SRR28106307
		3	1234488	1166734	609706	1184.1	164	GSM8113064	SRR28106308
	pap1 ChIP	1	976752	321579	80814	156.9	200	GSM8113047	SRR28106321
		2	611286	442617	116572	226.4	196	GSM8113048	SRR28106320
		3	908445	625126	175567	341.0	180	GSM8113049	SRR28106319
	pap1 input	1	772403	695333	344712	669.4	217	GSM8113065	SRR28106309
		2	1576591	1150485	604588	1174.1	197	GSM8113066	SRR28106299
		3	1100997	1026495	479496	931.2	155	GSM8113067	SRR28106292
	Col0 (control for pap7) ChIP	1	873695	220455	107336	208.4	163	GSM8113050	SRR28106318
		2	1236307	725191	539096	1046.9	150	GSM8113051	SRR28106317
		3	1897241	1376761	1102743	2141.6	165	GSM8113052	SRR28106316
	Col0 (control for pap7) input	1	1499182	927972	439784	854.1	190	GSM8113068	SRR28106302
		2	802892	636973	357116	693.5	191	GSM8113069	SRR28106301
		3	1217644	1063170	528814	1027.0	186	GSM8113070	SRR28106300
	pap7 ChIP	1	789424	255597	55548	107.9	192	GSM8113053	SRR28106315
		2	1931132	677442	143159	278.0	161	GSM8113054	SRR28106314
		3	764132	479414	130058	252.6	169	GSM8113055	SRR28106310
	pap7 input	1	1616121	1434207	583153	1132.5	191	GSM8113071	SRR28106289
		2	1980316	1793825	739957	1437.0	172	GSM8113072	SRR28106298
		3	1424172	1310173	654756	1271.6	179	GSM8113073	SRR28106297
PAP1 ChIP	Col0 ChIP	1	907232	747871	403277	783.2	161	GSM8113056	SRR28106311
		2	1593581	1408513	568131	1103.3	156	GSM8113057	SRR28106312
		3	765690	558636	225969	438.8	167	GSM8113058	SRR28106313
	Col0 input	1	727417	658851	256060	497.3	192	GSM8113074	SRR28106296
		2	1030683	987586	525899	1021.3	171	GSM8113075	SRR28106295
		3	871242	810989	296145	575.1	196	GSM8113076	SRR28106294
	pap1 ChIP	1	957762	639307	114661	222.7	195	GSM8113059	SRR28106303
		2	1260348	235369	52143	101.3	175	GSM8113060	SRR28106304
		3	1596702	1043009	157847	306.5	170	GSM8113061	SRR28106305
	pap1 input	1	1097835	1038704	485938	943.7	171	GSM8113077	SRR28106293
		2	1659165	1558283	764621	1484.9	187	GSM8113078	SRR28106291
		3	790264	734169	367346	713.4	179	GSM8113079	SRR28106290